1.6 矩阵A的LU分解

先告诉大家L是下三角矩阵，而u是上三角矩阵！

前置定理

在开始讲解A的分解之前，我们需要先明白几个定理。

一：

$$(AB)^{-1} = B^{-1}A^{-1}$$

记住他，这么记：脱下鞋子，再脱袜子，但是当你再想穿上时要先穿上袜子再穿鞋子！

二：

首先我们需要大致了解一下转置矩阵：
转置矩阵就是将原矩阵各行换成对应列，所得到的新矩阵：

$$\mathrm{A} = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix} \quad A^T = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 3 & 1 \end{bmatrix}$$

可以看作是沿着主对角线进行向里翻折！
介绍完了转置矩阵的基础，接下来看一看它和逆矩阵有什么联系:
以这个式子为例：

$$AA^{-1} = I$$

同时对这两边进行转置，得到的是：

$$(AA^{-1})^T = (A^{-1})^T A^T = I$$

为什么$(A^{-1})^T$会变换到A^T的前面来呢？有一个感性的直观理解：把等式左边看成是一个增广矩阵，可知单位矩阵是沿着主对角线进行向里翻折的，那么前面的增广矩阵也得一样，自然顺序就调换了！大家可以这么理解，但是我也不知道严谨不严谨！实在不行就按照脱鞋那么来记忆。

现在由于A是方阵，那么A的转置矩阵也是方阵，那么是不是就可以得到：

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

所以得到的结论是：
对于单个矩阵而言，转置与取逆两个运算顺序可颠倒。

A=LU

讲完这些前置知识，那么回到主题了。
首先在前面的消元法化矩阵A为上三角矩阵，我们都是通过对A的行进行操作，那么理所应当的应该把记录这些行操作的矩阵放到A的左边（左乘代表行操作希望大家没有忘记）。那么就会有这样的等式存在：$(E_{32}E_{21})A = U$，现在我们要把他们写到右边来A = LU。那为什么我们这么执着与写到右边来呢。给个例子说明：

现有 $E_{32}E_{31}E_{21}A = U$，已知

$$E_{32} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -5 & 1\end{bmatrix},\ E_{21} = \begin{bmatrix}1 & 0 & 0\\-2 & 1 & 0\\0 & 0 & 1\end{bmatrix},\ E_{31} = I。$$

求 A = LU分解后的 L。

- 思路：
  逆矩阵化简为：$A = (E_{21})^{-1}(E_{31})^{-1}(E_{32})^{-1}U$ （注意顺序！）

  计算出各个矩阵：

  $$(E_{21})^{-1} = \begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\0 & 0 & 1\end{bmatrix},\ (E_{32})^{-1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 5 & 1\end{bmatrix}$$

  。

  直接代入计算，$L = (E_{21})^{-1}(E_{31})^{-1}(E_{32})^{-1} = \begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\0 & 5 & 1\end{bmatrix}$
  那么重点来了：

