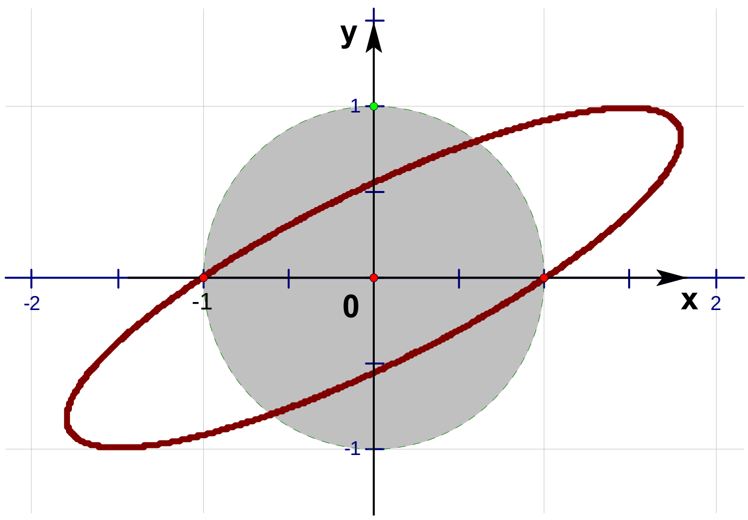
-----图解线性代数-----

线性代数的几何意义

之（5上）

任广千　编著



2010．08．16

几何意义名言录

没有任何东西比几何图形更容易印入脑际了，因此用这种方式来表达事物是非常有意义的。 -------笛卡尔

算术符号是文字化的图形，而几何图形则是图像化的公式；没有一个数学家能缺少这些图像化的公式。 --------希尔伯特

“如果代数与几何各自分开发展，那它的进步十分缓慢，而且应用范围也很有限，但若两者互相结合而共同发展，则就会相互加强，并以快速的步伐向着完善化的方向猛进。”

--------拉格朗日

不会几何学就不会正确的思考，而不会正确思考的人不过是行尸走肉。 --------柏拉图

无论是从事数学教学或研究， 我是喜欢直观的。学习一条数学定理及其证明， 只有当我能把定理的直观含义和证明的直观思路弄明白了， 我才认为真正懂了。--------中国当代数学家徐利治

第五章　矩阵的几何意义

通过前面的章节我们初步了解到，解线性方程组的克莱姆法则使用了行列式理论，但克莱姆法则只能用于解方程个数等于未知数个数的方程组，而且系数行列式不能等于 0。即使以上条件都满足，也要计算 n+1 个 n 阶行列式。实际工程中的 n 一般很大，即使在现代计算机技术面前，计算效率也不能使人满意。

在用消元法解各种类型的线性方程组时，一系列问题出现了：当系数行列式等于 0 时 ，方程组是否有解？若有解又如何求出？当未知量个数与方程的个数不等时，线性方程组的解又如何？

要深入探讨这些问题除了向量概念外还需要引入矩阵的理论。到 1858 年，哈密尔顿（W.R.Hamilton）和凯莱（A.Cay-ley）的著作中出现了矩阵的运算，从行列式到矩阵的出现，大约经过了 100 多年的时间。

我们知道，在直角坐标系中，一个有序的实数数组 (a,b) 和 (a,b,c) 分别代表了平面上和空间上的一个点，这就是实数组的几何意义。类似的，在线性空间中如果确定了一个基，线性映射就可以用确定的矩阵来表示，这就是矩阵的几何意义：线性空间上的线性映射。

矩阵独立的几何意义表现为对向量的作用结果。矩阵对一个向量是如何作用的？矩阵对多个向量是如何作用的？矩阵对一个几何图形（由无数向量组成的几何图形）是如何作用的？在矩阵对一个几何图形的作用研究中，我们会发现一些有规律的东西比如特征向量、秩等等。

5.1. 矩阵的概念

矩阵的本质就是一个长方形的数表，在生活中的所有长方形数表都可以看成是矩阵。矩阵和行列式相似，也是以行和列组织的矩形数字阵列，因此称为矩阵。与行列式的表示法不同，矩阵是用方括号把数字块括起来，表示一个有顺序有组织的数据块；而行列式是对这些数据块进行的一个运算，是一个算式，故称为行列式。矩阵的一个三阶例子如下：

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

如果用数组来统一定义标量、向量和矩阵的话就是：标量是一维向量，向量是标量的数组，矩阵则是向量的数组。例如上面介绍的矩阵我们如果使用列向量 $\mathbf{a}=(a_1,a_2,a_3)^T$ ，$\mathbf{b}=(b_1,b_2,b_3)^T$ ，$\mathbf{c}=(c_1,c_2,c_3)^T$ 来表示它，这个矩阵就可以写作：$\begin{bmatrix}\mathbf{a} & \mathbf{b} & \mathbf{c}\end{bmatrix}$。

