线性代数

---4 课　A 的 LU 分解

一、知识概要

这一节中首先完善之前讲到的逆矩阵内容，然后使用消元矩阵介绍 A 的 LU 分解，即：将矩阵 A 分解为矩阵 L 与上三角矩阵 U，介绍这种运算的普遍规律。最后再一次提起了之前介绍过的“行交换矩阵”，引入置换矩阵概念。

二．逆矩阵性质补充

首先考虑一个问题：方阵 A，B 都是可逆矩阵的话，AB 的逆矩阵是什么呢？这个问题并不复杂，想求出逆矩阵，无非就是令 AB*逆矩阵 = I，而我们不难想到，$ABB^{-1}A^{-1}$ = I。所以有：

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

由于下一章中要涉及到矩阵的转置问题，我们在这里一并讨论矩阵转置与矩阵的逆的关系。

首先介绍一下转置矩阵，转置矩阵就是将原矩阵各行换成对应列，所得到的新矩阵，如：

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

看起来就像是沿着左上角开始的一条对角线翻折了一样。

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

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。

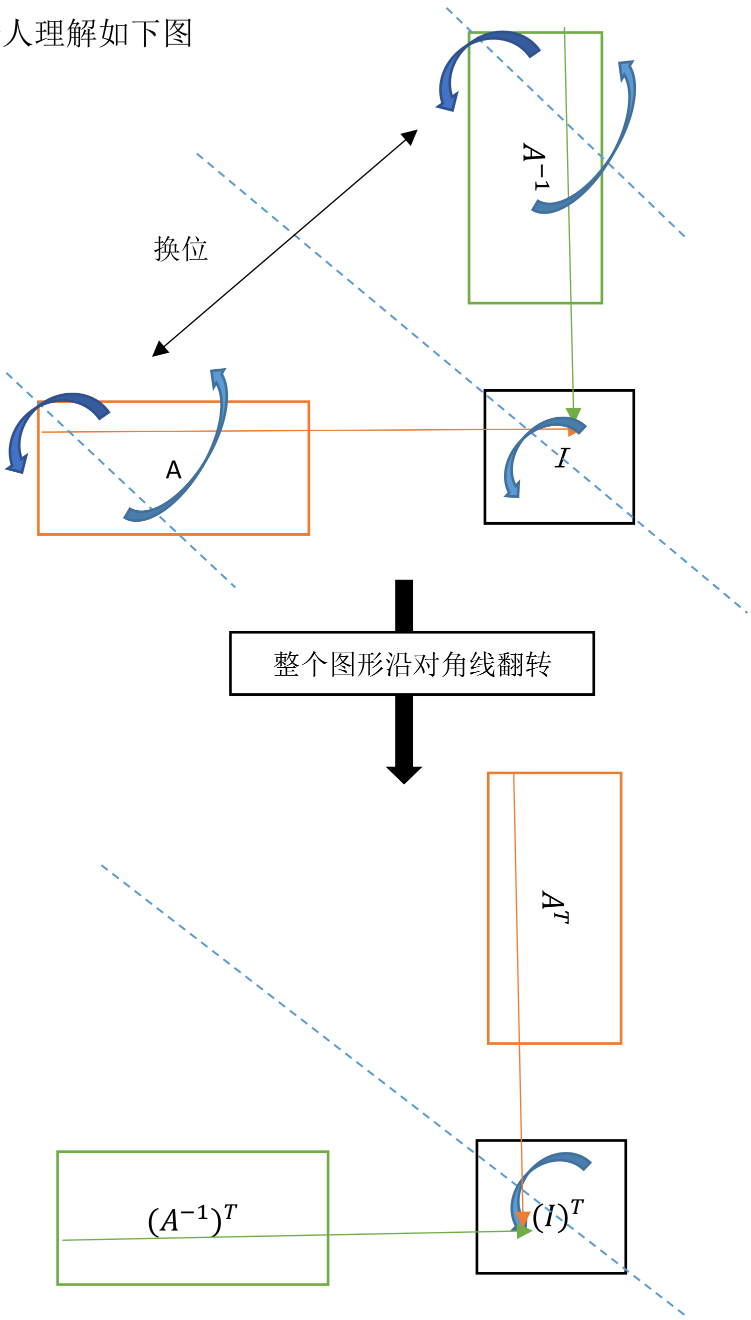
说到逆矩阵，最经典的式子无非就是：AA^{-1} = I 。为了找到转置矩阵A^T与逆矩阵A^{-1}间的关系，我们对AA^{-1} = I 两边同时进行转置运算，得到：

