MAT140 — 第16讲讲义

线性代数

上节课中，我们引入了矩阵和行列式。矩阵最早出现时，是一种方便整理线性方程组系数的方法，而行列式则给了我们一种从方阵中提取单个数字的方法。我们还看到，行列式并不只是某种"需要记住的代数公式"，而是描述了面积之类的几何量。

在这节课里，我们将完成对线性方程组的讨论，方法是稍微退后一步，从更宏观的角度来看我们已经做过的一切。事实上，在过去三节课的讨论中，我们一直在不断接触线性代数这一领域。线性代数研究向量、向量空间、矩阵、线性变换以及线性方程组。今天我们将开始更正式地介绍这个主题，把目前所见的所有内容串联成一个线性代数的叙事。一般来说，我们会重点关注这些内容背后的几何图景。

今天我们将：

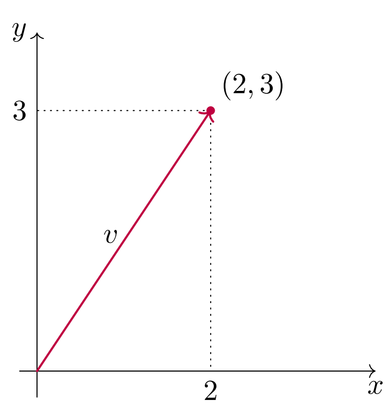
1. 定义一种叫做向量的新对象，并解释如何对它们进行运算。
2. 描述张成与线性组合，并引入标准基向量。
3. 定义矩阵加法与矩阵乘法，并解释为什么矩阵乘法是很特殊的。
4. 引入线性变换，并说明每一个这样的变换都可以由一个矩阵来描述。
5. 把行列式重新解释为面积缩放因子，并把矩阵与线性方程组的解集重新联系起来。

这节课可以看作是某种补充性的内容。我们不会对这些主题进行很深入的研究，而是主要介绍思想，并把它们和我们之前的直觉联系起来。对于那些想要进一步投入更技术性学科的同学来说，这节课可以视为对今后需要深入学习的思想的一次良好入门。对于那些并不打算继续学习比这更深数学内容的同学来说，这节课也可以被看作是本课程的数学顶峰，是我们在过去几节课中一直攀登的那座高山。当然，这两种看法并不互相排斥。

1 向量

1.1 什么是向量?

向量是一种同时具有大小和方向的数学对象。在二维平面中，我们通常把向量看成是一个以原点为起点的箭头。由于这个小箭头的尖端必须落在某个地方，所以我们可以把向量与点理解为彼此对应的一对对象。例如，点$(2,3)$ 确定了向量v，它可以画成从$(0,0)$ 到$(2,3)$ 的箭头：



因此，向量v 就是那个从原点出发并终止于点$(2,3)$ 的唯一箭头。“$x=2$” 和“$y=3$” 这两条信息唯一地确定了向量v，所以用某种有序方式把这两个数字写出来表示这个向量是很自然的。为了今后的方便，我们将使用竖直括号来表示向量，而不是水平括号。对于上面画出的向量v，我们写作：

$$v = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

点与向量的对应关系

在二维中，点与向量之间有一种自然的对应关系：点(x,y) 对应于向量

$$v = \begin{pmatrix} x \\ y \end{pmatrix}.$$

这里，v 可以解释为从原点指向点(x,y) 的有向箭头。

向量空间是一个由向量组成的集合，在其中两种基本运算是有意义的：

1. 向量可以相加；
2. 向量可以乘以实数，这些实数称为标量。

在这节课中，我们主要关注用实数进行缩放的二维和三维向量。数学上，这两个向量空间分别由$\mathbb{R}^2$ 和$\mathbb{R}^3$ 构成。现在我们来详细说明这些加法和乘法运算的本质。

1.2 向量运算

如果

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \qquad v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

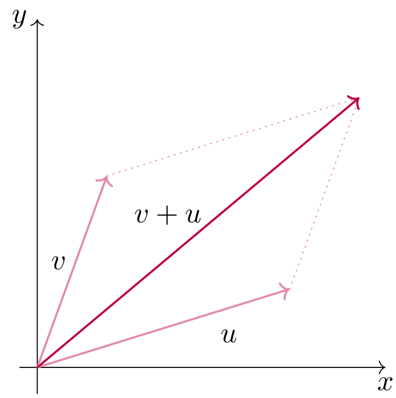
是一对向量，那么我们可以按分量相加，也就是说，把它们的x 分量彼此相加，再把它们的y 分量彼此相加：

$$u + v = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}.$$

同样地，如果a 是一个实数，那么标量乘法定义为

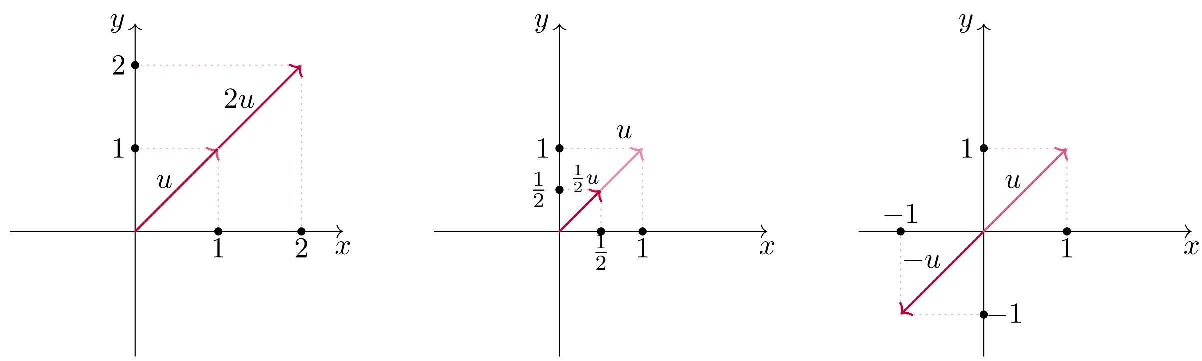
$$au = \begin{pmatrix} au_1 \\ au_2 \end{pmatrix}.$$

从几何上看，向量加法对应于把箭头首尾相接。这样会形成一个平行四边形，而两个向量的和就是从原点指向新产生顶点的那条对角线箭头：



观察上图可以看到，把u 的一个副本放到v 的箭头尖端，和把v 的一个副本放到u 的箭头尖端，都会得到平行四边形的同一个对顶点。这说明向量加法满足交换律，也就是说，$v+u=u+v$。

标量乘法则只是把一个向量拉长、缩短，或者反转它的方向。例如，考虑向量$u=\begin{pmatrix}1\\1\end{pmatrix}$。向量$2u$、$\frac{1}{2}u$ 和$-u$ 如下图所示：



