矩阵运算

矩阵的加（减）法

如果 A 和 B 是 $m \times n$ 的矩阵，那么 A 和 B 的**和**定义为通过加上对应元素得到的 $m \times n$ 矩阵 $A+B$。即，

$$[A+B]_{ij} = [A]_{ij} + [B]_{ij}$$

- 矩阵 $(-A)$ 称为 A 的**加法逆元**（additive inverse），定义为将 A 中的每个元素取反所得的矩阵。
- 也就是说，如果 $A = [a_{ij}]$，那么 $-A = [-a_{ij}]$。这使得矩阵减法可以用自然的方式定义。
- 对于形状相同的两个矩阵，$A - B$ 的差定义为矩阵 $A - B = A + (-B)$，因此

$$[A-B]_{ij} = [A]_{ij} - [B]_{ij}$$

矩阵加法的性质

对于 $m \times n$ 矩阵 A、B 和 C，以下性质成立：

- **闭合性**（Closure property）：$A+B$ 仍然是一个 $m \times n$ 的矩阵。
- **结合律**（Associative property）：$(A+B)+C = A+(B+C)$。
- **交换律**（Commutative property）：$A+B = B+A$。
- **加法单位元**（Additive identity）：$m \times n$ 的零矩阵 0 具有 $A+0 = A$ 的性质。
- **加法逆元**（Additive inverse）：$m \times n$ 矩阵 $-A$ 具有 $A+(-A) = 0$ 的性质。

数乘（Scalar Multiplication）

标量 α 与矩阵 A 的乘积，记作 αA，定义为通过将 A 的每个元素乘以 α 所得到的矩阵。即

$$[\alpha A]_{ij} = \alpha [A]_{ij} \quad \text{对于每个 } i \text{ 和 } j.$$

数乘的性质

对于 $m \times n$ 矩阵 A 和 B 以及标量 α 和 β，以下性质成立：

- **闭合性**（Closure property）：αA 仍然是一个 $m \times n$ 的矩阵。
- **结合律**（Associative property）：$(\alpha\beta)A = \alpha(\beta A)$。
- **分配律 1**（Distributive property）：$\alpha(A+B) = \alpha A + \alpha B$。标量乘法对矩阵加法具有分配性。
- **分配律 2**（Distributive property）：$(\alpha+\beta)A = \alpha A + \beta A$。标量乘法对标量加法具有分配性。
- **单位元性质**（Identity property）：$1A = A$。数字1在标量乘法中是单位元素。

转置（Transpose）

$A_{m\times n}$ 的**转置**定义为 $n \times m$ 矩阵 A^T，通过将 A 中的行和列互换得到。更准确地说，如果 $A = [a_{ij}]$，那么 $[A^T]_{ij} = a_{ji}$。例如，

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}.$$

显然，对于所有矩阵 A，都有 $(A^T)^T = A$。

共轭转置（Conjugate Transpose）

对于 $A = [a_{ij}]$，**共轭矩阵**（conjugate matrix）定义为 $\overline{A} = [\overline{a_{ij}}]$，而 A 的**共轭转置**（conjugate transpose）定义为 $\overline{A}^T = \overline{A^T}$。从现在开始，$\overline{A}^T$ 将记作 A^*，因此 $[A^*]_{ij} = \overline{a_{ji}}$。例如，

$$\begin{pmatrix} 1-4i & i & 2 \\ 3 & 2+i & 0 \end{pmatrix}^* = \begin{pmatrix} 1+4i & 3 \\ -i & 2-i \\ 2 & 0 \end{pmatrix}.$$

- 对于所有矩阵，$(A^*)^* = A$，并且当 A 仅包含实数元素时，$A^* = A^T$。
- 有时矩阵 A^* 被称为 A 的**伴随矩阵**（adjoint of A）。

转置的性质（Properties of the Transpose）

如果 A 和 B 是相同形状的两个矩阵，并且 α 是一个标量，那么以下各个陈述均为真：

- $(A+B)^T = A^T + B^T$ 并且 $(A+B)^* = A^* + B^*$。
- $(\alpha A)^T = \alpha A^T$ 并且 $(\alpha A)^* = \overline{\alpha} A^*$。

对称性（Symmetries）（不重要）

设 $A = [a_{ij}]$ 为一个方阵。

- 当 $A = A^T$ 时，称 A 为**对称矩阵**（symmetric matrix），即 $a_{ij} = a_{ji}$。
- 当 $A = -A^T$ 时，称 A 为**反对称矩阵**（skew-symmetric matrix），即 $a_{ij} = -a_{ji}$。
- 当 $A = A^*$ 时，称 A 为**厄米矩阵**（hermitian matrix），即 $a_{ij} = \overline{a_{ji}}$。这是对称性的复数类比。
- 当 $A = -A^*$ 时，称 A 为**反厄米矩阵**（skew-hermitian matrix），即 $a_{ij} = -\overline{a_{ji}}$。这是反对称性的复数类比。

