线性代数课程笔记

Linear Algebra 18.06 Course Notes

作者：1210

邮箱：

时间：March 1, 2020

组织：

别人都关心你飞的有多高，只有我关心你的翅膀好不好吃！

特别声明

线性代数（英语：linear algebra）是关于向量空间和线性映射的一个数学分支。它包括对线、面和子空间的研究，同时也涉及到所有的向量空间的一般性质。

坐标满足线性方程的点集形成 n 维空间中的一个超平面。n 个超平面相交于一点的条件是线性代数研究的一个重要焦点。此项研究源于包含多个未知数的线性方程组。这样的方程组可以很自然地表示为矩阵和向量的形式。

线性代数既是纯数学也是应用数学的核心。例如，放宽向量空间的公理就产生抽象代数，也就出现若干推广。泛函分析研究无穷维情形的向量空间理论。线性代数与微积分结合，使得微分方程线性系统的求解更加便利。线性代数的理论已被泛化为算子理论。

线性代数的方法还用在解析几何、工程、物理、自然科学、计算机科学、计算机动画和社会科学（尤其是经济学）中。由于线性代数是一套完善的理论，非线性数学模型通常可以被近似为线性模型。

课程网址:Linear Algebra18.06, 课程笔记是基于网友的笔记进行重置的.(基本是复制的)

language of flowers

目录

第一章 方程组的几何解释

1.1 概述

本节主要介绍线性代数的基础。首先从解方程开始，学习线性代数的应用之一就是求解复杂的方程问题，本节核心之一就是从 row picture (行图像) 和 columnpicture (列图像) 的角度解方程。

1.2 方程组的几何解释基础

1.2.1 二维行图像

我们首先通过一个例子了解二维方程组（2 个未知数，2 个方程），如下：

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases}$$

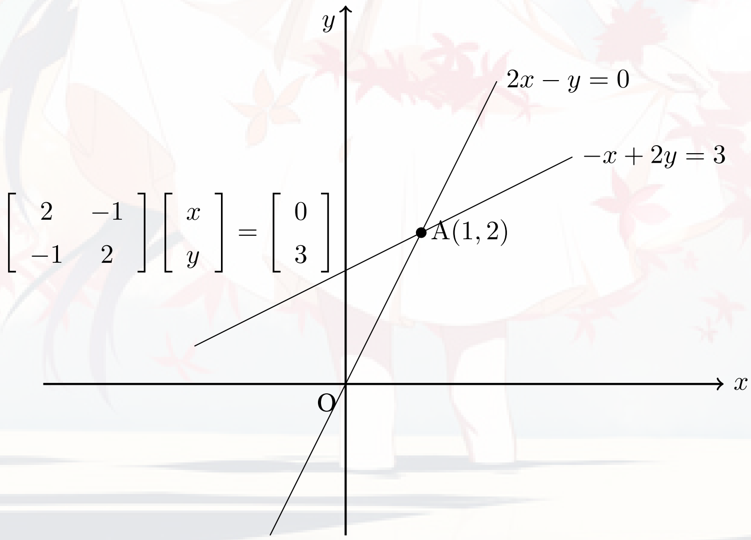
我们首先按 row (行) 将方程组写成矩阵形式：

$$\underbrace{\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}}_{\text{系数矩阵}} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\text{未知向量}} = \underbrace{\begin{bmatrix} 0 \\ 3 \end{bmatrix}}_{\text{向量}}$$

1. 系数矩阵 (A): 将方程组系数按行提取出来，构造完成的一个矩阵。
2. 未知向量 (x): 将方程组的未知数提取出来，按列构成一个向量。
3. 向量 (b): 将等号右侧结果按列提取，构成一个向量。

构造完成相应的矩阵形式了，我们将对应的行图像画出来。

所谓行图像，就是在系数矩阵上，一次取一行构成方程，在坐标系上做出如下图。和我们在初等数学中学习的作图求解方程的过程无异。



1.2.2 二维列图像

从列图像的角度，我们再次求解上面的方程：

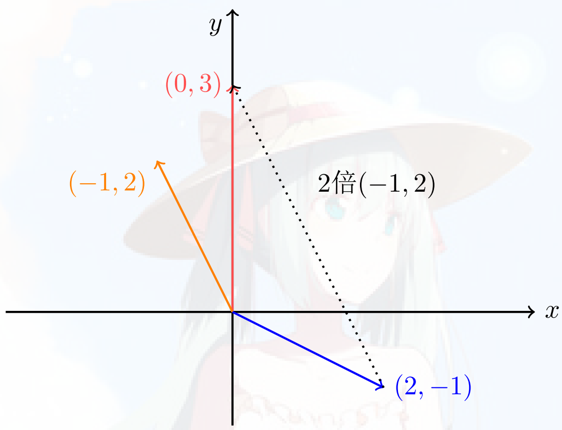
$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases}$$

在这一次的求解过程中，我们将方程按列提取，使用的矩阵为：

$$x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

如上所示，我们使用列向量构成系数矩阵，将问题转化为: 将向量 $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ 与向量 $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ 正确组合，使得其结果构成 $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$。

接下来我们使用列图像将方程组展现出来，并求解：



即寻找合适的 x，y 使得 x 倍的 $(2, -1) + y$ 倍的 $(-1, 2)$ 得到最终的向量 $(0, 3)$。在很明显能看出来，1 倍 $(2, -1)$ +2 倍 $(-1, 2)$ 即满足条件。

反映在图像上，明显结果正确即满足条件。

我们再想一下，仅仅对

$$x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

这个方程，如果我们任意取 x 和 y，那么我们得到的是什么呢？很明显，能得到任意方向的向量，这些向量能够布满整个平面。这里我们先不做展开，稍微有一些印象就好。

1.3 方程组的几何解释推广

1.3.1 高维行图像

我们将方程组的维数进行推广，从三维开始，

$$\begin{cases} 2x - y = 0 \\ -x + 2y - z = -1 \\ -3y + 4z = 4 \end{cases}$$

，如果我们继续使用上面介绍的做出行图像来求解问题，那么会得到一个很复杂的图像。

矩阵如下：

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

对应方程：$Ax = b$

$$\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

如果绘制行图像，很明显这是一个三个平面相交得到一点，我们想直接看出这个点的性质可谓是难上加难。

比较靠谱的思路是先联立其中两个平面，使其相交于一条直线，再研究这条直线与平面相交于哪个点，最后得到点坐标即为方程的解。

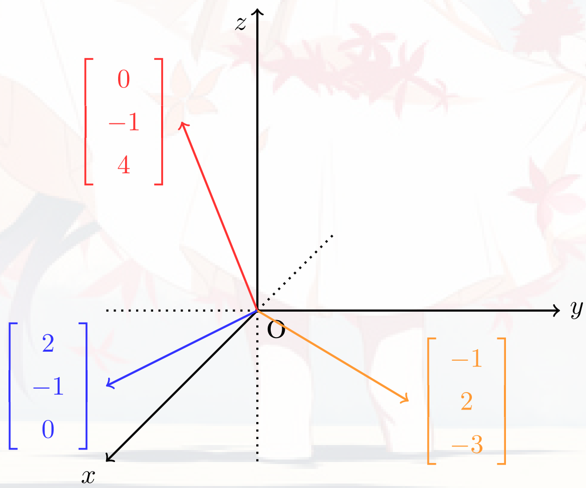
这个求解过程对于三维来说或许还算合理，那四维呢？五维甚至更高维数呢？直观上很难直接绘制更高维数的图像，这种行图像受到的限制也越来越多。