正如你所见，乘以2 会把长度加倍，乘以$\frac{1}{2}$ 会把长度减半，而乘以-1 则保持长度不变，但会反转u 的方向。

现在我们给出一个两个向量相加的例子。设

$$u=\begin{pmatrix}1\\1\end{pmatrix}, \qquad v=\begin{pmatrix}-2\\3\end{pmatrix}.$$

向量$u+v$、$v-u$ 和$3u-v$ 都可以通过按分量进行这些运算来求出。我们有：

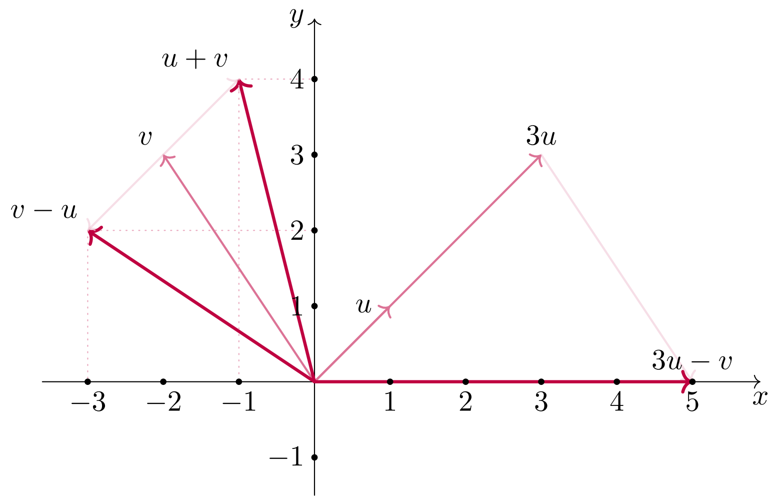
$$u+v=\begin{pmatrix}1\\1\end{pmatrix}+\begin{pmatrix}-2\\3\end{pmatrix}=\begin{pmatrix}-1\\4\end{pmatrix},$$

$$v-u=\begin{pmatrix}-2\\3\end{pmatrix}-\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}-3\\2\end{pmatrix},$$

并且

$$3u-v=3\begin{pmatrix}1\\1\end{pmatrix}-\begin{pmatrix}-2\\3\end{pmatrix}=\begin{pmatrix}3\\3\end{pmatrix}-\begin{pmatrix}-2\\3\end{pmatrix}=\begin{pmatrix}5\\0\end{pmatrix}.$$

从图形上看，这些向量如下所示：



1.3 向量的基本性质

向量运算的规则可以整理成向量空间的形式定义。这里我们不会强调完全抽象的表述，但下面这些性质是值得记住的。

向量的基本性质

对于向量u, v, w 和标量a, b，下面的性质成立：

1. 存在零向量0。
2. 加法交换律：$u+v=v+u$。
3. 加法结合律：$(u+v)+w=u+(v+w)$。
4. 加法逆元：$u+(-u)=0$。
5. 分配律：$a(u+v)=au+av$。
6. 分配律：$(a+b)u=au+bu$。
7. 标量乘法结合律：$(ab)u=a(bu)$。
8. 标量单位元：$1u=u$。

当然，这些性质都可以从几何上加以验证。例如，我们之前已经看到，把一个向量v 加到另一个向量w 上，对应于由v 和w 所构成平行四边形的对角线上的新向量。你以什么顺序把这两个向量相加并不重要，因为两种首尾相接的构造都会得到平行四边形的同一个对顶点。代数上，这一点也很容易看出来，因为向量加法只是把各个分量分别相加，因此它从实数那里继承了交换律：

$$u+v=\begin{pmatrix}u_1\\u_2\end{pmatrix}+\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}u_1+v_1\\u_2+v_2\end{pmatrix}\overset{(*)}{=}\begin{pmatrix}v_1+u_1\\v_2+u_2\end{pmatrix}=\begin{pmatrix}v_1\\v_2\end{pmatrix}+\begin{pmatrix}u_1\\u_2\end{pmatrix}=v+u,$$

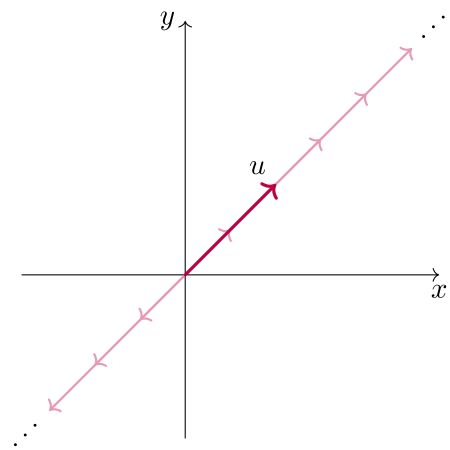
其中在步骤$(*)$ 中，我们只是用了实数加法满足交换律这一事实。

1.4 张成与线性组合

如果v 是一个非零向量，那么v 的所有标量倍数组成一条经过原点的直线。这个集合称为v 的张

成:

$$\mathrm{span}(v) = \{\text{所有}v\text{ 的标量倍数}\}.$$



更一般地，给定两个向量u 和v，任何形如

$$au + bv$$

的向量都称为u 和v 的一个线性组合。所有这类组合构成的集合为

$$\mathrm{span}(u, v) = \{au + bv : a, b \in \mathbb{R}\}.$$

如果u 和v 彼此是标量倍数，那么$\mathrm{span}(u, v)$ 仍然只是一条直线。如果它们指向真正不同的方向，那么它们的张成将铺满整个平面。

张成

在$\mathbb{R}^2$ 中，一个非零向量的张成是一条经过原点的直线。两个不平行向量的张成是整个平面。

1.5 基向量与坐标

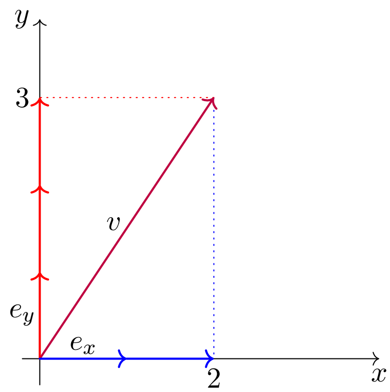
在平面中，有两个特别重要的向量:

$$e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

它们称为标准基向量。事实上，$\mathbb{R}^2$ 中的每一个向量都可以写成这两个向量的线性组合。例如:

$$v = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2e_x + 3e_y.$$

从图形上看，我们可以把这理解成一种运算，它和我们在二维空间中描点的方法非常相似:



利用这两个特殊向量，我们可以把任意向量的分量重新理解为：

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

在x 方向上的移动量（左/右）

在y 方向上的移动量（上/下）

练习1

设

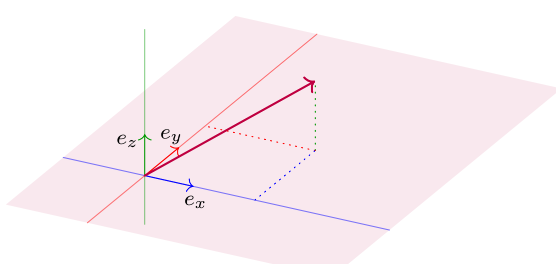
$$u = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \qquad v = \begin{pmatrix} -3 \\ 4 \end{pmatrix}.$$