矩阵乘法（Matrix Multiplication）

- 如果矩阵 A 和 B 满足 A 的列数等于 B 的行数（即 A 是 $m \times p$ 矩阵，B 是 $p \times n$ 矩阵），则称矩阵 A 和 B 在顺序 AB 上是**可乘的**（conformable）。
- 对于可乘矩阵 $A_{m\times p} = [a_{ij}]$ 和 $B_{p\times n} = [b_{ij}]$，**矩阵积**（matrix product）AB 定义为一个 $m \times n$ 的矩阵，其第 (i,j) 个元素是 A 的第 i 行与 B 的第 j 列的内积。即，

$$[AB]_{ij} = A_{i*}B_{*j} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ip}b_{pj} = \sum_{k=1}^{p} a_{ik}b_{kj}.$$

- 如果 A 和 B 不可乘（即 A 是 $m \times p$ 矩阵，而 B 是 $q \times n$ 矩阵，且 $p \neq q$），则不定义 AB 的积。

矩阵乘法不是交换律的（Matrix Multiplication Is Not Commutative）

矩阵乘法是一个**非交换运算**（noncommutative operation），即即使两个乘积都存在并且矩阵形状相同，也可能有 $AB \neq BA$。

(需要理解的，后面课件中经常出现的表达方式)在矩阵乘积中，有多种方法可以表示单个行和列。例如，AB 的第 i 行为：

$$[AB]_{i*} = [A_{i*}B_{*1} \mid A_{i*}B_{*2} \mid \cdots \mid A_{i*}B_{*n}] = A_{i*}B$$

$$= (a_{i1}\, a_{i2}\, \cdots\, a_{ip}) \begin{pmatrix} B_{1*} \\ B_{2*} \\ \vdots \\ B_{p*} \end{pmatrix} = a_{i1}B_{1*} + a_{i2}B_{2*} + \cdots + a_{ip}B_{p*}.$$

如下面所示，对于单个列也有类似的表示方式。

乘积的行与列（Rows and Columns of a Product）

假设 $A = [a_{ij}]$ 是 $m \times p$ 矩阵，$B = [b_{ij}]$ 是 $p \times n$ 矩阵。

- $[AB]_{i*} = A_{i*}B$（AB 的第 i 行）=（A 的第 i 行）×B。
- $[AB]_{*j} = AB_{*j}$（AB 的第 j 列）= A ×（B 的第 j 列）。
- $[AB]_{i*} = a_{i1}B_{1*} + a_{i2}B_{2*} + \cdots + a_{ip}B_{p*} = \sum_{k=1}^{p} a_{ik}B_{k*}$。
- $[AB]_{*j} = A_{*1}b_{1j} + A_{*2}b_{2j} + \cdots + A_{*p}b_{pj} = \sum_{k=1}^{p} A_{*k}b_{kj}$。

最后两个等式说明 AB 的行是 B 的行的组合，而 AB 的列是 A 的列的组合。

线性的概念是我们的主题的基本主题。在初等数学中，“线性函数”是指直线。在高等数学中，线性意味着更普遍的东西

线性函数（Linear Functions）

假设 D 和 R 是具有加法运算和标量乘法运算的集合——即，标量与集合成员之间的乘法运算。一个将 D 中的点映射到 R 中的点的函数 f 称为**线性函数**，当且仅当 f 满足以下条件：

$$f(x+y)=f(x)+f(y)$$

和

$$f(\alpha x)=\alpha f(x)$$

对于 D 中的每个 x 和 y 以及所有标量 α。这两个条件可以合并表示为：当且仅当

$$f(\alpha x+y)=\alpha f(x)+f(y)$$

对于所有标量 α 和 D 中的 x,y 时，f 是一个线性函数。

- 最简单的线性函数之一是 $f(x)=\alpha x$，其在 $\mathbb{R}^2$ 中的图像是一条经过原点的直线。
- 然而，$f(x)=\alpha x+\beta$ 并不符合线性函数的定义——它是一个通过常数 β 偏移的线性函数。
- 线性函数的偏移被称为**仿射函数**（affine functions）。
- 在 $\mathbb{R}^3$ 中，曲面 $f(x_1,x_2)=\alpha_1x_1+\alpha_2x_2$ 是一个经过原点的平面，并且容易验证 f 是一个线性函数。
- 当 $\beta\neq 0$ 时，$f(x_1,x_2)=\alpha_1x_1+\alpha_2x_2+\beta$ 不再是线性函数，而是一个仿射函数。
- 虽然我们无法通过肉眼观察到更高维度的情况，但建议认为形式为

  $$\begin{gathered}f(x_1,x_2,\cdots,x_n)=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_nx_n\\ f(\theta_1x_1,\theta_2x_2,\cdots,\theta_nx_n)=\theta_1\alpha_1x_1+\theta_2\alpha_2x_2+\cdots+\theta_n\alpha_nx_n\end{gathered}$$

  的函数是一般的线性函数。
- 对于标量 α_j 和矩阵 X_j，表达式

  $$\alpha_1X_1+\alpha_2X_2+\cdots+\alpha_nX_n=\sum_{j=1}^{n}\alpha_jX_j$$

  被称为 X_j 的**线性组合**（linear combination）。

单位矩阵（Identity Matrix）

一个 $n\times n$ 矩阵，其主对角线上的元素为1，其余元素为0，定义为

$$I_n=\begin{pmatrix}1&0&\cdots&0\\0&1&\cdots&0\\\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1\end{pmatrix}$$