1.3.2 高维列图像

我们使用上面同样的例子，$\begin{cases} 2x - y = 0 \\ -x + 2y - z = -1 \\ -3y + 4z = 4 \end{cases}$，如果我们使用列图像的思路进行计算，那矩阵形式就变为：

$$x \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \\ -3 \end{bmatrix} + z \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

左侧是线性组合，右侧是合适的线性组合组成的结果，这样一来思路就清晰多了，“寻找线性组合”成为了解题关键。



很明显这道题是一个特例，我们只需要取 $x = 0$，$y = 0$，$z = 1$。就得到了结。就得到了结果，这在行图像之中并不明显。当然, 之所以我们更推荐使用列求解方程，是因为这是一种更系统的求解方法，即寻找线性组合，而不用绘制每个行方程的图像之后寻找那个很难看出来的点。另外一

个优势在于，如果我们改变最后的结果 b，例如本题中，我们将其改为

$$x\begin{bmatrix}2\\-1\\0\end{bmatrix}+y\begin{bmatrix}-1\\2\\-3\end{bmatrix}+z\begin{bmatrix}0\\-1\\4\end{bmatrix}=\begin{bmatrix}1\\1\\-3\end{bmatrix}$$

, 那么我们就重新寻找一个线性组合就够了，但是如果我们使用的是行图像呢？那意味着我们要完全重画三个平面图像，就简便性来讲，两种方法高下立判。

另外，还要注意的一点是对任意的 b 是不是都能求解 $Ax=b$ 这个矩阵方程呢？也就是对 $3*3$ 的系数矩阵 A，其列的线性组合是不是都可以覆盖整个三维空间呢？对于我们举的这个例子来说，一定可以，还有我们上面 $2*2$ 的那个例子，也可以覆盖整个平面，但是有一些矩阵就是不行的，比如三个列向量本身就构成了一个平面，那么这样的三个向量组合成的向量只能活动在这个平面上，肯定无法覆盖一个三维空间，比如三个列向量分别为：$\begin{bmatrix}2, & -1, & 2\end{bmatrix}^T$, $\begin{bmatrix}-1, & -1, & 3\end{bmatrix}^T$, $\begin{bmatrix}1, & 1, & 3\end{bmatrix}^T$ 这三个向量就构成了一个平面。其中

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

这样的矩阵 $A\begin{bmatrix}2 & -1 & 1\\-1 & 2 & 1\\0 & -3 & -3\end{bmatrix}$ 构成的方程 $Ax=b$，其中的 b 就无法覆盖整个三维空间，也就无法实现：对任意的 b，都能求解 $Ax=b$ 这个方程。

1.3.3 矩阵乘法

例如 AX, 如果我们已知一个矩阵 A 和一个向量 x, 我们如何求解它们的积呢？例如 $A=\begin{bmatrix}2 & 5\\1 & 3\end{bmatrix}$, $x=\begin{bmatrix}1\\2\end{bmatrix}$, 我们这样求：

1.3.3.1 将矩阵 A 看做列向量的组合

$$\begin{bmatrix}2 & 5\\1 & 3\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}=1\begin{bmatrix}2\\1\end{bmatrix}+2\begin{bmatrix}5\\3\end{bmatrix}=\begin{bmatrix}12\\7\end{bmatrix}$$

即 x 每个分量与矩阵中各的列向量相乘，再将其求和。看做 A 各列的线性组合。

1.3.3.2 将矩阵 A 看做行向量的组合：

$$\begin{bmatrix}2 & 5\\1 & 3\end{bmatrix}\begin{bmatrix}1\\2\end{bmatrix}=\begin{bmatrix}(2,5)\cdot(1,2)\\(1,3)\cdot(1,2)\end{bmatrix}=\begin{bmatrix}12\\7\end{bmatrix}$$

即采用这个方式进行向量乘法：

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & * & 2 & + & 2 & * & 5 \\ 1 & * & 1 & + & 2 & * & 3 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & * & 2 & + & 2 & * & 5 \\ 1 & * & 1 & + & 2 & * & 3 \end{bmatrix}$$

笔记 这部分内容是对线性代数概念的初涉，从解方程谈起，引进列空间的概念，可以发现从列空间角度将求解方程变化为求列向量的线性组合，这个方式更加科学。

第二章 矩阵消元

2.1 Overview (概述)

在本节中，我们主要介绍消元法 (Elimination)，也就是我们在上一节中提到的“系统化”求解方程所用的方法。通过矩阵的消元运算可以很轻松地求解复杂方程。后半部分还介绍了消元矩阵，即我们的消元运算在矩阵乘法中所表现的形式。并从消元矩阵引入，介绍了一些逆矩阵的基础知识。

2.2 使用消元法求解方程

2.2.1 高斯消元法

笔记 能够使用消元法分为以下 2 种情况:

1. 成功: 矩阵的主元不包含 0，也就是说，可逆矩阵。
2. 失败: 举着的主元有至少一个 0，也就是说，不可逆矩阵。

数学上，高斯消元法（英语：Gaussian Elimination），是线性代数中的一个算法，可用来为线性方程组求解，求出矩阵的秩，以及求出可逆方阵的逆矩阵。当用于一个矩阵时，高斯消元法会产生出一个行梯阵式。

高斯消元法可用来找出下列方程组的解或其解的限制：

$$\begin{cases} 2x + y - z = 8 & (L_1) \\ -3x - y + 2z = -11 & (L_2) \\ -2x + y + 2z = -3 & (L_3) \end{cases}$$

这个算法的原理是：

首先，要将 L_1 以下的等式中的 x 消除，然后再 L_2 以下的等式中的 y 消除。这样可使整个方程组变成一个三角形似的格式。之后再将已得出的答案一个个地代入已被简化的等式中的未知数中，就可求出其余的答案了。

在刚才的例子中，我们将 $\frac{3}{2}L_1$ 和 L_2 相加，就可以将 L_2 中的 x 消除了。然后再将 L_1 和 L_3 相加，就可以将 L_3 中的 x 消除。