LU 分解的本质：消元过程的记录

LU 分解的目标是将一个矩阵 A 分解成两个三角矩阵：

$$A = LU$$

其中：

- L：一个**下三角矩阵**（Lower Triangular），对角线上的元素为 1。
- U：一个**上三角矩阵**（Upper Triangular）。

这个分解的核心思想是：

> **我们通过一系列行变换将矩阵 A 变成上三角矩阵 U，然后把这些行变换记录下来，就得到了下三角矩阵 L。**

高斯消元法回顾：从 A 到 U

在高斯消元过程中，我们会做以下类型的行变换：

- 用某一行减去另一行的某个倍数，例如：

$$R_2 \leftarrow R_2 - l_{21}R_1$$

这些操作的目的就是让某些位置变成 0，最终得到一个上三角矩阵 U。

假设我们有两个这样的操作：

1. 第一步：$R_2 \leftarrow R_2 - 2R_1$
2. 第二步：$R_3 \leftarrow R_3 - 5R_2$

那么我们就知道：

- 消元矩阵分别是 E_{21} 和 E_{32}
- 它们的逆矩阵分别是 $(E_{21})^{-1}$ 和 $(E_{32})^{-1}$

从 U 回到 A：需要哪些操作？

我们知道：

$$E_{32}E_{21}A = U \Rightarrow A = (E_{21})^{-1}(E_{32})^{-1}U$$

所以我们可以把 L 写成：

$$L = (E_{21})^{-1}(E_{32})^{-1}$$

而关键点来了：

> **$(E_{21})^{-1}$ 和 $(E_{32})^{-1}$ 都是下三角矩阵，它们的乘积仍然是下三角矩阵，也就是 L。**

更神奇的是：

> **我们不需要真的计算这些逆矩阵，只需要知道消元过程中用了什么系数，就能直接写出 L！**

如何从行变换直接写出 L？

关键观察：

在消元过程中，如果我们做了如下操作：

- 用第 2 行减去了第 1 行的 **2 倍**（即 $R_2 \leftarrow R_2 - 2R_1$）这里的系数1和2将填在第二行第一列和第二行第二列，因为这是为了消除第二行第一列元素的操作，以左乘（行视角）的角度去看的话就是这样。
- 用第 3 行减去了第 2 行的 **5 倍**（即 $R_3 \leftarrow R_3 - 5R_2$）这里的系数1和5将填在第三行第二列和第三行第三列，因为这是为了消除第三行第二列元素的操作！
  没有用到的行就填0！

那么这些"系数"就会出现在 L 矩阵的特定位置上：

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 5 & 1 \end{bmatrix}$$

也就是说：

- 在 L 中，第 2 行第 1 列的位置填的是我们在消元时使用的系数 **2**
- 第 3 行第 2 列的位置填的是我们在消元时使用的系数 **5**

题外话，这样子看L永远都是满秩的！（学完1.11之后来看就会知道！）

总结一下这个"秘诀"

步骤	操作	对应到 L
1	$R_2 \leftarrow R_2 - 2R_1$	$L[2,1] = 2$
2	$R_3 \leftarrow R_3 - 5R_2$	$L[3,2] = 5$

✅ 我们不需要真正计算逆矩阵
✅ 只需要记录每次行变换中用到了哪些系数
✅ 就可以直接构造出 L

打个比方，想象你在做蛋糕，步骤是：
把面粉加水揉成面团（→ 类似从 A 得到 U）
然后你发现："如果我要从面团还原回原始面粉+水的比例，我只要记得我在哪一步加了多少水就行了。"
同理：

- 从 A 到 U 的过程就像是"揉面"
- 从 U 到 A 的过程就像是"还原原料比例"

- 而那些“加水量”就是行变换的系数，它们构成了 L

相信大家通过上面的讲解，已经明白了A=LU的重要性！也希望大家领悟他的奇妙之处！至于为什么这么神奇，证明过程太复杂，就跳过了！其实这也是我喜欢这门课程的原因之一，没有复杂且冗余（对于非数学专业的我们来说这些证明就是冗余的）证明过程，有的只是恰到好处的例子与许多的类比说明！

矩阵A分解的运算量

就是把矩阵A消元成一个上三角矩阵的运算量是多少：
比如现在我们有一个100 * 100的超级大的矩阵(无0元素)。我们需要运算（将每个元素的一次乘法与减法作为一个过程，每一个这样的过程计为一次运算）多少次之后，才能将其化为上三角矩阵 U 呢？

$$A=\begin{bmatrix} L & \cdots & \cdots & \cdots & \cdots \\ \cdots & \ddots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \ddots & \cdots & \cdots \\ \vdots & \cdots & \cdots & \ddots & \cdots \\ \cdots & \cdots & \cdots & \cdots & L \end{bmatrix}$$

先消除第一列的所有元素，让A变成这样：

$$\begin{bmatrix} 1 & \cdots & \cdots & \cdots & \cdots \\ 0 & \ddots & \cdots & \cdots & \cdots \\ 0 & \cdots & \ddots & \cdots & \cdots \\ \vdots & \cdots & \cdots & \ddots & \cdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{bmatrix}$$