称为**阶数为 n 的单位矩阵**（identity matrix of order n）。对于每个 $m\times n$ 矩阵 A，

$$AI_n=A\quad\text{和}\quad I_mA=A.$$

当单位矩阵的大小在上下文中是显而易见时，通常会省略 I_n 上的下标。

通过将一个或两个因子划分为**子矩阵**（submatrices）来执行两个矩阵之间的乘法（即包含在另一个矩阵中的矩阵）可以是一种有用的技术。

分块矩阵乘法（Block Matrix Multiplication）

假设矩阵 A 和 B 被划分为子矩阵（通常称为块），如下所示：

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{sr} \end{pmatrix}, \quad B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1t} \\ B_{21} & B_{22} & \cdots & B_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ B_{r1} & B_{r2} & \cdots & B_{rt} \end{pmatrix}.$$

如果每一对 (A_{ik}, B_{kj}) 是可乘的，则称 A 和 B 为**可配对划分**（conformably partitioned）。对于这样的矩阵，乘积 AB 是通过按普通矩阵乘法中的方式组合各个块得到的。即，AB 中的 (i, j) 块为

$$A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ir}B_{rj}.$$

块乘法在待乘矩阵中存在某些模式时尤其有用。考虑以下分块矩阵：

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} C & I \\ I & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 2 & 1 & 2 \\ 3 & 4 & 3 & 4 \end{pmatrix} = \begin{pmatrix} I & 0 \\ C & C \end{pmatrix},$$

其中

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{和} \quad C = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

使用块乘法，可以轻松计算出乘积 AB 为

$$AB = \begin{pmatrix} C & I \\ I & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ C & C \end{pmatrix} = \begin{pmatrix} 2C & C \\ I & 0 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 1 & 2 \\ 6 & 8 & 3 & 4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

矩阵的逆（Matrix Inversion）-奇异矩阵、非奇异矩阵

- 如果 α 是一个非零标量，则对于每个数 β，方程 $\alpha x = \beta$ 有唯一解 $x = \alpha^{-1}\beta$。
- 性质 $\alpha\alpha^{-1} = 1$ 和 $\alpha^{-1}\alpha = 1$ 是关键因素。
- 我们希望用与求解标量方程相同的方式求解矩阵方程。

对于给定的方阵 $A_{n\times n}$，满足条件

$$AB = I_n \quad \text{和} \quad BA = I_n$$

的矩阵 $B_{n\times n}$ 称为 A 的**逆矩阵**（inverse of A），记作 $B = A^{-1}$。并不是所有方阵都是可逆的——零矩阵是一个简单的例子，但还有许多非零矩阵是不可逆的。一个可逆矩阵称为**非奇异矩阵**（nonsingular），而没有逆的方阵称为**奇异矩阵**（singular matrix）。

- **请注意，矩阵求逆仅对方阵定义。**
- 尽管并非所有矩阵都是可逆的，但当逆存在时，它是唯一的。
- 由于矩阵求逆的定义类似于标量求逆，因此我们有以下结论：
    i. 如果 A 是一个非奇异矩阵，那么在矩阵方程 $A_{n\times n}X_{n\times p} = B_{n\times p}$ 中，矩阵 X 有唯一解，其解为

    $$X = A^{-1}B.$$

    ii. 一个包含 n 个未知数的 n 元线性方程组可以写成一个矩阵方程 $A_{n\times n}x_{n\times 1} = b_{n\times 1}$，因此当 A 是非奇异矩阵时，该系统有唯一一解，解为

    $$x = A^{-1}b.$$

逆矩阵的存在性（Existence of an Inverse）

对于一个 $n \times n$ 矩阵 A，以下陈述是等价的：

- A^{-1} 存在（即 A 是非奇异的）。

- $\mathrm{rank}(A) = n$。
- A 经过高斯-约当消元可以化为单位矩阵 I。
- $Ax = 0$ 推出 $x = 0$。

矩阵求逆的定义表明，为了计算 A^{-1}，需要求解矩阵方程 $AX = I$ 和 $XA = I$。

如果 A 和 X 是方阵，

$$AX = I \implies XA = I.$$

换句话说，如果 A 和 X 是方阵且 $AX = I$，那么 $X = A^{-1}$。

尽管我们通常尽量避免计算矩阵的逆，但在某些情况下必须求出逆矩阵。

计算逆矩阵（Computing an Inverse)

高斯-约当消元法可用于通过如下化简来求 A 的逆：

$$[A \mid I] \xrightarrow{\text{Gauss-Jordan}} [I \mid A^{-1}].$$

这种化简唯一可能失败的情况是，在增广矩阵的左侧出现全零的行，这仅在 A 是奇异矩阵时发生。

Exp：

问题：如果可能，求 $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$ 的逆矩阵。

解答：

$$[A \mid I] = \begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ 1 & 2 & 2 & | & 0 & 1 & 0 \\ 1 & 2 & 3 & | & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ 0 & 1 & 1 & | & -1 & 1 & 0 \\ 0 & 1 & 2 & | & -1 & 0 & 1 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 0 & 0 & | & 2 & -1 & 0 \\ 0 & 1 & 0 & | & -1 & 2 & -1 \\ 0 & 0 & 1 & | & 0 & -1 & 1 \end{pmatrix}.$$