我们可以这样写：

$$L_2 + \frac{3}{2}L_1 \to L_2$$
$$L_3 + L_1 \to L_3$$

结果就是：

$$2x + y - z = 8$$
$$\frac{1}{2}y + \frac{1}{2}z = 1$$
$$2y + z = 5$$

现在将 $-4L_2$ 和 L_3 相加，就可将 L_3 中的 y 消除：

$$L_3 + -4L_2 \to L_3$$

其结果是：

$$\begin{aligned}&2x+y-z=8\\&\frac{1}{2}y+\frac{1}{2}z=1\\&-z=1\end{aligned}$$

这样就完成了整个算法的初步，一个三角形的格式（指：变量的格式而言，上例中的变量各为 3,2,1 个）出现了。

第二步，就是由尾至头地将已知的答案代入其他等式中的未知数。第一个答案就是：

$$z=-1 \quad (L_3)$$

然后就可以将 z 代入 L_2 中，立即就可得出第二个答案：

$$y=3 \quad (L_2)$$

之后，将 z 和 y 代入 L_1 之中，最后一个答案就出来了：

$$x=2 \quad (L_1)$$

就是这样，这个方程组就被高斯消元法 解决了。

笔记 注：并不是所有的 A 矩阵都可消元处理，需要注意在我们消元过程中，如果主元位置（左上角）为 0，那么意味着这个主元不可取，需要进行“换行”处理：

首先看它的下一行对应位置是不是 0，如果不是，就将这两行位置互换，将非零数视为主元。

如果是，就再看下下行，以此类推。若其下面每一行都看到了，仍然没有非零数的话，那就意味着这个矩阵不可逆，消元法求出的解不唯一。

2.2.2 回带求解

这里其实是高斯-若尔当消元法，在高斯消元法的基础上多了一步回代求解过程。

通常人或电脑在应用高斯消元法的时候，不会直接写出方程组的等式来消去未知数，反而会使用矩阵来计算。以下就是使用矩阵来计算的例子：

$$\begin{pmatrix}2&1&-1&8\\-3&-1&2&-11\\-2&1&2&-3\end{pmatrix}$$

该矩阵为增广矩阵，跟着以上的方法来运算，这个矩阵可以转变为以下的样子：

$$\begin{pmatrix}2&1&-1&8\\0&\frac{1}{2}&\frac{1}{2}&1\\0&0&-1&1\end{pmatrix}$$

这矩阵叫做行梯阵式。

最后，可以利用同样的算法产生以下的矩阵，便可把所得出的解或其限制简明地表示出来：

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

最后这矩阵叫做简化行梯阵式，亦是高斯-若尔当消元法指定的步骤。

笔记 以上高斯消元法的内容基本是回忆求解线性方程的步骤，是我们很熟悉的东西。在线性代数中比较重要的就是将之前所说的“第二行减去第一行的 3 倍”这种操作条例变为矩阵化的数学语言。

2.3 消元矩阵

2.3.1 行向量与矩阵乘法

上面的消元法是从简单的变换角度介绍了消元的具体操作，接下来我们需要用矩阵来表示变换的步骤，这也十分有必要，因为这是一种“系统地”变换矩阵的方法。

矩阵运算的核心内容就是对“行”或者“列”进行独立操作。

如前一节课“列图像”部分所言，系数矩阵乘以未知数向量，相当于对系数矩阵的列向量进行线性组合。

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix}\begin{bmatrix}3\\4\\5\end{bmatrix} = 3\underset{\text{列 1}}{\begin{bmatrix}*\\*\\*\end{bmatrix}} + 4\underset{\text{列 2}}{\begin{bmatrix}*\\*\\*\end{bmatrix}} + 5\underset{\text{列 3}}{\begin{bmatrix}*\\*\\*\end{bmatrix}} = \begin{bmatrix}8\\8\\8\end{bmatrix}$$

笔记 矩阵与列向量的乘积，就是矩阵各列的线性组合

这并不能解决我们现在的问题，因为在消元过程中我们用到的是行变换，那么我们考虑这个问题，行向量与矩阵的乘积是什么？

$$\begin{bmatrix}1 & 2 & 7\end{bmatrix}\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \begin{array}{l} 1\,\underset{\text{行 1}}{\begin{bmatrix} * & * & * \end{bmatrix}} \\ + \\ 2\,\underset{\text{行 2}}{\begin{bmatrix} * & * & * \end{bmatrix}} \\ + \\ 7\,\underset{\text{行 3}}{\begin{bmatrix} * & * & * \end{bmatrix}} \end{array} = \begin{bmatrix}\otimes & \otimes & \otimes\end{bmatrix}$$

笔记 行向量与矩阵的乘积，就是矩阵各行的线性组合

2.3.2 消元矩阵介绍

好的，接下来是重点。学会了行向量与矩阵之间的乘法，我们就可以使用行向量对矩阵的行做操作了。所谓消元矩阵，就是将消元过程中的行变换转化为矩阵之间的乘法形式。

首先我们要知道，

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

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

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

此时将 $\begin{bmatrix}1 & 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0 & 1\end{bmatrix}$ 构成一个矩阵：单位矩阵 $\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$，我们很

容易验证，这样得单位矩阵与矩阵相乘不改变矩阵。消元矩阵就是它得变形形式。可以看如下示例：

$$\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}$$

首先明确：

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix}$$

我们消元过程是将第一行乘以 -3 加到第二行，这是对第二行的操作，那么就从单位阵的第二行着手

$$\begin{bmatrix} -3 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \end{bmatrix}$$

从而得出：

$$\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 3 & 8 & 1 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}$$

所以，这一步得消元矩阵就是 $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$，计为 E_{21}, 意义是将矩阵 A 中 2 行 1 列 (2,1) 位置变为零得消元矩阵。

矩阵消元的第二步是完成矩阵 E_{21} 的第三行减去第二行的 2 倍，通过左乘矩阵 E_{32} 来实现这一过程。

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 5 \end{bmatrix}$$

这一步的消元矩阵就是 $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$, 计为 E_{32}，此时得到最终矩阵为一个阶梯型矩阵。

3×3 矩阵的消元本来应该分三步完成，最终得到 $E_{32}E_{21}$。所以结果变为 $E_{32}(E_{21A}) = U$，因为矩阵运算符合结合律，也可写作 $(E_{32}E_{21})A = U$。可以记作 $EA = U$。

方程 $Ax = b$ 的解也满足方程 $Ux = EAx = Eb = c$，因此我们将问题转化为 $Ux = c$。

笔记 核心：求解消元矩阵就是从单位阵 I 入手，按照 A 每次变换得消元步骤做 I 矩阵，能分别得到 $E_{\text{某行某列}}$，最后累计 E 即可。

2.3.3 行变换和列变换

有了上面消元矩阵的启发，不难得到，能够交换 2×2 矩阵中两行的矩阵为：

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

而交互 2×2 矩阵中两列的矩阵为：

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

笔记 左乘等于行变换，右乘等于列变换

2.4 逆矩阵初探

可以说我们学会了消元矩阵，就相当于我们可以用矩阵乘法对一个矩阵进行任何变化了，那么我们考虑一个反过程，即我们把一个消元结束的矩阵 U 如何变为未经消元的矩阵 A 呢？