计算：

1. $u + v$
2. $2u$
3. $3u - v$

1.6 理解更高维空间

如果我们有两个向量e_x 和e_y，那么我们就可以生成一个二维的向量平面。现在想象一下，我们把这个平面取出来，再添加另一个长度为1、从这个平面中朝另一个方向伸出去的向量：



我们把这个新向量记作e_z。从上图中你可以看出，这个新向量的张成方向与e_x 和e_y 都是完全垂直的。换句话说，e_z 让我们获得了另一个维度，向量有可能朝这个维度的方向指去。于是，我们可以把由e_x 和e_y 张成的二维平面理解为一种特殊情形，即z 方向的分量恰好等于零。一般来说，通

过改变任意给定向量中所包含的e_z 的多少，我们就可以把这个平面在第三个维度中上下移动。为了记录这额外的一点信息，我们可以用三个数字而不是之前的两个数字来表示三维向量。从我们的特殊基向量开始，我们写作：

$$e_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad e_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

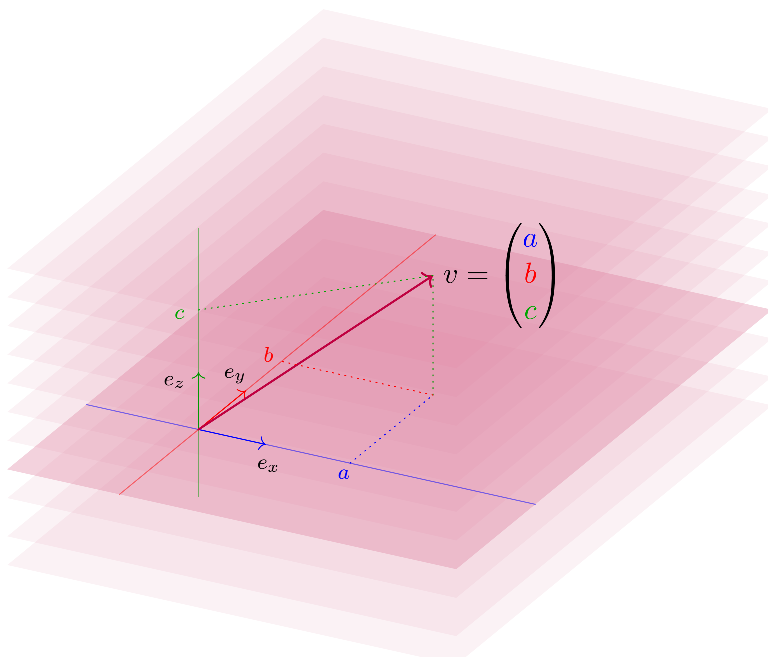
而任意向量

$$v = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

都可以写成

$$v = ae_x + be_y + ce_z.$$

从某种意义上说，我们可以把所得的空间想象为许多个$\mathrm{span}(e_x, e_y)$ 的副本一层一层叠在一起：



就像一条直线由无限多个点连续地排在一起组成一样，我们也可以把三维空间理解为无限多个二维空间连续地上下叠放而成。

尽管同样的概念可以继续推广到更高维，但我们对于向量行为的几何直觉很快就会失效。于是，我们不得不依赖向量的代数描述。从代数上说，一个n 维向量具有形式

$$v = a_1e_1 + a_2e_2 + \cdots + a_ne_n.$$

其中标准基向量为

$$e_1 = \begin{pmatrix}1\\0\\\vdots\\0\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\\\vdots\\0\\0\end{pmatrix}, \quad \cdots \quad e_{n-1} = \begin{pmatrix}0\\0\\\vdots\\1\\0\end{pmatrix}, \quad \text{以及} \quad e_n = \begin{pmatrix}0\\0\\\vdots\\0\\1\end{pmatrix}.$$

这n 个基向量分别指向n 个彼此垂直的不同方向。这个一般维数的向量空间记作$\mathbb{R}^n$。

练习2

1. 将$\begin{pmatrix}4\\-2\end{pmatrix}$ 写成$ae_x + be_y$ 的形式。
2. 将$\begin{pmatrix}1\\0\\5\end{pmatrix}$ 写成$ae_x + be_y + ce_z$ 的形式。
3. 从几何上描述$\text{span}\left(\begin{pmatrix}2\\1\end{pmatrix}\right)$。

2 矩阵

我们之前已经看到，矩阵是一个由实数构成的长方形数组。回忆一下，如果一个矩阵有m 行和n 列，那么我们就说它是一个$m \times n$ 矩阵，或者说这个矩阵的阶是$m \times n$。

现在我们将说明，矩阵本身也有属于它们自己的运算方式。

2.1 矩阵加法

如果我们有两个阶相同的矩阵，那么就可以按元素逐项相加。对于2×2 矩阵，加法看起来是这样的：

$$\begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix} + \begin{pmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{pmatrix} = \begin{pmatrix}a_{11}+b_{11} & a_{12}+b_{12}\\a_{21}+b_{21} & a_{22}+b_{22}\end{pmatrix}.$$

例如，

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix} + \begin{pmatrix}0 & 2\\2 & 1\end{pmatrix} = \begin{pmatrix}1 & 4\\5 & 4\end{pmatrix}.$$

我们一步一步地展示这个过程：

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

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}+\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1+0 & 2+2\\ & \end{pmatrix}=\begin{pmatrix}1 & 4\\ & \end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}+\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1+0 & 2+2\\3+2 & \end{pmatrix}=\begin{pmatrix}1 & 4\\5 & \end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}+\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1+0 & 2+2\\3+2 & 3+1\end{pmatrix}=\begin{pmatrix}1 & 4\\5 & 4\end{pmatrix}$$

需要强调的是，矩阵加法只有在两个矩阵具有相同阶时才有定义，否则两个矩阵中的元素无法"一一对应"，也就没有东西可以相加。

2.2 矩阵乘法

在合适的条件下，两个矩阵也可以相乘。不过，与矩阵加法不同，这种运算不是按元素逐项进行的。矩阵乘法要更微妙一些。对于2×2 矩阵，我们按如下方式相乘：

$$\begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}b_{11} & b_{12}\\b_{21} & b_{22}\end{pmatrix}=\begin{pmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{pmatrix}.$$

这里，我们取第一个矩阵的行，并将其有选择地与第二个矩阵的列相乘。例如，乘积左上角的元素由第一个矩阵的第一行与第二个矩阵的第一列相乘得到。当我们让一行与一列相乘时，必须注意元素的顺序不能弄错。