因此，该矩阵是非奇异的，且

$$A^{-1} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

如果我们希望验证此结果，只需检查 $AA^{-1} = I$。

- [ ] 矩阵求逆的复杂度

和的逆矩阵与敏感性（Inverses of Sums and Sensitivity)

- 逆序律使得求积的逆矩阵比较容易处理，但求和的逆矩阵则要困难得多。
- 首先，$(A + B)^{-1}$ 可能不存在，即使 A^{-1} 和 B^{-1} 都存在。（（AB）$^{-1} = B^{-1}A^{-1}$）
- 而且，即使 $(A + B)^{-1}$ 存在，在少数例外情况下，$(A + B)^{-1} \neq A^{-1} + B^{-1}$。
- 对于 $(A + B)^{-1}$ 没有通用的有用公式，但对于一些特定的和可以得出一些结论。
- 最容易求逆的和之一是 $I + cd^T$，其中 c 和 d 是 $n \times 1$ 的非零列，满足 $1 + d^T c \neq 0$。
- 通过直接乘法可以验证：

$$(I + cd^T)^{-1} = I - \frac{cd^T}{1 + d^T c}.$$

Sherman–Morrison 公式

- 如果 $A_{n\times n}$ 是非奇异的，并且 c 和 d 是 $n \times 1$ 的列，满足 $1 + d^T A^{-1} c \neq 0$，那么矩阵 $A + cd^T$ 是非奇异的，且

$$(A+cd^T)^{-1} = A^{-1} - \frac{A^{-1}cd^TA^{-1}}{1+d^TA^{-1}c}.$$

- Sherman–Morrison–Woodbury 公式是这一结果的推广。如果 C 和 D 是 $n \times k$ 的矩阵，且 $(I+D^TA^{-1}C)^{-1}$ 存在，那么

$$(A+CD^T)^{-1} = A^{-1} - A^{-1}C(I+D^TA^{-1}C)^{-1}D^TA^{-1}.$$

- **Sherman–Morrison–Woodbury 公式**可以通过直接乘法验证。
- **应用场景**:假设 A^{-1} 已知于之前的计算，但 A 中的某个元素需要更改或更新——我们需要将 α 加到 a_{ij} 上。
- Sherman–Morrison 公式展示了如何通过已知的 A^{-1} 信息更新以获得新的逆矩阵。

问题：已知矩阵 A 和 A^{-1} 如下。通过对 a_{21} 加上1来更新 A，然后使用 Sherman–Morrison 公式来更新 A^{-1}：

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} \quad \text{和} \quad A^{-1} = \begin{pmatrix} 3 & -2 \\ -1 & 1 \end{pmatrix}.$$

解答：更新后的矩阵为

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = A + e_2e_1^T.$$

应用 Sherman–Morrison 公式得到更新的逆矩阵：

$$B^{-1} = A^{-1} - \frac{A^{-1}e_2e_1^TA^{-1}}{1+e_1^TA^{-1}e_2} = A^{-1} - \frac{[A]^{-1}_{*2}[A]^{-1}_{1*}}{1+[A^{-1}]_{12}}.$$

计算得到：

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

因此，

$$B^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}.$$

另一个常需要求逆的和是 $I - A$，但需要小心，因为 $(I-A)^{-1}$ 不一定总是存在。
然而，当 A 中的元素足够小时，求逆是安全的。
实际使用是$(A+B)^{-1}$

Neumann 级数（Neumann Series）

如果 $\lim_{n\to\infty} A^n = 0$，则 $I - A$ 是非奇异的，并且

$$(I-A)^{-1} = I + A + A^2 + \cdots = \sum_{k=0}^{\infty} A^k.$$

这称为 **Neumann 级数**。当 A 的元素具有小量级时，它提供了 $(I-A)^{-1}$ 的近似。例如，一阶近似为 $(I-A)^{-1} \approx I + A$。

- 虽然一般情况下不存在 $(A+B)^{-1}$ 的有用公式，但 Neumann 级数允许我们在 B 的元素相对于 A 较小时，或反之，得出一些结论。
- 例如，如果 A^{-1} 存在，并且 B 中的元素足够小以保证 $\lim_{n\to\infty}(A^{-1}B)^n = 0$，则可以应用 Neumann 级数。

$$(A+B)^{-1} = \left(A\left(I - [-A^{-1}B]\right)\right)^{-1} = (I - [-A^{-1}B])^{-1}A^{-1}$$

$$= \left(\sum_{k=0}^{\infty} [-A^{-1}B]^k\right) A^{-1}$$

一阶近似为：

$$(A+B)^{-1} \approx A^{-1} - A^{-1}BA^{-1}.$$

- 因此，如果 A 被一个小矩阵 B 扰动，可能由于测量误差不精确或舍入误差导致，则 A^{-1} 的变化大约为 $A^{-1}BA^{-1}$。
- 换句话说，一个小的扰动（或误差）B 的影响会因与 A^{-1} 的乘法（在两侧）而被放大。

因此，如果 A^{-1} 有较大的元素，A 中的小扰动（或误差）可能会在得到的逆矩阵中产生较大的扰动（或误差）。

灵敏度与条件数（Sensitivity and Conditioning）

- 当非奇异矩阵 A 的一个小的相对变化会在 A^{-1} 中引起大的相对变化时，称 A 是**病态的**（ill conditioned）。病态程度通过一个**条件数** $\kappa = \|A\|\|A^{-1}\|$ 来衡量，其中 $\| * \|$ 是矩阵范数。
- 线性方程 $Ax = b$ 的解的灵敏度取决于 A 是病态矩阵的程度。