答案就是乘上一个逆矩阵。

消元矩阵之逆矩阵的实施效果就是抵消原矩阵的消元操作。消元矩阵实现了对原矩阵 $A=\begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & -2 \\ 0 & 4 & 1 \end{bmatrix}$ 的操作，使第二行行向量 $[3,8,1]$ 减掉了第一行 $[1,2,1]$ 的 3 倍变为 $[0,2,-2]$，则逆向操作就应该是把现在的第二行行向量 $[0,2,-2]$ 加上第一行 $[1,2,1]$ 的 3 倍，从而变回原来的第二行 $[3,8,1]$。所以对于

$$E_{21}=\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

有

$$E_{21}^{-1}=\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

满足

$$E_{21}^{-1}E_{21}=I\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

笔记 这一个章节，我们从矩阵消元的角度，介绍解方程的通用做法，并介绍了消元矩阵，使我们从矩阵乘法层面理解了消元的过程，并延伸了消元矩阵的应用：就是基于单位阵 I 的变化，对矩阵 A 进行行列变换的过程。

第三章 矩阵的乘法和逆矩阵

3.1 Overview(概述)

前面介绍了向量和矩阵的乘法，这一节我们要介绍一下两个矩阵之间的乘法。并讨论逆矩阵存在的条件。最后再介绍求解逆矩阵的方法。

3.2 矩阵乘法 Matrix multiplication

我们通过四种方法讨论如何使矩阵 A 与 B 相乘得到矩阵 C。其中 A 为 $m \times n$(m 行 n 列) 矩阵，而 B 为 $n \times p$ 矩阵，则 C 为 $m \times p$ 矩阵，记 c_{ij} 为矩阵 C 中第 i 行第 j 列的元素。

3.2.1 标准方法 (行乘以列)

矩阵乘法的标准计算方法是通过矩阵 A 第 i 行的行向量和矩阵 B 第 j 列的列向量点积得到 c_{ij}。

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \ldots\ldots$$

笔记 矩阵运算怎么这么奇怪，只有同样型号的矩阵（指同为 mn）可以做加法，但是不同型号的矩阵却可以做乘法，话说这不同型号的矩阵为啥要做这种复杂的乘法？数乘一个矩阵为啥其中每一个元素都要乘以这个数，而行列式则不是？”

A.D. 亚历山大洛夫编著的《数学：它的内容，方法和意义》第三卷中找到了答案。先补一下矩阵的加法和数乘规则：

$$\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -3 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 0 & 1 \\ 1 & 6 \end{bmatrix} \quad 2\begin{bmatrix} 2 & 1 \\ 3 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & 0 \\ 0 & 8 \end{bmatrix}$$

本质上矩阵的代数运算就是线性方程组的运算：

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots\ldots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{cases}$$

可以简写做矩阵形式 $y = Ax$,A 为 $m \times n$ 矩阵

$$\begin{cases} z_1 = b_1x_1 + \ldots + b_1x_2 \\ z_n = b_2x_1 + \ldots + b_nx_n \end{cases}$$

可以简写做矩阵形式 $z = Bx$,B 为 $m \times n$ 矩阵

$$\begin{cases} y_1 + z_1 = (a_{11} + b_{11})\, x_1 + \ldots + (a_{1n} + b_{1n})\, x_n \\ \ldots\ldots \\ y_m + z_m = (a_{m1} + b_{m1})\, x_1 + \ldots + (a_{mn} + b_{mn})\, x_n \end{cases}$$

$y+z=(A+B)x$ 所以矩阵加法运算规则就, 只能形状相同的矩阵进行加和

$$\begin{cases} \alpha y_1 = \alpha a_{11}x_1 + \ldots + \alpha a_{1n}x_n \\ \ldots\ldots \\ \alpha y_m = \alpha a_{m1}x_1 + \ldots + \alpha a_{mn}x_n \end{cases}$$

$\alpha y = \alpha Ax$ 数乘要在矩阵每个元素上乘以 α

而矩阵乘法可以视为给线性方程组做变量替换

$y = Ax$,A 为 $m \times n$ 矩阵

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1n}x_n \\ \ldots\ldots \\ y_m = a_{m1}x_1 + \ldots + a_{mn}x_n \end{cases}$$

$w = Dy$,D 为 $k \times m$ 矩阵

$$\begin{cases} w_1 = d_{11}y_1 + \ldots + d_{1m}y_m \\ \ldots\ldots \\ w_k = d_{k1}y_1 + \ldots + d_{km}y_m \end{cases}$$

将左侧 y1 至 ym 的表达式代入右侧的方程，则有

$$\begin{cases} w_1 = d_{11}\left(a_{11}x_1 + \ldots + a_{1n}x_n\right) + \ldots + d_{1m}\left(a_{m1}x_1 + \ldots + a_{mn}x_n\right) \\ \ldots\ldots \\ w_k = d_{k1}\left(a_{11}x_1 + \ldots + a_{1n}x_n\right) + \ldots + d_{km}\left(a_{m1}x_1 + \ldots + a_{mn}x_n\right) \end{cases}$$

重新整理一下

$$\begin{cases} w_1 = \left(d_{11}a_{11} + \ldots + d_{1m}a_{m1}\right)x_1 + \ldots + \left(d_{11}a_{1n} + \ldots + d_{1m}a_{mn}\right)x_n = c_{11}x_1 + \ldots + c_{1n}x_n \\ \ldots\ldots \\ w_k = \left(d_{k1}a_{11} + \ldots + d_{kn}a_{m1}\right)x_1 + \ldots + \left(d_{k1}a_{1n} + \ldots + d_{km}a_{mn}\right)x_n = c_{k1}x_1 + \ldots + c_{bn}x_n \end{cases}$$

$w = Dy = DAx = Cx$ $C = DA$ 为 $k \times n$ 矩阵，从上式可得到 c_{ij} 的公式，即矩阵的乘法运算规则，可以看到这是变量代换的结果。所以矩阵乘法中两矩阵形状可以不同，但要求左侧矩阵的列数要等于右侧矩阵的行数，即 $w = Dy$ 中 y 的分量个数要等于 $y = Ax$ 中 y 的分量数，不如此则线性方程的变量代换无法进行。

3.2.2 列操作

列操作是指矩阵 C 的第 j 列是通过矩阵 A 乘以矩阵 B 第 j 列的列向量得到的。这表明矩阵 C 的列向量是矩阵 A 列向量的线性组合，组合的“权”就是矩阵 B 第 j 列的各个分量。

$$AB = A\begin{bmatrix} b_1 & b_2 & \ldots & b_p \end{bmatrix} = \begin{bmatrix} Ab_1 & Ab_2 & \ldots & Ab_p \end{bmatrix}$$

上式中 b_i，Ab_i 均为列向量，其中

$$Ab_j = \begin{bmatrix} | & | & | & | & | \\ a_1 & a_2 & \cdots & \cdots & a_n \\ | & | & | & | & | \end{bmatrix}\begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ \vdots \\ b_{nj} \end{bmatrix} = b_{1j}\begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} + b_{2j}\begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} + \cdots + b_{nj}\begin{bmatrix} | \\ a_n \\ | \end{bmatrix}$$