例如，我们计算：

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1\cdot 0+2\cdot 2 & 1\cdot 2+2\cdot 1\\3\cdot 0+3\cdot 2 & 3\cdot 2+3\cdot 1\end{pmatrix}=\begin{pmatrix}4 & 4\\6 & 9\end{pmatrix}.$$

为了看清这个过程是如何进行的，下面是逐步计算：

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1\cdot 0+2\cdot 2 & \\ & \end{pmatrix}=\begin{pmatrix}4 & \\ & \end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1\cdot 0+2\cdot 2 & 1\cdot 2+2\cdot 1\\ & \end{pmatrix}=\begin{pmatrix}4 & 4\\ & \end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1\cdot 0+2\cdot 2 & 1\cdot 2+2\cdot 1\\3\cdot 0+3\cdot 2 & \end{pmatrix}=\begin{pmatrix}4 & 4\\6 & \end{pmatrix}$$

$$\begin{pmatrix}1 & 2\\3 & 3\end{pmatrix}\cdot\begin{pmatrix}0 & 2\\2 & 1\end{pmatrix}=\begin{pmatrix}1\cdot 0+2\cdot 2 & 1\cdot 2+2\cdot 1\\3\cdot 0+3\cdot 2 & 3\cdot 2+3\cdot 1\end{pmatrix}=\begin{pmatrix}4 & 4\\6 & 9\end{pmatrix}$$

与矩阵加法不同，不同阶的矩阵有时也可以相乘，唯一的要求是左边矩阵的行长度必须与右边矩阵的列长度相匹配。

矩阵乘法的尺寸规则

一个$m \times n$ 矩阵可以和一个$n \times \ell$ 矩阵相乘，结果是一个$m \times \ell$ 矩阵。

2.3 零矩阵与单位矩阵

就像数字有0 和1 一样，矩阵通常也有特殊的中性元素。对于2×2 矩阵，零矩阵是

$$0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

并且对于每一个2×2 矩阵A，都有$A + 0 = A$。阶为2×2 的乘法单位矩阵是:

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

并且对于每一个2×2 矩阵A，都有

$$AI = IA = A.$$

2.4 矩阵运算的性质

在整个课程中，我们已经见过几种不同的运算：一开始我们简单讨论了实数的运算，而在课程较后部分，我们又开始讨论多项式的运算。对于后者，我们曾使用过这样一种叙事：多项式的行为在很多方面很像整数 $\mathbb{Z}$。从这个意义上说，我们曾指出，多项式全体并不真正和数字一样，但它们有相似之处。

与多项式不同，矩阵并不像普通数字那样“相似”。首先，并不是任意两个矩阵都可以相加或相乘，它们的阶必须以特定方式兼容才行。然而，即便假设这一点成立，例如我们只把注意力限制在所有2×2 的方阵上，矩阵运算仍然会表现出一些奇怪的现象。

矩阵运算的不寻常性质

对于矩阵A 和B:

1. 乘法不一定满足交换律，因此$AB \neq BA$;
2. 不是每个矩阵都有乘法逆元。换句话说，有些矩阵A 不存在逆矩阵A^{-1}，使得$AA^{-1} = I$。

为了说明矩阵乘法不一定可交换，考虑下面这一对2×2 矩阵:

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

那么，它们的乘积AB 和BA 不相等:

$$AB = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad BA = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

也就是说，$AB \neq BA$。

练习3

计算下列各项：

1. $\begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 3 & 2 \end{pmatrix}$
2. $\begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 1 & 5 \end{pmatrix}$
3. $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 0 & 3 \\ 1 & 1 \end{pmatrix}$
4. $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

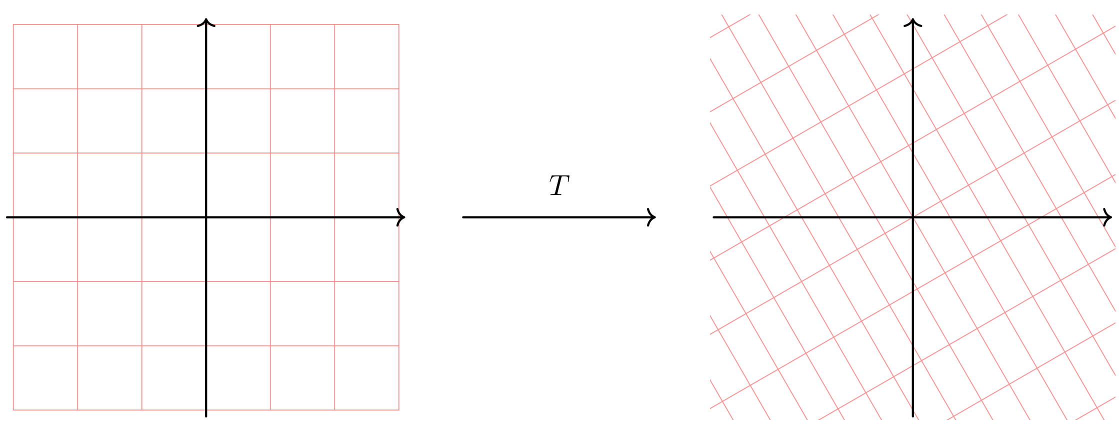
练习4

1. 一个2×3 矩阵与一个3×4 矩阵的乘积是否有定义？如果有，结果的大小是多少？
2. 一个2×3 矩阵与一个2×2 矩阵的乘积是否有定义？
3. 用文字解释，单位矩阵与另一个矩阵相乘时会起什么作用。

3 线性变换

3.1 几何思想

线性变换是一个把向量作为输入并产生向量作为输出的函数。从几何上看，线性变换就是那些保持空间“直线性”的映射：直线仍然变成直线，坐标平面中的方格网会变成另一张由直线构成的网格，而原点保持不动。例如，下图描绘了一个线性变换：