如果非奇异系统 $Ax = b$ 稍微扰动得到系统 $(A + B)\tilde{x} = b$，则 $x = A^{-1}b$ 且 $\tilde{x} = (A + B)^{-1}b$，因此（如果要应用一阶诺曼级数）

$$x - \tilde{x} = A^{-1}b - (A + B)^{-1}b \approx A^{-1}Bx.$$

结果有：

$$\|x - \tilde{x}\| \lesssim \|A^{-1}\|\|B\|\|x\|,$$

$$\frac{\|x - \tilde{x}\|}{\|x\|} \lesssim \|A^{-1}\|\|B\| = \kappa\left(\frac{\|B\|}{\|A\|}\right).$$

notes：：干嘛要这样来定义条件数。。。纯纯的设置专业壁垒

- **经验法则**：如果使用带部分选主元的高斯消元法求解非奇异系统 $Ax = b$，并使用 t 位数的浮点运算，那么在假设不存在其他误差来源的情况下，可以认为当条件数 κ 的数量级为 10^p 时，计算的解预计具有至少 $t - p$ 位有效数字的精度。

初等矩阵与等价性（Elementary Matrices and Equivalence）

- 数学中的一个常见主题是将复杂的对象分解为更基本的成分。
- 将大的多项式分解为较小的多项式的乘积。
- 本节旨在为矩阵代数中的类似思想奠定基础，通过考虑如何将一个一般矩阵分解为多个更基本的矩阵的乘积。
- 形如 $I - uv^T$ 的矩阵，其中 u 和 v 是 $n \times 1$ 的列，且满足 $v^Tu \neq 1$，称为**初等矩阵**（elementary matrices）。
- 所有这样的矩阵都是非奇异的，且

$$(I - uv^T)^{-1} = I - \frac{uv^T}{v^Tu - 1}.$$

- [ ] 证明等式

- 请注意，初等矩阵的逆也是初等矩阵。
- 我们主要关注与以下三种初等行（或列）操作相关的初等矩阵：
  - **类型 I**：交换第 i 行（列）和第 j 行（列）。
  - **类型 II**：将第 i 行（列）乘以 $\alpha \neq 0$。
  - **类型 III**：将第 i 行（列）的某个倍数加到第 j 行（列）。
- 类型 I、II 或 III 的初等矩阵是通过在单位矩阵上执行类型 I、II 或 III 的初等操作得到的。
- 例如：

$$E_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix}.$$

- $E_1 = I - uu^T$，其中 $u = e_1 - e_2$。
- $E_2 = I - (1 - \alpha)e_2e_2^T$ 且 $E_3 = I + \alpha e_3e_1^T$。

初等矩阵的性质（Properties of Elementary Matrices）

- 当作为**左乘**运算符时，类型 I、II 或 III 的初等矩阵执行相应的行操作。
- 当作为**右乘**运算符时，类型 I、II 或 III 的初等矩阵执行相应的列操作。

用于将 $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 3 & 6 & 13 \end{pmatrix}$ 化简为 E_A 的行操作序列如下：

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ 3 & 6 & 13 \end{pmatrix} \xrightarrow{R_2 - 2R_1} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \\ 3 & 6 & 13 \end{pmatrix} \xrightarrow{R_3 - 3R_1} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{交换} R_2 \text{和} R_3} \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 - 4R_2} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = E_A$$

这种化简可以通过与相应的初等矩阵的左乘序列来实现，如下所示：

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

定理：A 是非奇异矩阵当且仅当 A 是类型 I、II 或 III 的初等矩阵的乘积。

证明：如果 A 是非奇异的，那么通过高斯-约当方法可以通过行操作将 A 化为单位矩阵 I。设 $G_1, G_2, \ldots, G_k$ 是对应于所使用的初等行操作的初等矩阵序列，则

$$G_k \cdots G_2 G_1 A = I$$

或等价地，

$$A = G_1^{-1} G_2^{-1} \cdots G_k^{-1}.$$

由于初等矩阵的逆也是同类型的初等矩阵，这证明了 A 是类型 I、II 或 III 的初等矩阵的乘积。反之，如果 $A = E_1 E_2 \cdots E_k$ 是初等矩阵的乘积，那么 A 必须是非奇异的，因为 E_i 是非奇异矩阵，而非奇异矩阵的乘积也是非奇异的。