为矩阵 A 中列向量的线性组合

3.2.3 行操作

行操作是指矩阵 C 的第 i 行是通过矩阵 A 的第 i 行乘以矩阵 B 得到的。这表明矩阵 C 的行向量是矩阵 B 行向量的线性组合。

$$AB=\begin{bmatrix}a_1\\a_2\\\vdots\\\vdots\\a_m\end{bmatrix}B=\begin{bmatrix}a_1B\\a_2B\\\vdots\\\vdots\\a_mB\end{bmatrix}$$

此处 a_i 和 a_iB 为行向量

$$\begin{bmatrix}1&2&7\end{bmatrix}\begin{bmatrix}\cancel{*\ \ *\ \ *}\\\cancel{*\ \ *\ \ *}\\\cancel{*\ \ *\ \ *}\end{bmatrix}=\begin{array}{l}1\ [\cancel{*\ \ *\ \ *}]\\\quad+\ \text{行 1}\\2\ [\cancel{*\ \ *\ \ *}]=[\otimes\otimes\otimes]\\\quad+\ \text{行 2}\\7\ [\cancel{*\ \ *\ \ *}]\\\quad\ \ \ \text{行 3}\end{array}$$

笔记 差不多就是这样了

$$\begin{bmatrix}a&a&a&a\\b&b&b&b\\c&c&c&c\end{bmatrix}\begin{bmatrix}?&?&?&?\\?&?&?&?\\?&?&?&?\\?&?&?&?\end{bmatrix}=\begin{bmatrix}[\ a&a&a&a\ ]&B\\[\ b&b&b&b\ ]&B\\[\ c&c&c&c\ ]&B\end{bmatrix}=C$$

3.2.4 列乘以行

矩阵 A 的第 k 列是一个 $m\times1$ 的向量，而矩阵 B 第 k 行是一个 $1\times p$ 的向量，两向量相乘会得到一个矩阵 C_k，将所有的 n 个矩阵相加记得到 C。

$$AB=\sum_{k=1}^{n}\begin{bmatrix}a_{1k}\\\vdots\\a_{mk}\end{bmatrix}\begin{bmatrix}b_{k1}&\cdots&b_{kp}\end{bmatrix}$$

可以从矩阵乘法对加法的分配率推导出来。

这里我们通过一个例子来看：

求解：

$$\begin{bmatrix}2&7\\3&8\\4&9\end{bmatrix}\begin{bmatrix}1&6\\0&0\end{bmatrix}$$

这里使用列乘以行

$$\begin{bmatrix}2&7\\3&8\\4&9\end{bmatrix}\begin{bmatrix}1&6\\0&0\end{bmatrix}=\begin{bmatrix}2\\3\\4\end{bmatrix}\begin{bmatrix}1&6\end{bmatrix}+\begin{bmatrix}7\\8\\9\end{bmatrix}\begin{bmatrix}0&0\end{bmatrix}=\begin{bmatrix}2&12\\3&18\\4&24\end{bmatrix}$$

注意这里每一次都是用列向量与行向量相乘得到一个矩阵，而每次得到的矩阵都是有特点的，比

如 $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$, 这其中得到的矩阵 $\begin{bmatrix} 2 & 12 \\ 3 & 18 \\ 4 & 24 \end{bmatrix}$ 每一列都在和向量 $\begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ 同向，可以说列向量都在 $\begin{bmatrix} 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}$ 这条直线上，列空间是一条直线。同理，行向量都在 $\begin{bmatrix} 1 & 6 \end{bmatrix}$ 这条直线上，行空间 (矩阵行所有可能的线性组合) 是一条直线.

3.2.5 分块矩阵

分块乘法就是宏观上的矩阵乘法，比如现在有一个 $50 * 50$ 的矩阵与 $50 * 50$ 矩阵相乘，一个一个进行运算很麻烦，尤其是如果矩阵在某一区域上有一定的性质，那么我们可以将其分块，如：

$$\begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} = \begin{bmatrix} C_1 & C_2 \\ C_3 & C_4 \end{bmatrix}$$

在数学的矩阵理论中，一个分块矩阵或是分段矩阵就是将矩阵分割出较小的矩形矩阵，这些较小的矩阵就称为区块。换个方式来说，就是以较小的矩阵组合成一个矩阵。分块矩阵的分割原则是以水平线和垂直线进行划分。分块矩阵中，位在同一行（列）的每一个子矩阵，都拥有相同的列数（行数）。

通过将大的矩阵通过分块的方式划分，并将每个分块看做另一个矩阵的元素，这样之后再参与运算，通常可以让计算变得清晰甚至得以大幅简化。例如，有的大矩阵可以通过分块变为对角矩阵或者是三角矩阵等特殊形式

如下矩阵

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

可以分成四个 2×2 区块

$$P_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, P_{12} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, P_{21} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}, P_{22} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

分块后的矩阵可以写作

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

一个分块的矩阵乘法可以仅用包含算符的子矩阵来表述。给定一个 $(m \times p)$ 矩阵 $\mathbf{A}$ 有 q 行 s 列

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{q1} & A_{q2} & \cdots & A_{qs} \end{bmatrix}$$

另外一个 $(p\times n)$ 矩阵 $\mathbf{B}$ 有 s 行且 r 列

$$B=\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ B_{21} & B_{22} & \cdots & B_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{sr} \end{bmatrix}$$

矩阵乘积

$$C=AB$$

可被分成块来计算，矩阵 $\mathbf{C}$ 是 $(m\times n)$ 的矩阵有 q 行 r 列，你的矩阵 $\mathbf{C}$ 中的分割矩阵可以在乘法中被相乘:

$$C_{\alpha\beta}=\sum_{\gamma=1}^{s}A_{\alpha\gamma}B_{\gamma\beta}$$

3.3 逆矩阵 Inverse

之前介绍过一点逆矩阵的定义，对于一个方阵 A, 如果 A 可逆，就有这样一个 A^{-1} 使

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

如果 A 是非方阵，左侧的 A^{-1} 与右侧的 A^{-1} 不可能相同，因为这时左右侧 A^{-1} 形状一定不相同，那么这就违背了我们说的“一个” A^{-1} 。

逆矩阵（inverse matrix），又称反矩阵。在线性代数中，给定一个 n 阶方阵 $\mathbf{A}$，若存在一 n 阶方阵 $\mathbf{B}$ ，使得 $\mathbf{AB}=\mathbf{BA}=\mathbf{I}_n$，其中 $\mathbf{I}_n$ 为 n 阶单位矩阵，则称 $\mathbf{A}$ 是可逆的，且 $\mathbf{B}$ 是 $\mathbf{A}$ 的逆矩阵，记作 $\mathbf{A}^{-1}$。

只有方阵 $n\times n$ 的矩阵才可能有逆矩阵。若方阵 $\mathbf{A}$ 的逆矩阵存在，则称 $\mathbf{A}$ 为非奇异方阵或可逆方阵。

与行列式类似，逆矩阵一般用于求解联立方程组。

再举一个没有逆的矩阵：$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ 学习过行列式的同学一下子就可以算出来，这个矩阵对应行列式为 0, 所以矩阵不可逆。但是本门课程并不是从行列式开始介绍，那么我们换一种看法。我们看到这个矩阵中含两个列向量 $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ 这两个向量，它们之间互成倍数，也就是说这两个向量之一对其线性组合无意义，那么这个 A 不可能有逆。换句话说：