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

为什么$(A^{-1})^T$会变换到A^T的前面来呢？我们想象一下，最后乘积所得的单位矩阵 I 中每一个元素都是由 A 的行向量与A^{-1}的列向量构成，当做转置运算时，

I 沿对角线翻折，可以理解为整个乘法运算图形也要沿着 I 的对角线进行翻折，这样就解释了$(AA^{-1})^T = (A^{-1})^T A^T$的原因。

个人理解如下图



再观察此式：$(A^{-1})^T A^T = \mathrm{I}$，由于 A 是方阵，则$A^T$定然也是方阵，那么我们发现意外得到了$A^T$的逆矩阵。即为$(A^{-1})^T$。也就是说：

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

这个结果告诉我们：对于单个矩阵，转置与取逆两个运算顺序可颠倒。

三. A 的 LU 分解

我们熟悉的的消元法都是直接使用行变换得来的。而由于消元矩阵的存在，说明用矩阵乘法也可以达与之到一样的消元效果，所以，现在假设有可逆矩阵 A，若有：$A \xrightarrow[\text{初等行变换}]{} U$(上三角矩阵)。就一定有类似于这样的形式：$E_{21} * A = U$ 的等式存在，使 A 相当于进行了初等行变换成为 U。而我们已经学习了逆矩阵，E_{21} 这样的矩阵一定有逆矩阵，因为它本身就是单位阵变化过来的。所以原式可以改写成：$A = (E_{21})^{-1} U$。这一形式即为 A = LU 形式，这个过程就是分解过程。

那么矩阵 L 是不是有什么特殊之处呢？我们通过一道例题来探讨下 。

【例】

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，$L = (E_{21})^{-1}(E_{31})^{-1}(E_{32})^{-1} = \begin{bmatrix}1 & 0 & 0\\2 & 1 & 0\\0 & 5 & 1\end{bmatrix}$

注：

观察发现，L 具有一个非常重要的特点，L 矩阵中各个元素都是$(E_{21})^{-1}$与$(E_{32})^{-1}$中对应位置的元素：

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

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

这就给了我们启示，在使用 A = LU 分解矩阵的时候，我们只需要从 U 入手，反过来考虑，**看如何通过行变换可以将上三角矩阵 U 变为 A，然后再将单位阵按此形式变化，就得到了 L 矩阵**。这个性质也是 A = LU 形式分解矩阵的最大优点，我们甚至不需要知道类似的值到底是什么，我们只需要知道变换形式，即可求出 L，写出 A = LU 等式。

以上，我们已经学会了 A = LU 分解矩阵方法，那么现在有一个额外问题，就是消元的运算量问题，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（将一行乘一定倍数后加到另一行上消元，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？

这个问题我们先从列的角度进行考虑，第一列消元运算结束之后，矩阵将会变为$\begin{bmatrix} 1 & \dots & \cdots & \cdots & \cdots \\ 0 & \dots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & \cdots \\ \dots & \cdots & \cdots & \cdots & \cdots \\ 0 & \dots & \cdots & \cdots & \cdots \end{bmatrix}$形式，这一步中，第一列的元素运算了 100 次，而第一行一共有 100 个元素，于是仅第一行与第一列的消元结束后，我们就运算了100^2次。之后我们要研究的就变成了剩下的 99*99 的矩阵。以此类推，可知，最后的运算量为：$\sum_{\mathrm{k}=1}^{\mathrm{n}} k^2$

这个写法看上去比较难以计算，那么我们有没有什么估计大致值的方法呢？回想我们微积分中学到的知识，如果我们计算的不是离散的点（$1^2+2^2+3^2+\cdots$）的值之和，而是连续区域上函数的黎曼和的话，我们可以通过积分来计算区域面积值的和。而这里我们的离散点很多，近似可以看做黎曼和，所以我们可以采用积分来估计和，也就是这样估算：$\sum_{\mathrm{k}=1}^{\mathrm{n}} k^2 = \int_1^n x^2 dx$

三. 置换矩阵

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。

【例】

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：

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

这可以理解为一个群，很明显，我们任取两个矩阵相乘，结果仍在这个群中。

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。我自己的理解是：单看第一行，有 n 种排列方式，再看除去第一行，第一列的(n-1)阶矩阵，再看其第一行，有(n-1)种排列方式。以此类推，直到最后的 1 阶，有 1 种排列方式，由乘法原理，就有了 n!个置换矩阵。

四、学习感悟

线性代数的前面这部分基本是一些技巧的运算。本节我们对矩阵的转置，逆矩阵性质进行了部分介绍，学习了矩阵的 A = LU 分解，了解了这种分解方式的优点所在，并学会了直接构造 L 矩阵，简化消元过程。这些技巧与知识都是我们接下来学习的重要基础。