等价性（Equivalence）

- 当 B 可以通过对 A 进行初等行和列操作得到时，我们记作 $A \sim B$，并称 A 和 B 是**等价矩阵**（equivalent matrices）。由于初等行和列操作分别对应于左乘和右乘初等矩阵，因此我们可以说：

$$A \sim B \iff PAQ = B$$

  对于非奇异矩阵 P 和 Q。
- 当矩阵 B 可以通过对矩阵 A 执行一系列的初等**行**操作得到时，我们记作 $A \stackrel{\text{row}}{\sim} B$，并称 A 和 B 是**行等价**的（row equivalent）。换句话说：

$$A \stackrel{\text{row}}{\sim} B \iff PA = B$$

  其中 P 是非奇异矩阵。
- 当矩阵 B 可以通过对矩阵 A 执行一系列的初等**列**操作得到时，我们记作 $A \stackrel{\text{col}}{\sim} B$，并称 A 和 B 是**列等价**的（column equivalent）。换句话说：

$$A \stackrel{\text{col}}{\sim} B \iff AQ = B$$

  其中 Q 是非奇异矩阵。
- 如果可以通过初等行和列操作从 A 到达 B，显然也可以从 B 回到 A，因为初等操作是可逆的。
- 因此，讨论矩阵之间的等价关系时无需考虑顺序：$A \sim B \iff B \sim A$。
- 每种类型的等价关系都是传递的，即：

$$A \sim B \quad \text{且} \quad B \sim C \implies A \sim C.$$

秩标准形（Rank Normal Form）

如果 A 是一个 $m \times n$ 的矩阵，且 $\operatorname{rank}(A) = r$，则

$$A \sim N_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}.$$

N_r 称为 A 的**秩标准形**，它是通过行和列操作对 A 进行完全化简的最终结果。

- 给定矩阵 A 和 B，如何判断 $A \sim B$ 是否成立？
  - 我们可以通过尝试将 A 通过初等操作化简为 B 来进行试验。
  - $A \sim B$ 当且仅当 $\mathrm{rank}(A) = \mathrm{rank}(B)$。
  - **推论**：非奇异矩阵的乘法不会改变秩。因为非奇异矩阵可以由初等矩阵组成。
- 转置不改变秩——即，对于所有 $m \times n$ 的矩阵，有

$$\mathrm{rank}(A) = \mathrm{rank}(A^T) \quad \text{和} \quad \mathrm{rank}(A) = \mathrm{rank}(A^*).$$

LU 分解（LU Factorization）

- 现在，我们回到使用高斯消元法和回代法求解非奇异线性方程组。这一次，目标是在矩阵的上下文中描述和理解该过程。
- 如果 $Ax = b$ 是一个非奇异系统，那么高斯消元法的目标是通过初等行操作将 A 化为上三角矩阵。
- **如果没有遇到零主元，则不需要行交换。行仅交换破坏LU分解**
- 这种化简可以仅使用类型 III 的初等行操作来实现。
  例如，考虑将矩阵化简：

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix} \xrightarrow{R_2 - 2R_1} \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 12 & 16 \end{pmatrix} \xrightarrow{R_3 - 3R_1} \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 12 & 16 \end{pmatrix} \xrightarrow{R_3 - 4R_2} \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} = U.$$

我们在上一节中了解到，这些类型 III 操作可以通过与相应初等矩阵 G_i 的左乘来实现。这些 G_i 的乘积为：

$$\mathbf{G}_3\mathbf{G}_2\mathbf{G}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & -4 & 1 \end{pmatrix}$$

换句话说，$G_3G_2G_1A = U$，因此 $A = G_1^{-1}G_2^{-1}G_3^{-1}U = LU$，其中 L 是下三角矩阵：

$$L = G_1^{-1}G_2^{-1}G_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}.$$

因此，$A = LU$ 是一个下三角矩阵 L 与一个上三角矩阵 U 的乘积。这自然被称为 A 的 **LU 分解**。

观察到U 是高斯消去法的最终结果，且其对角线上的位置为主元，同时 L 的对角线元素为 1。

- **L** 具有显著的性质：在其对角线下方，每个元素 l_{ij} 恰好是消去法中用于消去 (i, j) 位置的倍数。(实际上不重要，不要死记这个性质。)

$$LU\text{分解}$$

如果 A 是一个 $n \times n$ 矩阵，且在应用 Type III 操作的高斯消去法时未遇到零主元，那么 A 可以被分解为乘积 $A = LU$，其中满足以下条件：

- L 是下三角矩阵，U 是上三角矩阵。
- $l_{ii} = 1$ 且 $u_{ii} \neq 0$ 对于每个 $i = 1, 2, \ldots, n$。
- 在 L 的对角线以下，元素 l_{ij} 是第 j 行的倍数，该倍数从第 i 行中减去，以在高斯消去法中消去 (i, j) 位置的元素。
- U 是高斯消去法应用于 A 的最终结果。

- A 的分解 $A = LU$ 被称为 **LU 分解**，矩阵 L 和 U 被称为 A 的 LU 因子。

一旦为一个非奇异矩阵 $A_{n\times n}$ 获得了 LU 因子，就可以相对容易地解线性方程组 $Ax = b$。通过将 $Ax = b$ 重写为

$$L(Ux) = b$$

并设 $y = Ux$，我们可以看到 $Ax = b$ 等价于两个三角方程组

$$Ly = b \quad \text{和} \quad Ux = y.$$

首先，通过**前向替换**解下三角系统 $Ly = b$ 得到 y。也就是说，如果

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \ell_{21} & 1 & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix}$$

则

$$y_1 = b_1, \quad y_2 = b_2 - \ell_{21}y_1, \quad y_3 = b_3 - \ell_{31}y_1 - \ell_{32}y_2$$

