

第4课 A 的 LU 分解

本章介绍的主要内容是 A 的 LU 分解。在最开始部分首先补充介绍了一些逆矩阵的内容。然后我们研究了矩阵的 LU 分解，主要体现在求得一个一击致命的 E (其逆即为 L)，同时我们探究了为什么 L 比 E 要好。最后我们考虑了行交换的情况，介绍了置换矩阵。

- 矩阵乘积的逆：

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= I \\ (B^{-1}A^{-1})(AB) &= I\end{aligned}$$

老师给了一个生动有趣的比喻：“这就像先脱鞋子，再脱袜子，然后其逆动作应该是先穿袜子，再穿鞋子。”

- 转置矩阵和逆矩阵的关系：

注意，矩阵乘积的逆和矩阵乘积的转置都调换了矩阵的顺序： AB 的逆为 $B^{-1}A^{-1}$ ，其转置为 $B^T A^T$ 。

$$\begin{aligned}\text{已知 } AA^{-1} &= I \\ \therefore (AA^{-1})^T &= I^T = I \\ \therefore (A^{-1})^T A^T &= I \\ \therefore (A^T)^{-1} &= (A^{-1})^T\end{aligned}$$

由上可知，一个矩阵 A 的转置的逆等于 A 的逆的转置。

从 $(AB)(B^{-1}A^{-1}) = I$ 我们很容易明白为什么矩阵的逆调换了矩阵的顺序，但矩阵的转置为什么调换了矩阵的顺序呢？实际上简单地从矩阵乘法的定义就能直接推出来了。

令 $A = (a_{ij})_{m \times n}$ ， $B = (b_{ij})_{n \times s}$ ，显然 $(AB)^T$ 与 $B^T A^T$ 同型，由矩阵乘法定义及转置定义有：

$$\begin{aligned}((AB)^T)_{ij} &= (AB)_{ji} = \sum_{k=1}^n a_{jk} b_{ki} \\ (B^T A^T)_{ij} &= \sum_{k=1}^n a_{jk} b_{ki}\end{aligned}$$

所以， $(AB)^T = B^T A^T$ 。

- 关于矩阵 A 的 LU 分解中 E 和 L 的关系以及为什么选择 $A = LU$ 而不是 $EA = U$

- 我们知道，通过不断左乘初等矩阵可使得矩阵 A 变换为 U 。假设所有左乘的初等矩阵的总乘积为 E ，那么有 $EA = U$ 。

$$\begin{aligned}\therefore EA &= U \\ \therefore E^{-1}EA &= E^{-1}U \\ \therefore IA &= E^{-1}U \\ \therefore A &= E^{-1}U = LU\end{aligned}$$

显然， L 和 E 互为逆矩阵。

- $A = LU$ 比 $EA = U$ 更好。下面举个简单的例子，假设从 A 到 U 只需要进行两次行变换，其中一次是对矩阵 $(2, 1)$ 部分进行处理（记为 E_{21} ），另外一次是对矩阵 $(3, 2)$ 部分进行处理（记为 E_{32} ）。

那么有 $E_{32}E_{21}A = U$ 。

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -5 & 1 \end{bmatrix}}^{E_{32}} \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{E_{21}} = \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 10 & -5 & 1 \end{bmatrix}}^E$$

注意到 E 中出现了 10。它是如何出现的呢？这是因为在第一步中 E_{21} 有 2 倍的行一从行二中减去了，然后在第二步 E_{32} 中又有 5 倍的行二从行三中减去，因此总共在行三中加上了 10 倍行一。这也即第一步的操作会影响的第二步的操作。

注意到，我们反向计算可以避免这些影响。

也即 $A = E_{21}^{-1} E_{32}^{-1} U$ 。

$$\overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}^{E_{21}^{-1}} \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}}^{E_{32}^{-1}} = \overbrace{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}}^L$$

显然， L 中矩阵的顺序非常好。在第一步中 E_{32}^{-1} 将 5 倍的行二与行三相加，在第二步中 E_{21}^{-1} 将 2 倍的行一与行二相加。注意到，第一步中的行二并没有被任何操作所修改，所以自然也就不会有 10 的出现。此外，容易发现的是， L 矩阵各个元素都是 E_{21}^{-1} 与 E_{32}^{-1} 中对应位置的元素！

- 对于一个 $n \times n$ 矩阵 A ，我们需要进行大约 $\frac{1}{3}n^3$ 次操作将其变换为 U 。而对于右侧向量 b 的变换和回代求解我们需要大约各 $\frac{1}{2}n^2$ 共 n^2 次操作。

在上面对矩阵 A 的 LU 分解我们没有考虑到允许行交换的情况而是假定 A 是一个相当好的矩阵（也即在消元过程中不需要进行与行之间的交换）。在下一讲中我们将会考虑行交换的情况，而在进入下一讲之前，我们先介绍一点点置换矩阵的知识。

- **置换矩阵：用于实现行交换**

对于 3×3 矩阵，我们能找到 6 个置换矩阵（有穷的）：

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

这 6 个置换矩阵组成了一个很有意思的矩阵群。这六个矩阵不管怎么相乘或者取逆，其结果仍然在这个群中。这样的情况是很自然的，因为对矩阵的各行进行多次交换存在最后矩阵回到最初的可能。

通过观察列出的 6 个置换矩阵，我们能发现一条很奇妙的关于置换矩阵的性质：**置换矩阵的逆等于其转置，也即 $P^{-1} = P^T$ 。**