首先为了消除第二行的第一个元素，那么第二行需要乘以一个系数，再减去第一行。那么第二行有100个元素需要这样的操作，那么有100次这个过程！接着到第三行，以此类推一共有99行，每行有100个这样的操作。位99 *100，我们省略位100* 100。

然后消除第二列的所有元素，矩阵大小就变成了99阶的了，按照上面的方法，操作就需要99 * 99次了。一直到变换位上三角矩阵，一共有：$1^2+2^2+\cdots+(n-1)^2+n^2$次。
那么我们近似看成是求连续区间上的黎曼和，而并非是离散的点的值，通过微积分的求和公式可以得到：$1^2+2^2+\cdots+(n-1)^2+n^2=\sum_{i=1}^{n} i^2 \approx \int_0^n x^2dx=\frac{1}{3}n^3$。
当然我们还可以用高中的数列求和的方法来写（裂项相消）：不知道大家记不记得，结果也是一样的！

那么在一个AX = b的求解中，b是随着A一起变动的，根据上面的推到方法，b的操作次数大概是n^2，这比起A来说可以忽略不记！

置换矩阵群

我们之前接触过行变换所用到的矩阵，即是将单位阵 I 按照对应行变换方式进行操作之后得到的矩阵。它可以交换矩阵中的两行，代替矩阵行变换。什么时候我们需要使用矩阵行变换呢？一个经典的例子就是：在消元过程中，当矩阵主元位置上面为 0 时，我们就需要用行变换将主元位置换为非 0 数。这样的由单位阵变换而来的矩阵，通过矩阵乘法可以使被乘矩阵行交换。我们将这样的矩阵称为置换矩阵 P。我们通过一个例子来熟悉一下置换矩阵。
求矩阵

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

的所有置换矩阵，并判断其性质。 一共有 6 个置换矩阵：

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

这可以理解为一个矩阵群，很明显，我们任取两个矩阵相乘，结果仍在这个矩阵群中。

注： 推广到 n 阶矩阵，n 阶矩阵有 n! 个置换矩阵，就是将单位矩阵 I 各行重新排列后所有可能的情况数量。这个应该不难理解，就是不同行之间的一个排列问题，结果就是排列公式的结果！

而且这个矩阵群中的每一个矩阵都有一个神奇的地方，矩阵的逆等于转置矩阵，大家可以计算一个试试看！像这样的矩阵我们叫做正交矩阵！

讨论课内容：

有这么个矩阵：

$$A=\begin{bmatrix}1 & 0 & 1\\a & a & a\\b & b & a\end{bmatrix}$$

找出他的LU分解（就是找出LU）。以及说出a，b满足什么条件的时候A会有LU分解。
相信大家有一个大概一样的思路：先把A转换为上三角矩阵，然后用矩阵记录下这消元的过程。再求逆！在上面写过太多的消元过程，本来不想写，但是为了强调出两个重点，还是把他写出来吧。

第一步，消除第二行第一列的元素，第二行减去第一行的a倍，A此时此刻为：

$$A=\begin{bmatrix}1 & 0 & 1\\0 & a & 0\\b & b & a\end{bmatrix}$$

那么E_{21}为：

$$E_{21}=\begin{bmatrix}1 & 0 & 0\\ -a & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

第二步：消除第三行第一列的元素，第三行减去第一行的b倍，A为：

$$A=\begin{bmatrix}1 & 0 & 1\\ 0 & a & 0\\ 0 & b & a-b\end{bmatrix}$$

那么E_{31}为：

$$E_{31}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -b & 0 & 1\end{bmatrix}$$

第三步：消除第三行第二列的元素，第三行减去第二行的$\frac{b}{a}$倍，A为：

$$A=\begin{bmatrix}1 & 0 & 1\\ 0 & a & 0\\ 0 & 0 & a-b\end{bmatrix}$$

那么E_{32}为：

$$E_{32}=\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & \frac{b}{a} & 1\end{bmatrix}$$

那么

$$U=\begin{bmatrix}1 & 0 & 1\\ 0 & a & 0\\ 0 & 0 & a-b\end{bmatrix}$$

关键点一来了：视线回到第二步，发现当a=0时，需要换行进行正常消元，但是我们需要注意的是，换行之后的LU就不能是A的分解了，准确的说是PA = LU，这添加了一个置换矩阵！