前向替换算法可以更简洁地表示为

$$y_1 = b_1 \quad 和 \quad y_i = b_i - \sum_{k=1}^{i-1} \ell_{ik}y_k \quad 对于 \quad i = 2, 3, \ldots, n.$$

在求得 y 之后，利用标准的回代方法从 $x_n = y_n/u_{nn}$ 开始解上三角系统 $Ux = y$，并设置

$$x_i = \frac{1}{u_{ii}} \left(y_i - \sum_{k=i+1}^{n} u_{ik}x_k \right) \quad 对于 \quad i = n-1, n-2, \ldots, 1.$$

问题：使用 A 的 LU 分解来解方程 $Ax = b$，其中

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 7 & 7 \\ 6 & 18 & 22 \end{pmatrix}, \quad b = \begin{pmatrix} 12 \\ 24 \\ 12 \end{pmatrix}.$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix}.$$

策略是设 $Ux = y$ 并通过解两个三角方程组来求解 $Ax = L(Ux) = b$:

$$Ly = b \quad 和 \quad Ux = y.$$

首先，通过**前向替换**解下三角系统 $Ly = b$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 12 \\ 24 \\ 12 \end{pmatrix} \Rightarrow \begin{cases} y_1 = 12, \\ y_2 = 24 - 2y_1 = 0, \\ y_3 = 12 - 3y_1 - 4y_2 = -24. \end{cases}$$

然后，使用回代法解上三角系统 $Ux = y$:

$$\begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 12 \\ 0 \\ -24 \end{pmatrix} \Rightarrow \begin{cases} x_1 = \frac{(12-2x_2-2x_3)}{2} = 6, \\ x_2 = \frac{(0-3x_3)}{3} = 6, \\ x_3 = -24/4 = -6. \end{cases}$$

- 如果只需要解一个方程组 $Ax = b$，那么将增广矩阵 $[A|b]$ 化为行阶梯形式的方法与 LU 分解方法之间没有显著差异。但是，假设之后有必要解其他方程组 $Ax = \tilde{b}$，其系数矩阵相同，但右侧不同的情形，这种情况在应用工作中经常出现。
  如果在原方程组求解时已经计算并保存了 A 的 LU 因子，则不需要重新计算这些因子，因此后续方程组 $Ax = \tilde{b}$ 的求解相对成本较低。

总结

- 要使用 LU 分解 $A = LU$ 来解非奇异方程组 $Ax = b$，首先使用前向替换算法解 $Ly = b$ 得到 y，然后使用回代法解 $Ux = y$ 得到 x。
- 该方法的优势在于，一旦 A 的 LU 因子计算完毕，任何其他线性方程组 $Ax = \tilde{b}$ 都可以仅使用 n^2 次乘法/除法和 $n^2 - n$ 次加法/减法来解出。

并非所有非奇异矩阵都拥有 LU 分解。

显然不存在非零的 u_{11} 值使得成立。

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l_{21} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} \\ 0 & u_{22} \end{pmatrix}$$

问题在于 $(1,1)$ 位置的零主元。

一个非奇异矩阵 A 存在 LU 分解当且仅当在使用 Type III 操作进行行化简至上三角形形式的过程中，不会出现零主元。

还有另一种有趣的方式来描述 LU 因子的存在性。此描述是基于矩阵 A 的主子矩阵的，定义为从矩阵 A 的左上角提取的那些子矩阵。即

$$A_1 = (a_{11}), \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \ldots, A_k = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}, \ldots$$

LU 因子的存在性

以下每个陈述都等价于一个非奇异矩阵 $A_{n\times n}$ 拥有 LU 分解的条件：

- 在使用 Type III 操作进行行化简至上三角形形式的过程中不会出现零主元。(充分不必要条件，如果是用TYPE III 操作，可以确保没有问题)
- 每个主子矩阵 A_k 是非奇异的。

到目前为止，我们避免了处理行交换，因为如果需要通过行交换来去除零主元，那么就不可能进行 LU 分解。
然而，我们知道在实际计算中，部分主元法的行交换是必要的.因此，即使没有出现零主元，通常我们也需要以某种方式考虑行交换。
当允许行交换时，只要原始矩阵 A 是非奇异的，就可以始终避免零主元的出现。

因此，我们可以得出结论：对于每一个非奇异矩阵 A，都存在一个置换矩阵 P（一系列基本行交换矩阵的乘积），使得 PA 具有 LU 分解。

这意味着我们可以像在没有行交换的情况下那样进行操作，逐步覆盖原始包含 A 的数组，每个乘数替换它所消去的位置。每当发生行交换时，相应的乘数也会被正确地交换。

置换矩阵 P 只是所使用的各种交换的累计记录，矩阵 P 中的信息可以通过一个简单的技术来处理，下面的例子中展示了这种方法。

问题：对矩阵进行部分选主元操作

$$A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 2 & 1 \\ -3 & -1 & 1 & -4 \end{pmatrix}$$