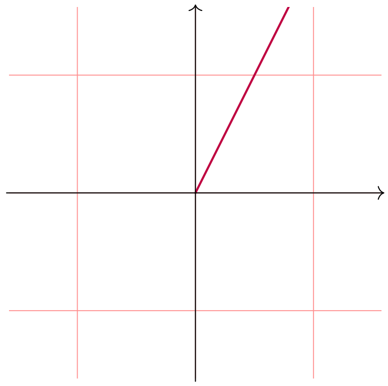


典型的例子包括：

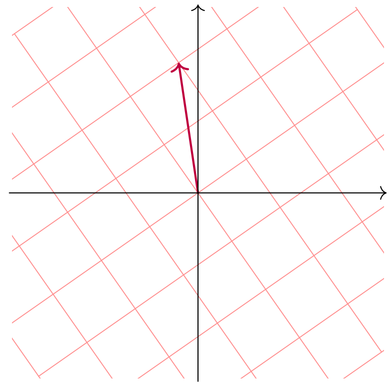
- 缩放，
- 旋转，
- 反射，
- 投影，

- 错切。

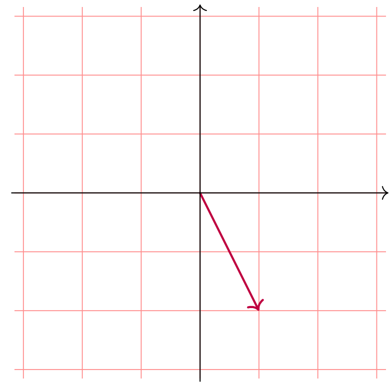
一个把直线弯成曲线的映射就不是线性的。



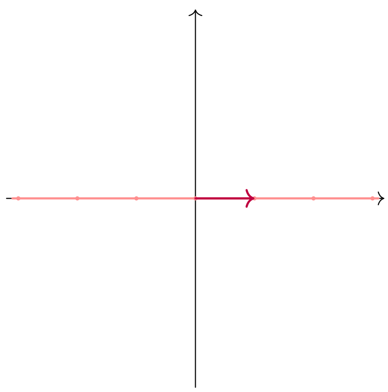
(i) 按2 缩放



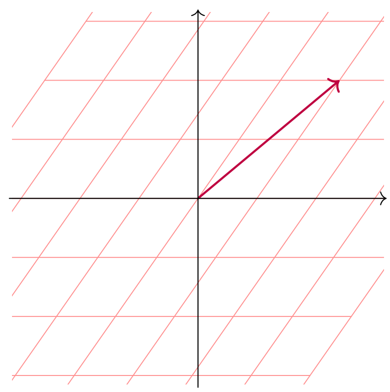
(ii) 旋转



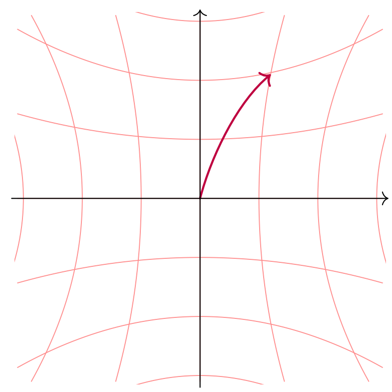
(iii) 关于x 轴的反射



(iv) 投影到x 轴上



(v) 错切



(vi) 一个非线性变换

3.2 形式定义

我们将重点关注平面中向量的变换，因此输入和输出向量都位于$\mathbb{R}^2$ 中。变换T 是一个函数$T : \mathbb{R}^2 \to \mathbb{R}^2$，如果它保持向量加法和标量乘法，那么就称它是线性的:

$$T(u+v) = T(u) + T(v), \qquad T(av) = aT(v).$$

线性变换的定义

若一个向量空间之间的函数T 对所有向量u, v 和所有标量a 都满足

$$T(u+v) = T(u) + T(v) \qquad \text{且} \qquad T(av) = aT(v),$$

则称T 是线性的。

这两条规则正是线性变换能够保持平面方格结构的代数原因。

3.3 线性变换与矩阵

一个非常有用的事实是，在$\mathbb{R}^2$ 中，一个线性变换完全由它对两个标准基向量

$$e_x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

所做的事情来决定。

事实上，每个向量都可以写成

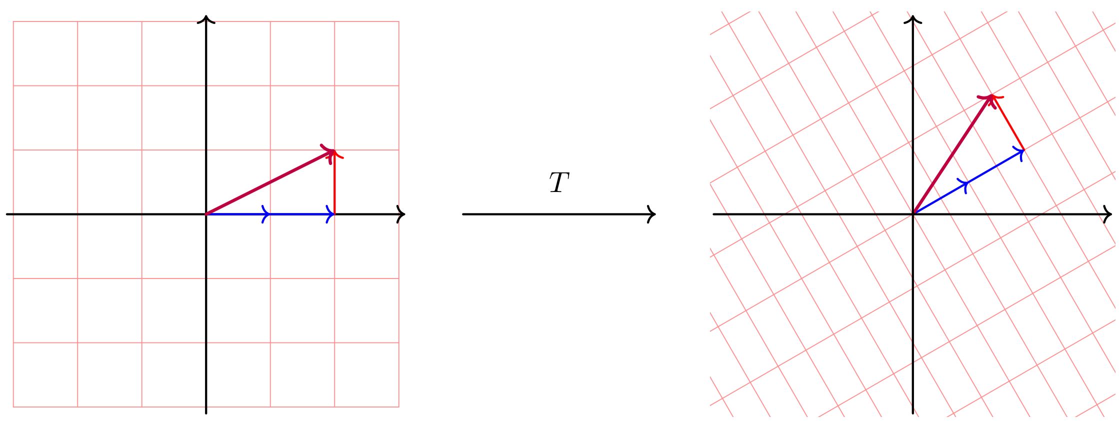
$$\begin{pmatrix} x \\ y \end{pmatrix} = xe_x + ye_y.$$

因此我们可以利用T 的线性性质得到：

$$T\begin{pmatrix} x \\ y \end{pmatrix} = T(xe_x + ye_y) = xT(e_x) + yT(e_y).$$

从几何上看，任意线性变换都会把我们的坐标网格送到另一张坐标网格上。由于输出网格仍由直线构成，我们仍然可以利用标量乘法与加法在这张新网格中移动。基向量e_x 和e_y 会被变换成两个向量$T(e_x)$ 和$T(e_y)$，它们可以用来"走遍"新网格，并描述任意其他向量的像$T(v)$ 会落在什么地方。

例如，如果我们有向量$v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$，并且想知道它在线性变换$T$ 下会落到哪里，那么我们可以查看输出网格，沿着变换后的e_x 方向走两步，再沿着变换后的e_y 方向走一步。终点就是$T(v)$：



一般地，假设一个线性变换T 把基向量e_x 和e_y 分别送到：

$$T(e_x) = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad T(e_y) = \begin{pmatrix} b \\ d \end{pmatrix},$$

其中a, b, c, d 是实数。由于任意其他向量v 都可以唯一地写成分量形式：

$$v = \begin{pmatrix} x \\ y \end{pmatrix} = xe_x + ye_y,$$

我们可以写出：

$$T\begin{pmatrix} x \\ y \end{pmatrix} = T\left(xe_x + ye_y\right) = xT(e_x) + yT(e_y) = x\begin{pmatrix} a \\ c \end{pmatrix} + y\begin{pmatrix} b \\ d \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}.$$

现在我们做出一个重要观察：如果我们把向量v 视作一个2×1 的矩阵，那么就可以进行矩阵乘法:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix},$$

这恰好与$T(v)$ 完全相同。这意味着，T 可以由矩阵

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

来描述，而它的列恰好就是基向量e_x 和e_y 在线性变换T 下的像。

平面上的每个线性变换都对应一个矩阵

如果

$$T(e_x) = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad T(e_y) = \begin{pmatrix} b \\ d \end{pmatrix},$$

那么

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

3.4 变换矩阵的例子

第3.1 节中所画出的五个线性变换例子，对应的变换矩阵如下:

(i) 按2 缩放:

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

(ii) 按角θ 旋转:

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

(iii) 关于x 轴的反射:

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

(iv) 投影到x 轴上:

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

(v) 错切:

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

作为一个明确的计算例子，如果

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix}, \qquad v = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

那么

$$Av = \begin{pmatrix} 2 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 + 1 \cdot 2 \\ -1 \cdot 1 + 3 \cdot 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

3.5 变换的复合

线性变换把一个向量空间映射到另一个向量空间。原则上，我们可以把这些变换复合起来，也就是先做一个变换，再做另一个。给定两个线性变换T 和S，我们把它们的复合记作$S \circ T$。这个记号应读作"先按T 变换这个向量空间，再按S 变换所得结果"。如果T 由矩阵A 表示，而S 由矩阵B 表示，那么复合变换$S \circ T$ 对应于从右侧开始相乘得到的矩阵乘积，也就是BA。由于矩阵乘法不满足交换律，这就意味着顺序很重要。

为了说明这一点，我们来看一个例子。设：

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

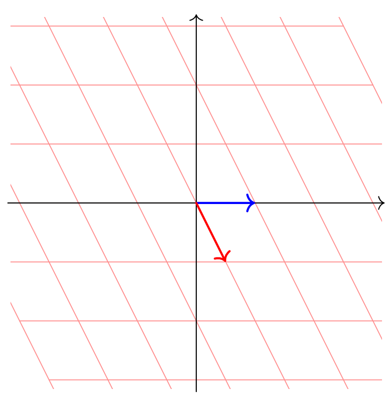
它表示关于x-轴的反射；再设

$$B = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix},$$

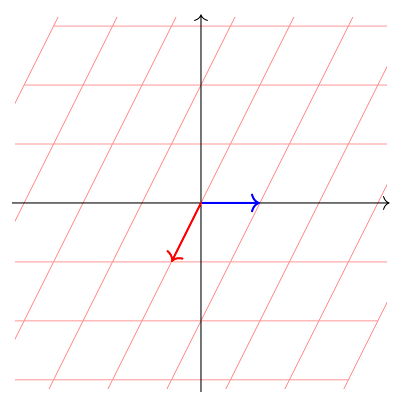
它表示一个错切。以不同顺序相乘这两个矩阵，会得到：

$$AB = \begin{pmatrix} 1 & \frac{1}{2} \\ 0 & -1 \end{pmatrix}, \qquad BA = \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & -1 \end{pmatrix},$$

因此$AB \neq BA$。由于A 是反射而B 是错切，所以乘积BA 的意思是先做反射再做错切，而乘积AB 的意思是先做错切再做反射。$AB \neq BA$ 这一事实意味着，先做反射再做错切，与先做错切然后再做反射，结果是不同的。从图形上看，这一点非常明显：



线性变换AB 的输出



线性变换BA 的输出

练习5

设

$$T = \begin{pmatrix} 3 & -1 \\ 2 & 4 \end{pmatrix}.$$

计算：

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

练习6

设一个线性变换T 满足

$$T(e_x) = \begin{pmatrix}2\\1\end{pmatrix}, \qquad T(e_y) = \begin{pmatrix}-1\\3\end{pmatrix}.$$

1. 写出T 的矩阵。
2. 求$T\left(\begin{pmatrix}4\\2\end{pmatrix}\right)$。

4 重新审视行列式

4.1 行列式作为面积

回忆上一节课，对于

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

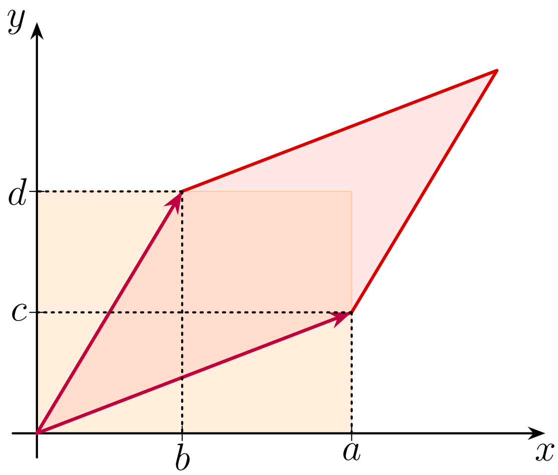
行列式为

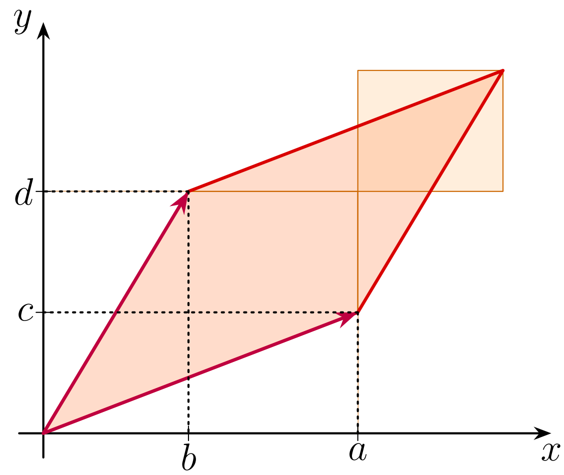
$$\det(A) = ad - bc.$$

如果我们把A 的列解释为向量

$$u = \begin{pmatrix}a\\c\end{pmatrix}, \qquad v = \begin{pmatrix}b\\d\end{pmatrix},$$

那么$\det(A)$ 与由u 和v 张成的平行四边形的面积密切相关。在图中的这种布局下，可以安排标记，使得一个面积为ad 的矩形包含这个平行四边形，而多出来的部分面积为bc。这可以通过把面积为ad 的某些部分平移穿过平行四边形来理解：