关键点二：求L，我们应该去求E们的逆吗，这样太过于麻烦，但是大家还记得A=LU的目的吗。那么我来演示一下:

$$L=\begin{bmatrix} & 0 & 0\\ a & 1 & 0\\ b & \frac{b}{a} & 1\end{bmatrix}$$

在特定系数填完后，还剩下两个系数，那么讲LU放在一起对比，发现U与A第一行是不变的。所以第一个空填1。

$$L=\begin{bmatrix}1&0&0\\a&1&0\\b&\frac{b}{a}&1\end{bmatrix}$$

那么

$$L=\begin{bmatrix}1&0&0\\a&1&0\\b&\frac{b}{a}&1\end{bmatrix}$$

比起求逆是不是简单许多！所以大家记住这个简便的方法！当然，也要勤加练习求逆的运算！
大家发现L的主对角线元素都是1，这其实是因为我们在消元过程中就永远把要消元的行系数设为1！所以在可以正常消元情况下的L，一般主对角线元素都为1！

习题课

第一题：
有一个矩阵：

$$A=\begin{bmatrix}1&3&0\\2&4&0\\2&0&1\end{bmatrix}$$

求这个矩阵分解后的L！
这个小菜一碟，我们再来重复上面的过程！
第一步：第二行减去第一行的两倍。得到：

$$A=\begin{bmatrix}1&3&0\\0&-2&0\\2&0&1\end{bmatrix}$$

第二步：第三行减去第一行的两倍。得到：

$$A=\begin{bmatrix}1&3&0\\0&-2&0\\0&-6&1\end{bmatrix}$$

第三步：第三行减去第二行的3倍。得到：

$$A=\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix}$$

那么最后：

$$U=\begin{bmatrix}1&3&0\\0&-2&0\\0&0&1\end{bmatrix}$$

现在我们先得到L：

$$L=\begin{bmatrix} & 0 & 0\\ & & 0\\ & & \end{bmatrix}$$

再填系数：

$$L=\begin{bmatrix} & 0 & 0\\ 2 & 1 & 0\\ 2 & 3 & 1\end{bmatrix}$$

同样的，意味当把残缺的LU 放在一起时，发现U与A的第一行一样，那么L：

$$L=\begin{bmatrix}1 & 0 & 0\\ 2 & 1 & 0\\ 2 & 3 & 1\end{bmatrix}$$

那么我们就求出L了！

第二题：

$$A=\begin{bmatrix}a & a & a & a\\ a & b & b & b\\ a & b & c & c\\ a & b & c & d\end{bmatrix}$$

求LU，并告诉我，a和b满足什么条件下U会有4个主元（就是U的主对角线上都有元素）。
其实这里我们需要解读一下题意，一般可以正常消元情况下，U主对角线上的元素都不为0，因为如果出现了0，我们必须要置换才能继续消元或者根本就消不了元。所有这是在告诉我们什么时候A可以正常消元！这么一说，是不是与朗读课上的那道题一模一样了！
首先求出LU，由于前面重复了很多次，我们这里就简写：

$$A\rightarrow\begin{bmatrix}a & a & a & a\\ 0 & b-a & b-a & b-a\\ 0 & b-a & c-a & c-a\\ 0 & b-a & c-a & d-a\end{bmatrix}\rightarrow\begin{bmatrix}a & a & a & a\\ 0 & b-a & b-a & b-a\\ 0 & 0 & c-b & c-b\\ 0 & 0 & c-b & d-b\end{bmatrix}\rightarrow\begin{bmatrix}a & a & a & a\\ 0 & b-a & b-a & b-a\\ 0 & 0 & c-b & c-b\\ 0 & 0 & 0 & d-c\end{bmatrix}=U,L=$$

观察这几个步骤，看到最后U，那么就是：
$a\neq 0$，$b\neq a$，$c\neq b$，$d\neq c$，这就是答案！

希望大家可以通过这一讲体会到A分解的奥秘！