并确定 LU 分解 $PA = LU$，其中 P 是相应的置换矩阵。

$$[A|p] = \begin{pmatrix} 1 & 2 & -3 & 4 & | & 1 \\ 4 & 8 & 12 & -8 & | & 2 \\ 2 & 3 & 2 & 1 & | & 3 \\ -3 & -1 & 1 & -4 & | & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & 8 & 12 & -8 & | & 2 \\ 1 & 2 & -3 & 4 & | & 1 \\ 2 & 3 & 2 & 1 & | & 3 \\ -3 & -1 & 1 & -4 & | & 4 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 4 & 8 & 12 & -8 & | & 2 \\ 1/4 & 0 & -6 & 6 & | & 1 \\ 1/2 & -1 & -4 & 5 & | & 3 \\ -3/4 & 5 & 10 & -10 & | & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & 8 & 12 & -8 & | & 2 \\ -3/4 & 5 & 10 & -10 & | & 4 \\ 1/2 & -1 & -4 & 5 & | & 3 \\ 1/4 & 0 & -6 & 6 & | & 1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 4 & 8 & 12 & -8 & | & 2 \\ -3/4 & 5 & 10 & -10 & | & 4 \\ 1/4 & 0 & -6 & 6 & | & 1 \\ 1/2 & -1/5 & -5 & 3 & | & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 4 & 8 & 12 & -8 & | & 2 \\ -3/4 & 5 & 10 & -10 & | & 4 \\ 1/4 & 0 & -6 & 6 & | & 1 \\ 1/2 & -1/5 & 1/3 & 1 & | & 3 \end{pmatrix}.$$

因此，

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -3/4 & 1 & 0 & 0 \\ 1/4 & 0 & 1 & 0 \\ 1/2 & -1/5 & 1/3 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 4 & 8 & 12 & -8 \\ 0 & 5 & 10 & -10 \\ 0 & 0 & -6 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad P = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

带行交换的 LU 分解

- 对于每一个非奇异矩阵 A，存在一个置换矩阵 P，使得 PA 拥有 LU 分解 $PA = LU$。
- 为了计算 L、U 和 P，逐步覆盖最初包含 A 的数组。将每个被消去的元素替换为执行消去操作时使用的乘数。当执行部分选主元的行交换时，数组中的乘数将自动按正确的方式交换。
- 尽管很少需要完整的置换矩阵 P，但它可以通过根据所使用的各种交换对单位矩阵 I 的行进行置换来构造。这些交换可以累积在一个“置换计数列” p 中，该列最初按自然顺序排列（$1, 2, \ldots, n$）。
- 为了使用带部分选主元的 LU 分解求解非奇异线性方程组 $Ax = b$，首先根据交换顺序调整 b 中的分量，即构造 $\tilde{b}$，并使其对应于 p 的顺序——然后通过前向替代法求解 $Ly = \tilde{b}$。

很容易将部分选主元的优势与 LU 分解结合，以便求解非奇异系统 $Ax = b$。

由于置换矩阵是非奇异的，系统 $Ax = b$ 等价于

$$PAx = Pb.$$

因此，我们可以使用之前讨论的 LU 解法技术来求解该置换系统。

也就是说，如果我们已经完成了因式分解 $PA = LU$，那么可以通过前向替代求解 $Ly = Pb$ 得到 y，然后通过回代求解 $Ux = y$。

问题：使用通过部分选主元获得的 LU 分解来求解系统 $Ax = b$，其中

$$A = \begin{pmatrix} 1 & 2 & -3 & 4 \\ 4 & 8 & 12 & -8 \\ 2 & 3 & 1 & 2 \\ -3 & -1 & -1 & 4 \end{pmatrix} \quad \text{和} \quad b = \begin{pmatrix} 3 \\ 60 \\ 1 \\ 5 \end{pmatrix}.$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ -3/4 & 1 & 0 & 0 \\ 1/4 & 0 & 1 & 0 \\ 1/2 & -1/5 & 1/3 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 60 \\ 5 \\ 3 \\ 1 \end{pmatrix} \Rightarrow y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 60 \\ 50 \\ -12 \\ -15 \end{pmatrix}.$$

然后通过回代求解 $Ux = y$：

$$\begin{pmatrix} 4 & 8 & 12 & -8 \\ 0 & 5 & 10 & -10 \\ 0 & 0 & -6 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 60 \\ 50 \\ -12 \\ -15 \end{pmatrix} \Rightarrow x = \begin{pmatrix} 12 \\ 6 \\ -13 \\ -15 \end{pmatrix}.$$

LDU 分解

- 在 LU 分解中存在一些不对称性，因为下三角矩阵的对角线上是 1，而上三角矩阵的对角线上不是单位对角。
- 这可以通过将对角线元素从上三角矩阵中提取出来来解决，如下所示：

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} = \begin{pmatrix} u_{11} & 0 & \cdots & 0 \\ 0 & u_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{pmatrix} \begin{pmatrix} 1 & u_{12}/u_{11} & \cdots & u_{1n}/u_{11} \\ 0 & 1 & \cdots & u_{2n}/u_{22} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

- 设 $D = \mathrm{diag}(u_{11}, u_{22}, \cdots, u_{nn})$（主元的对角矩阵），并重新定义 U 为上面所示的最右边的上三角矩阵，这样可以将任何 LU 分解写为：

$$A = LDU.$$

- 其中，L 和 U 分别是下三角和上三角矩阵，且它们的对角线元素均为 1。
- 这称为矩阵 A 的 LDU 分解。
- 此分解是唯一的，当 A 是对称矩阵时，LDU 分解可以写为 $A = LDL^T$。