更准确地说，

$$\text{平行四边形的面积} = \big|\det(A)\big|.$$

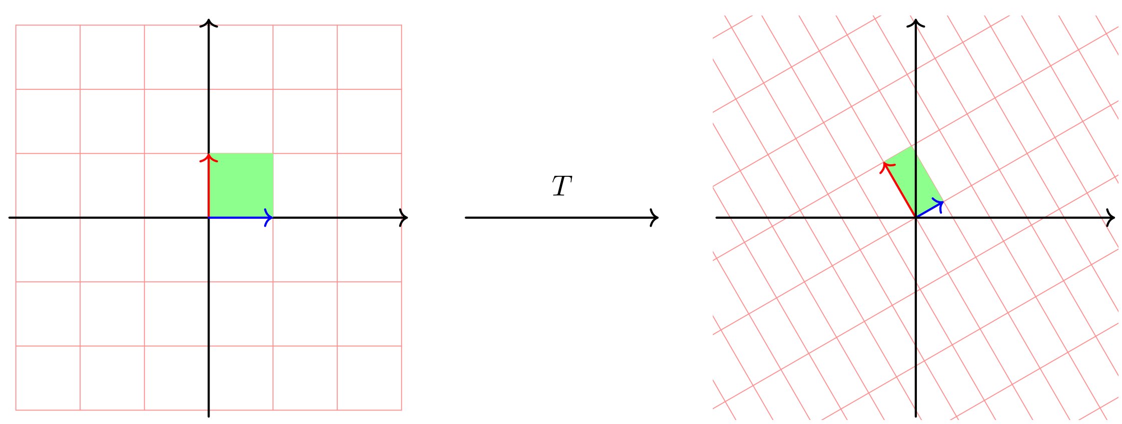
因此，2×2 行列式具有几何意义：它们衡量由两个向量构成的平行四边形的面积。

4.2 行列式作为缩放因子

还有另一种解释在线性代数中更加重要。矩阵

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

描述了一个线性变换。在这个变换下，平面中的单位正方形会被送到由A 的列张成的平行四边形。因此，$\det(A)$ 告诉我们面积在这个变换下是如何变化的。



行列式作为面积缩放因子

如果一个线性变换由矩阵A 表示，那么

$$|\det(A)|$$

告诉我们面积被缩放了多少倍。

例如:

1. 如果
$$A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$
那么
$$\det(A) = 4.$$
所以面积变为原来的4 倍。
2. 如果
$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$
那么
$$\det(A) = \cos^2\theta + \sin^2\theta = 1.$$
所以旋转保持面积不变。
3. 如果
$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$
那么
$$\det(A) = 0.$$
所以所有面积都会塌缩为零。这与把整个平面投影到一条直线上时的情形完全一致。
4. 如果
$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$
那么
$$\det(A) = -1.$$
其绝对值为1，所以面积保持不变，但负号表示这个变换翻转了定向。

4.3 3×3 行列式的几何解释

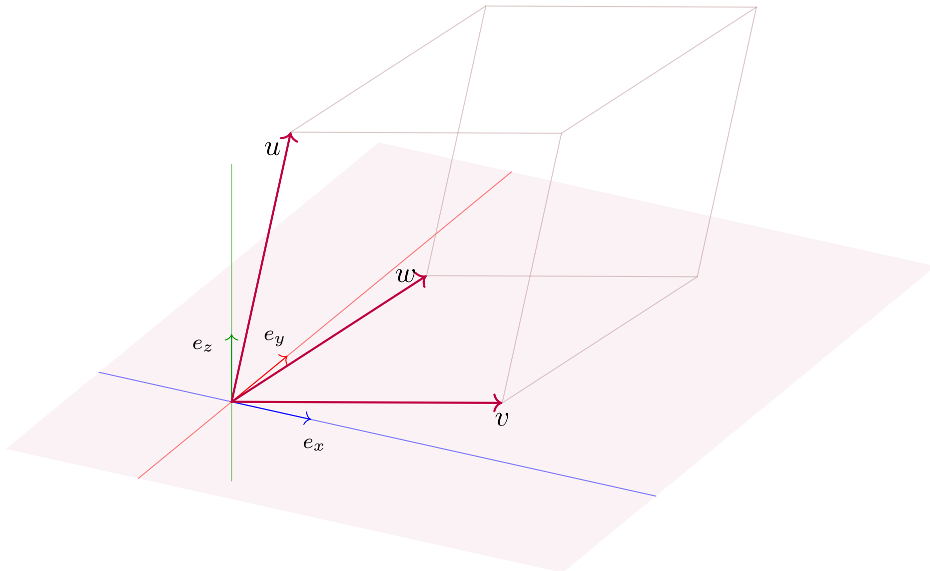
回忆一下，对于一个3×3 矩阵，我们可以利用余子式展开来计算行列式。例如，沿第三行展开时，我们得到:

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} - a_{32} \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} + a_{33} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

与二维情形完全类似，3×3 矩阵的行列式给出了由其列向量在三维空间中构成的平行六面体的有向体积。它的绝对值给出通常意义下的体积，而符号记录定向。如果我们令

$$u = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad v = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \text{以及} \quad w = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix},$$

那么这样一个平行六面体的例子可以是:



当我们沿第三行展开时，我们实际上是在把这个平行六面体分解成一系列的2×2 面积，再把它们乘上某个长度，从而形成棱柱。事实上，在沿第三行展开的情形中，这些平行四边形是这样构造出来的：从$\{u, v, w\}$ 中选取其中两个向量，用它们构成一个平行四边形，再把它投影到xy 平面上。对应的2×2 行列式就是这些投影平行四边形的有向面积。用注释的方式来看，沿第三行展开所计算的是：

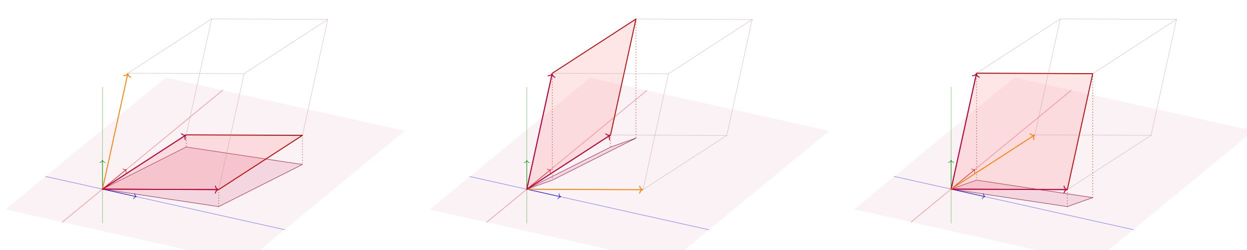
a_{31} 是第1 列向量的 z-坐标

a_{33} 是第3 列向量的 z-坐标

$$\det\begin{pmatrix}a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix} = a_{31}\overbrace{\begin{vmatrix}a_{12} & a_{13}\\ a_{22} & a_{23}\end{vmatrix}}^{\substack{\text{由第2 列和第3 列形}\\ \text{成的平行四边形}\\ \text{在投影到}xy\text{ 平面后}\\ \text{得到的有向面积}}} - a_{32}\underbrace{\begin{vmatrix}a_{11} & a_{13}\\ a_{21} & a_{23}\end{vmatrix}}_{\substack{\text{由第1 列和第3 列形成的平行四边形，}\\ \text{在投影到}xy\text{ 平面后得到的有向面积}}} + a_{33}\overbrace{\begin{vmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{vmatrix}}^{\substack{\text{由第1 列和第2 列形成的平行四边形，}\\ \text{在投影到}xy\text{ 平面后得到的有向面积}}}.$$

a_{32} 是第2 列向量的 z-坐标

这三个投影平行四边形如下图所示：



把这些相关的平行四边形棱柱重新拼接成恰好等于该平行六面体体积的过程，在几何上是相当繁琐的，也超出了本课程的范围。不过，这的确是可以做到的。

练习7

1. 求$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ 的行列式。这说明面积缩放发生了什么?
2. 求$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ 的行列式。面积会发生什么变化?
3. 沿第一行展开，计算$\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3 & 1 \\ 2 & 0 & 1 \end{vmatrix}$。