<u>若存在非零向量 x, 使得 $Ax=0$, 那么 A 就不可能有逆矩阵</u>

为什么呢？这样，如果 A 有逆，在 $Ax=0$ 这个等式两端同时乘上 A^{-1} 就有：

$$A^{-1}Ax=Ix=\text{零向量}$$

而 Ix 不可能是零向量，自相矛盾。所以此时 A 没有逆矩阵再看看矩阵 $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$ 由于两个列向量线性相关，一定有一个 x, 使得 $Ax=0$, 如 $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$。所以此矩阵不可逆。

3.4 矩阵逆的求解

3.4.1 伴随矩阵法

如果矩阵 A 可逆, 则 $A^{-1}=\frac{A^*}{|A|}$ 其中 $A^*=adj(A)$ 是 A 的伴随矩阵。

注意：A^* 中元素的排列特点是 A^* 的第 k 列元素是 A 的第 k 行元素的代数余子式。要求得 A^* 即为求解 A 的余因子矩阵的转置矩阵。

3.4.2 初等变换法

如果矩阵 A 和 B 互逆，则 $AB=BA=I$。由条件 $AB=BA$ 以及矩阵乘法的定义可知，矩阵 A 和 B 都是方阵。再由条件 $AB=I$ 以及定理“两个矩阵的乘积的行列式等于这两个矩阵的行列式的乘积”可知，这两个矩阵的行列式都不为 0。也就是说，这两个矩阵的秩等于它们的级数（或称为阶，也就是说，A 与 B 都是 $n\times n$ 方阵，且 $rank(A)=rank(B)=n$. 换而言之, A 与 B 均为满秩矩阵）。换句话说，这两个矩阵可以只经由初等行变换，或者只经由初等列变换，变为单位矩阵。

因为对矩阵 A 施以初等行变换（初等列变换）就相当于在 A 的左边（右边）乘以相应的初等矩阵，所以我们可以同时对 A 和 I 施以相同的初等行变换（初等列变换）。这样，当矩阵 A 被变为 I 时，I 就被变为 A 的逆阵 B。

这里我们使用初等变换法来求解矩阵，其实求逆矩阵就是解方程组的过程, 举例说明:

求解 $\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}$ 的逆矩阵

$$\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix}=I=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

从列向量的角度看来，得到两个方程:

$$\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}\begin{bmatrix}a\\ c\end{bmatrix}=\begin{bmatrix}1\\ 0\end{bmatrix}$$

$$\begin{bmatrix}1 & 3\\ 2 & 7\end{bmatrix}\begin{bmatrix}b\\ d\end{bmatrix}=\begin{bmatrix}0\\ 1\end{bmatrix}$$

解这个方程就行了，但是这样做低阶矩阵还好，高阶矩阵计算量未免太大了。所以这里可以使用，高斯-若尔当消元法

3.5 补充内容

3.5.1 余子式

在线性代数中，一个矩阵 A 的余子式（又称余因式，英语：minor）是指将 A 的某些行与列去掉之后所余下的方阵的行列式。相应的方阵有时被称为余子阵。

将方阵 A 的一行与一列去掉之后所得到的余子式可用来获得相应的代数余子式（英语：cofactor），后者在可以通过降低多阶矩阵的阶数来简化矩阵计算，并能和转置矩阵的概念一并用于逆矩阵计算。

不过应当注意的是，余子式和代数余子式两个概念的区别。在数值上，二者的区别在于，余子式只计算去掉某行某列之后剩余行列式的值，而代数余子式则需要考虑去掉的这一个元素对最

后值正负所产生的影响。

3.5.1.1 定义

设 A 为一个 $m\times n$ 的矩阵，k 为一个介于 1 和 m 之间的整数，并且 $k\leq n$。A 的一个 k 阶子式是在 A 中选取 k 行 k 列之后所产生的 k^2 个交点组成的方块矩阵的行列式。

A 的一个 k 阶余子式是 A 去掉了 k 行与 k 列之后得到的 $(m-k)\times(n-k)$ 矩阵的行列式。

由于一共有 $\binom{m}{k}$ 种方法来选择该保留的行，有 $\binom{n}{k}$ 种方法来选择该保留的列，因此 A 的 k 阶余子式一共有 $\binom{m}{k}\cdot\binom{n}{k}$ 个。

如果 $m=n$，那么 A 关于一个 k 阶子式的余子式，是 A 去掉了这个 k 阶子式所在的行与列之后得到的 $(n-k)\times(n-k)$ 矩阵的行列式，简称为 A 的 k 阶余子式。

3.5.1.2 代数余子式定义

一个矩阵 A 的 (i,j) 代数余子式：C_{ij} 是指 A 的 (i,j) 余子式 M_{ij} 与 $(1)^{i+j}$ 的乘积：

$$C_{ij}=(-1)^{i+j}M_{ij}$$

$n\times n$ 的方块矩阵 A 关于第 i 行第 j 列的余子式 M_{ij} 是指 A 中去掉第 i 行第 j 列后得到的 $n-1$ 阶子矩阵的行列式。有时可以简称为 A 的 (i,j) 余子式。

3.5.1.3 例子

对矩阵 $\begin{pmatrix}1&4&7\\3&0&5\\-1&9&11\end{pmatrix}$ 要计算代数余子式 C_{23}。首先计算余子式 M_{23}，也就是原矩阵去掉第 2 行和第 3 列后的子矩阵的行列式：$\begin{vmatrix}1&4&\square\\\square&\square&\square\\-1&9&\square\end{vmatrix}$，即 $\begin{vmatrix}1&4\\-1&9\end{vmatrix}=9-(-4)=13$

因此，C_{23} 等于 $(-1)^{2+3}M_{23}=-13$

3.5.2 伴随矩阵

在线性代数中，一个方形矩阵的伴随矩阵（英语：adjugate matrix）是一个类似于逆矩阵的概念。如果矩阵可逆，那么它的逆矩阵和它的伴随矩阵之间只差一个系数。然而，伴随矩阵对不可逆的矩阵也有定义，并且不需要用到除法。