5 矩阵与解空间

线性代数与方程组有着深刻联系。像

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

这样的方程组可以改写成矩阵乘法:

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

从这个意义上说，我们可以把这个方程组理解为：寻找那个输入向量，使它在系数矩阵所描述的变换作用下，产生给定的输出向量。
这种紧凑的记号正是矩阵如此有用的一个主要原因。我们不必把若干个方程一条接一条地写出来，而是可以用一个矩阵方程来表示整个系统:

$$A\mathbf{x} = \mathbf{b}.$$

对于一个一般的方程组，更合适的说法是它的解集：即所有满足这个方程的向量$\mathbf{x}$ 所构成的集合。正如我们在前面的课程中已经看到的，只有三种可能:

1. 一个解，
2. 无穷多个解，
3. 无解。

对于方阵系统，行列式通常能帮助我们迅速识别第一种情形。类似于我们在第15 讲中讲过的克莱姆法则，这里有如下结论。

行列式与唯一性

如果A 是一个方形系数矩阵，并且$\det(A) \neq 0$，那么方程组

$$A\mathbf{x} = \mathbf{b}$$

有唯一解。
如果$\det(A) = 0$，那么这个方程组没有唯一解。它可能有无穷多个解，也可能无解。

练习8

使用行列式或简单观察，判断每个方程组是否有唯一解。

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

练习解答

练习1

$$u+v=\begin{pmatrix}2\\-1\end{pmatrix}+\begin{pmatrix}-3\\4\end{pmatrix}=\begin{pmatrix}-1\\3\end{pmatrix}.$$

$$2u=2\begin{pmatrix}2\\-1\end{pmatrix}=\begin{pmatrix}4\\-2\end{pmatrix}.$$

$$3u-v=3\begin{pmatrix}2\\-1\end{pmatrix}-\begin{pmatrix}-3\\4\end{pmatrix}=\begin{pmatrix}6\\-3\end{pmatrix}-\begin{pmatrix}-3\\4\end{pmatrix}=\begin{pmatrix}9\\-7\end{pmatrix}.$$

练习2

1. $$\begin{pmatrix}4\\-2\end{pmatrix}=4e_x-2e_y.$$
2. $$\begin{pmatrix}1\\0\\5\end{pmatrix}=1e_x+0e_y+5e_z=e_x+5e_z.$$
3. $$\operatorname{span}\left(\begin{pmatrix}2\\1\end{pmatrix}\right)$$

   是一条经过原点、方向为$\begin{pmatrix}2\\1\end{pmatrix}$的直线。

练习3

1. $$\begin{pmatrix}2&2\\3&4\end{pmatrix}+\begin{pmatrix}1&1\\3&2\end{pmatrix}=\begin{pmatrix}3&3\\6&6\end{pmatrix}.$$

2.
$$\begin{pmatrix}1 & 2\\ 1 & 5\end{pmatrix} + \begin{pmatrix}1 & 2\\ 1 & 5\end{pmatrix} = \begin{pmatrix}2 & 4\\ 2 & 10\end{pmatrix}.$$

3.
$$\begin{aligned}\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}0 & 3\\ 1 & 1\end{pmatrix} &= \begin{pmatrix}1\cdot 0+1\cdot 1 & 1\cdot 3+1\cdot 1\\ 1\cdot 0+2\cdot 1 & 1\cdot 3+2\cdot 1\end{pmatrix}\\ &= \begin{pmatrix}1 & 4\\ 2 & 5\end{pmatrix}.\end{aligned}$$

4.
$$\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}.$$

练习4

1. 有定义。一个2×3 矩阵乘以一个3×4 矩阵，结果是一个2×4 矩阵。
2. 没有定义。因为中间的维数不匹配，即$3 \neq 2$。
3. 与单位矩阵相乘会保持原矩阵不变。

练习5

1.
$$T\begin{pmatrix}1\\ 0\end{pmatrix} = \begin{pmatrix}3\\ 2\end{pmatrix}.$$

2.
$$T\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}-1\\ 4\end{pmatrix}.$$

3.
$$\begin{aligned}T\begin{pmatrix}2\\ -1\end{pmatrix} &= \begin{pmatrix}3 & -1\\ 2 & 4\end{pmatrix}\begin{pmatrix}2\\ -1\end{pmatrix}\\ &= \begin{pmatrix}3\cdot 2+(-1)(-1)\\ 2\cdot 2+4(-1)\end{pmatrix} = \begin{pmatrix}7\\ 0\end{pmatrix}.\end{aligned}$$

练习6

1. 因为这两个列向量分别是$T(e_x)$ 和$T(e_y)$，所以矩阵为
$$\begin{pmatrix}2 & -1\\ 1 & 3\end{pmatrix}.$$

2.
$$\begin{aligned}T\begin{pmatrix}4\\ 2\end{pmatrix} &= \begin{pmatrix}2 & -1\\ 1 & 3\end{pmatrix}\begin{pmatrix}4\\ 2\end{pmatrix}\\ &= \begin{pmatrix}8-2\\ 4+6\end{pmatrix} = \begin{pmatrix}6\\ 10\end{pmatrix}.\end{aligned}$$

练习7

1. $$\det\begin{pmatrix}2 & 0\\0 & 3\end{pmatrix} = 2\cdot 3 - 0\cdot 0 = 6.$$

   因此面积会按6 倍缩放。

2. $$\det\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} = 1(-1) - 0 = -1.$$

   它的绝对值是1，所以面积保持不变，不过整个平面被翻转了。

3. $$\begin{aligned}\begin{vmatrix}1 & 2 & 0\\0 & 3 & 1\\2 & 0 & 1\end{vmatrix} &= 1\begin{vmatrix}3 & 1\\0 & 1\end{vmatrix} - 2\begin{vmatrix}0 & 1\\2 & 1\end{vmatrix} + 0\begin{vmatrix}0 & 3\\2 & 0\end{vmatrix}\\ &= 1(3) - 2(0-2) + 0\\ &= 3 + 4 = 7.\end{aligned}$$

练习8

1. 系数矩阵为

   $$\begin{pmatrix}1 & 1\\1 & -1\end{pmatrix}, \qquad \det = -1 - 1 = -2 \neq 0.$$

   所以这个方程组有唯一解。

2. 系数矩阵为

   $$\begin{pmatrix}1 & 1\\2 & 2\end{pmatrix}, \qquad \det = 2 - 2 = 0.$$

   第二个方程只是第一个方程的两倍，所以这个方程组有无穷多个解。

3. 系数矩阵仍然是

   $$\begin{pmatrix}1 & 1\\2 & 2\end{pmatrix}, \qquad \det = 0.$$

   但这一次方程组不相容，因为把第一个方程乘以2 会得到$2x + 2y = 2$，而不是$2x + 2y = 3$。所以这个方程组无解。