设 R 是一个交换环，A 是一个以 R 中元素为系数的 $n\times n$ 的矩阵。A 的伴随矩阵可按如下步骤定义：

定义 3.1. 余子式

A 关于第 i 行第 j 列的余子式（记作 M_{ij}）是去掉 A 的第 i 行第 j 列之后得到的 $(n-1)\times(n-1)$ 矩阵的行列式。 ♣

定义 3.2. 代数余子式

A 关于第 i 行第 j 列的代数余子式是：

$$C_{ij}=(-1)^{i+j}M_{ij}$$

♣

定义 3.3. 余子矩阵

A 的余子矩阵是一个 $n \times n$ 的矩阵 C，使得其第 i 行第 j 列的元素是 A 关于第 i 行第 j 列的代数余子式

引入以上的概念后，可以定义：

定义 3.4. 余子矩阵

矩阵 A 的伴随矩阵是 A 的余子矩阵的转置矩阵：

$$\mathrm{adj}(A) = C^T$$

也就是说，A 的伴随矩阵是一个 $n \times n$ 的矩阵（记作 $adj(A)$），使得其第 i 行第 j 列的元素是 A 关于第 j 行第 i 列的代数余子式。简言之，伴随矩阵就是把原来矩阵每一行的代数余子式竖着写：

$$[\mathrm{adj}(A)]_{ij} = C_{ji}$$

3.5.3 例子

一个 2×2 矩阵 $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ 的伴随矩阵是

$$\mathrm{adj}(A) = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

对于 3×3 的矩阵，情况稍微复杂一点：

$$a = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

其伴随矩阵是：

$$\mathrm{adj}(A) = \begin{pmatrix} +\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & +\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & +\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ +\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & +\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{pmatrix}$$

其中

$$\begin{vmatrix} a_{im} & a_{in} \\ a_{jm} & a_{jn} \end{vmatrix} = \det\begin{pmatrix} a_{im} & a_{in} \\ a_{jm} & a_{jn} \end{pmatrix}$$

第四章 矩阵 A 的 LU 分解

4.1 Overview 概述

在上一篇 03-乘法和逆矩阵中，我们在末尾介绍了逆矩阵的一些内容，今天我们首先完善之前讲到的逆矩阵的内容，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。

4.2 逆矩阵性质补充

1. $\left(A^{-1}\right)^{-1} = A$
2. $(\lambda A)^{-1} = \frac{1}{\lambda} \times A^{-1}$
3. $(AB)^{-1} = B^{-1}A^{-1}$
4. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$
5. $\det\left(A^{-1}\right) = \frac{1}{\det(A)}$

4.3 转置矩阵基础

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:

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

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

4.3.1 转置矩阵的性质

1. $\left(A^T\right)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(AB)^T = B^T A^T$

   注意因子反转的次序。以此可推出方块矩阵 A 是可逆矩阵，当且仅当 AT 是可逆矩阵，在这种情况下有 $(A1)^T = (AT)^1$。相对容易的把这个结果扩展到矩阵相乘的一般情况，可得出 $(ABC...XYZ)^T = Z^T Y^T X^T ... C^T B^T A^T$。
4. $(cA)^T = cA^T$
5. $\det\left(A^T\right) = \det(A)$
6. $a \cdot b = a^T b$
7. 如果 A 只有实数元素，则 $A^T A$ 是正半定矩阵。
8. 如果 A 是在某个域上，则 A 相似于 A^T。

4.3.2 特殊转置矩阵

其转置等于自身的方块矩阵叫做对称矩阵；就是说 A 是对称的，如果

$$A^T = A_0$$

其转置也是它的逆矩阵的方块矩阵叫做正交矩阵；就是说 G 是正交的，如果

$$GG^T = G^T G = I_n$$

其转置等于它的负矩阵的方块矩阵叫做斜对称矩阵；就是 A 是斜对称的，如果

$$A^T = -A$$

复数矩阵 A 的共轭转置，写为 A^H，是 A 的转置后再取每个元素的共轭复数:

$$A^H = (\bar{A})^T = \overline{(A^T)}$$

4.3.3 转置矩阵的逆矩阵

$\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$ 在 $AA^{-1} = I$ 两边取转置，并应用积的转置公式，得到 $\left(A^{-1}\right)^T A^T = I$ 根据逆矩阵定义得到 $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

4.4 A 的 LU 分解

我们熟悉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到与之一样的消元效果。所以，现在的假设有可逆矩阵 A ，若有矩阵 A 转变上三角阵 U,。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？我们通过下面一道例题来探讨一下。

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

$$A = (E_{21})^{-1} (E_{31})^{-1} (E_{32})^{-1} U$$

计算出各个矩阵：

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

直接带入计算，

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

笔记 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

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

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

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 L，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。

4.5 置换矩阵

百度百科上对于置换矩阵的定义：在数学上，特别是在矩阵理论中，置换矩阵是一个方形二进制矩阵，它在每行和每列中只有一个 1，而在其他地方则为 0。

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不是 1 时，我们就需要用行变换将主元位置换回 1。

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：

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

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。

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。

第五章 转置，置换，向量空间 R

5.1 置换

置换矩阵另一个优点就是可逆，因为置换矩阵各行还原后可以得到单位矩阵。而且对于置换矩阵 P, 有 $PP^T = I$, 即 $P^{-1} = P^T$ 这个性质其实很好理解，首先明确，P 是置换矩阵，所以 P 的每个列向量中只有一个分量是 L 其余分量均为 0。而既然要求 $PP^{-1} = I$ 那就说明 P 中每一行的行向量与 P^{-1} 中每一列的列向量的数量积为 1, 意味着 P 中每一行的行向量与 P^{-1} 中每一列的列向量中分量 1 出现位置相同，这就意味着 P 与 P^{-1} 沿对角线对称，所以 $P^{-1} = P^T$

5.1.1 置换矩阵的使用

在讲消元法的时候，主元位置为 0 是一件很让人头疼的事情，这时就需要考虑置换矩阵 P 来完成行交换，确保消元过程顺利进行。上一节我们学习 A = LU 分解时，我们没有考虑要交换行的过程，如果我们想写出更加普适的 LU 分解式的话，必须把行交换情况考虑进去，即:

$$PA = LU$$

先用行交换使得主元位置不为 0，行顺序正确。其后，再用 LU 分解。

5.2 转置矩阵

这部分在上一节已经描述过：之前简单介绍过转置矩阵，即

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

用符号来表示就是对 A 矩阵以及 A^T 矩阵中每一个元素，都有：

$$\left(A^T\right)_{ij} = A_{ji}$$

也就是说，转置矩阵中，行元素与列元素交换了，理解转置很简单。

5.3 对称矩阵

对称矩阵，顾名思义，就是对角线两侧元素对应相等的矩阵。或者说，对矩阵 A, 如果有：

$$A^T = A$$

那么矩阵 A 就是一个对称矩阵。

例如：

$$\begin{bmatrix} 3 & 1 & 7 \\ 1 & 2 & 9 \\ 7 & 9 & 4 \end{bmatrix}$$

如何得到对称矩阵呢？很简单，矩阵 A 与 A^T 相乘得到的方阵一定是对称短阵，因为我们从对称

矩阵的定义来看，取 $(A^TA)^T$ 根据转置的运算规律，可知，$(AB)^T = B^TA^T$，所以有：

$$(A^TA)^T = A^TA^{TT} = A^TA$$

所以任何的 A^TA, 转置任然是其本身，故为对称矩阵

5.4 向量空间

向量空间是现代数学中的一个基本概念。是线性代数研究的基本对象。

向量空间的一个直观模型是向量几何，几何上的向量及相关的运算即向量加法，标量乘法，以及对运算的一些限制如封闭性，结合律，已大致地描述了“向量空间”这个数学概念的直观形象。

在现代数学中，“向量”的概念不仅限于此，满足下列公理的任何数学对象都可被当作向量处理。譬如，实系数多项式的集合在定义适当的运算后构成向量空间，在代数上处理是方便的。单变元实函数的集合在定义适当的运算后，也构成向量空间，研究此类函数向量空间的数学分支称为泛函分析。

定义 5.1. 向量空间

给定域 F，F 上的向量空间 V 是一个集合，其上定义了两种二元运算：
向量加法 $+ : V \times V \to V$，把 V 中的两个元素 u 和 v 映射到 V 中另一个元素，记作 $u+v$；
标量乘法 $\cdot : F \times V \to V$ 把 F 中的一个元素 α 和 V 中的一个元素 u 变为 V 中的另一个元素，记作 $a \cdot u$。
V 中的元素称为向量，相对地，F 中的元素称为标量。
且满足 V 公理才构成一个向量空间 (对 F 中的任意元素 a、b 以及 V 中的任意元素 uvw 都成立)：

1. 向量加法的结合律 $u + (v + w) = (u + v) + w$
2. 向量加法的交换律 $u + v = v + u$
3. 向量加法的单位元, 存在一个叫做零向量的元素 $0 \in V$，使得对任意 $u \in V$ 都满足 $u + 0 = u$
4. 向量加法的逆元素, 对任意 $v \in V$ 都存在其逆元素 $v \in V$ 使得 $v + (v) = 0$
5. 标量乘法与标量的域乘法相容,$a(bv) = (ab)v$
6. 标量乘法的单位元, 域 F 存在乘法单位元 1 满足 $1v = v$
7. 标量乘法对向量加法的分配律,$a(u + v) = au + av$
8. 标量乘法对域加法的分配律,$(a + b)v = av + bv$

♣

前四个公理说明装备了向量加法的 V 是交换群，余下的四个公理应用于标量乘法。需要注意的是向量之间的加法“+”和标量之间的加法“+”是不一样的，标量与向量之间的标量乘法 • 和两个标量之间的乘法（域 F 中自带的乘法）也是不一样的。

5.4.1 基本性质

1. 零向量 0 是唯一的；
2. 对任意 $a \in F$，$a \cdot 0 = 0$；
3. 对任意 $u \in V$，$0 = 0 \cdot u$,0 是 F 的加法单位元。
4. 如果 $a \cdot u = 0$，则要么 $a = 0$，要么 $u = 0$。
5. 向量加法的逆向量 v 是唯一的，记作 v。$u + (v)$ 也可以写成 uv，两者都是标准的。

6. 对任意 $u \in V$，$u = 1 \cdot u$.
7. 对任意 $a \in F$ 以及 $u \in V$，$(a) \cdot u = (a \cdot u) = a \cdot (u)$.

5.4.2 概念化及额外结构

1. 一个实数或复数向量空间加上长度概念（就是范数）则成为赋范向量空间。
2. 一个实数或复数向量空间加上长度和角度的概念则成为内积空间。
3. 一个向量空间加上拓扑结构并满足连续性要求（加法及标量乘法是连续映射）则成为拓扑向量空间。
4. 一个向量空间加上双线性算子（定义为向量乘法）则成为域代数。

我们可以对向量进行所谓“线性运算”，即通过加和 $(v + w)$ 与数乘运算（$(3v)$ 得到向量的线性组合。向量空间对线性运算封闭，即空间内向量进行线性运算得到的向量仍在空间之内。R^2 即为向量空间，它是具有两个实数分量的所有向量（二维实向量）的集合。

例如：$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$，$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$，$\begin{bmatrix} \pi \\ e \end{bmatrix}$

向量 $\begin{bmatrix} a \\ b \end{bmatrix}$ 的图像是从原点出发到点 (a, b) 的箭头，其中第一分量 a 为横轴坐标，第二分量 b 为纵轴坐标。空间 R^2 的图像为整个 $x - y$ 平面。

所有向量空间必然包含零向量，因为任何向量数乘 0 或者加上反向量都会得到零向量，而因为向量空间对线性运算封闭，所以零向量必属于向量空间。

R^3 是向量空间，它是具有三个实数分量的所有向量的集合。

R^n 是向量空间，它是具有 n 个实数分量的所有向量的集合。

反例：R^2 中的第一象限则不是一个向量空间。

很明显，这部分空间无法满足“线性组合仍在空间中”的要求，比如数乘运算时，随便取个负数，向量就跑到第三象限去，脱离 D 空间范围内了。

5.4.3 子空间

线性子空间（或向量子空间）在线性代数和相关的数学领域中是重要的。在没有混淆于其他子空间的时候通常简称为“子空间”。

定义 5.2. 线性子空间

在线性代数和其他数学相关领域，一个线性子空间 (或向量子空间) U 是给定域 $\mathfrak{R}$ 向量空间 V 的一个子集，并且它还是 V 的加法子群，同时，在纯量乘下回到自身，那么，V 上运算在 U 上的限制导出 U 的向量空间结构，我们把 U 称为 V 上的向量（或线性）子空间。♣

定理 5.1. 线性子空间

设 V 是在域 K 上的向量空间，并设 W 是 V 的子集。则 W 是个子空间，当且仅当它满足下列三个条件:

零向量 0 在 W 中。

如果 u 和 v 是 W 的元素，则向量和 $u + v$ 是 W 的元素。

如果 u 是 W 的元素而 c 是来自 K 的标量，则标量积 cu 是 W 的元素。♡

包含于向量空间之内的一个向量空间称为原向量空间的一个子空间。例如用实数 c 数乘 R^2 空间中向量 v 所得到的向量集合就是 R^2 空间的一个子空间，其图像为二维平面上穿过原点的一

条直线，它对于线性运算封闭。

反例：R^2 中不穿过原点的直线就不是向量空间。子空间必须包含零向量，原因就是数乘 0 的到的零向量必须处于子空间中。

R^2 的子空间包括：

1. R^2 空间本身
2. 过原点的一条直线 (这是 R^2 空间中的一条直线，与 R^1 空间有区别)
3. 原点仅包含 0 向量 Z

R^3 的子空间包括：

1. R^3 空间本身
2. 过原点的一个平面 2 维
3. 过原点的一条直线 1 维
4. 原点仅包含 0 向量 Z

5.5 列空间 Column spaces 概述

给定矩阵 A，其列向量属于 R^3 空间，这些列向量和它们的线性组合张成了 R^3 空间中的一个子空间，即矩阵 A 的列空间 $C(A)$。

如果 $A = \begin{bmatrix} 1 & 3 \\ 2 & 3 \\ 4 & 1 \end{bmatrix}$，则 A 的列空间是 R^3 空间中包含向量 $\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$ 和 $\begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix}$ 并穿过原点的平面，空间内包含两向量的所有线性组合。

下面课程的任务就是在列空间和子空间的基础上理解 $Ax = b$。

第 六 章　列空间和零空间