当然，矩阵不只是只有几何意义，也具有现实的物理意义，矩阵的运算也都可以从实践中找到。下面有个例子：

比如某家电器公司的制造厂有几个生产线，产线在 2009 年和 2010 年的上半年的产出量的统计表如下：

顺序	产线名	2009 年上半年的每月产出量

		1 月	2 月	3 月	4 月	5 月	6 月
1	冰箱线	22	35	30	23	25	12
2	吸尘器线	25	43	32	34	35	30
3	电视线	23	23	34	44	40	45

顺序	产线名	2010 年上半年的每月产出量					
		1 月	2 月	3 月	4 月	5 月	6 月
1	冰箱线	22	34	30	23	25	12
2	吸尘器线	24	43	32	34	35	34
3	电视线	23	23	34	45	41	45
4	手机线	34	34	35	45	23	43
5	VCD 线	45	24	31	34	45	12

我们将第一个表格对应的矩阵记为 $\mathbf{A}$，第二个表格对应的矩阵记为 $\mathbf{B}$，则有：

$$\mathbf{A}=\begin{bmatrix} 22 & 35 & 30 & 23 & 25 & 12 \\ 25 & 43 & 32 & 34 & 35 & 30 \\ 23 & 23 & 34 & 44 & 40 & 45 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{和} \mathbf{B}=\begin{bmatrix} 22 & 34 & 30 & 23 & 25 & 12 \\ 24 & 43 & 32 & 34 & 35 & 34 \\ 23 & 23 & 34 & 45 & 41 & 45 \\ 34 & 34 & 35 & 45 & 23 & 43 \\ 45 & 24 & 31 & 34 & 45 & 12 \end{bmatrix}。$$

那么有：

$\mathbf{A}+\mathbf{B}$ 实际意义是：2009、2010 年 1~6 月各产线每月产量的和(2001 年手机，VCD 机的产量为 0)；

$\mathbf{B}-\mathbf{A}$ 实际意义是： 2010 年 1~6 月各产线每月产量比上年同期的增产情况；

设冰箱、吸尘器、电视机、手机、VCD 机的价格分别是 l500 元台、900 元/台、300 元/台、2500 元/台、980 元/台，则有向量 $\mathbf{c}=(1500\ \ 900\ \ 3000\ \ 2500\ \ 980)$，那么矩阵乘法：

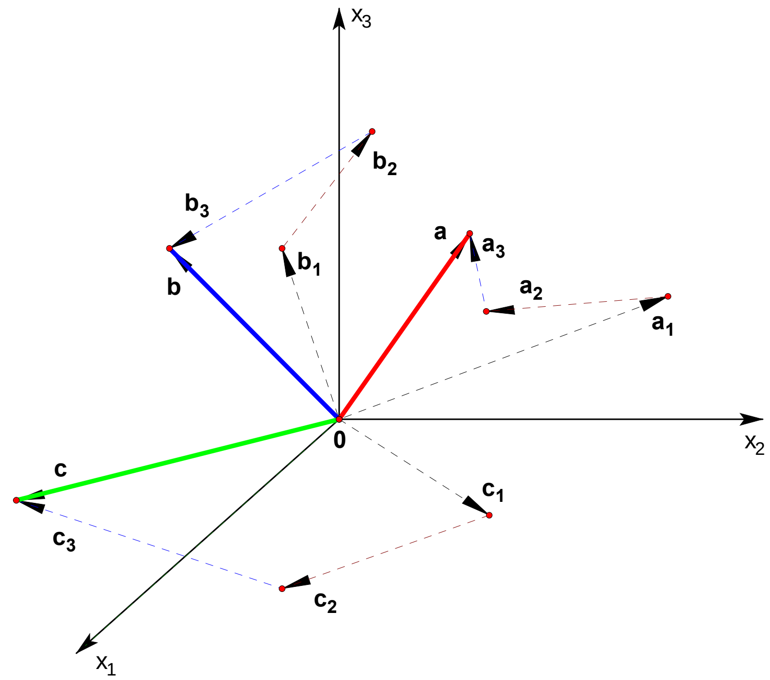
$\mathbf{cB}$ 的实际意义是：2010 年 1~6 月该制造厂每种产品的月产值。

如果每生产一件产品职工得到的奖励积分为 3 分，则数乘运算 $3\mathbf{A}$ 实际意义是 2009 年 1~6 月该厂各车间的职工月积分。

5.2. 矩阵加法的几何意义

矩阵的加法和乘法等简单运算可看作来自于线性方程组的简单运算，读者可以参看第七章的 7.1 节的详细介绍。在下面介绍矩阵的加法和乘法几何意义时，我们仍然不能离开向量的有力帮助。矩阵中的行向量或列向量的意义可以有效地帮助我们看清矩阵所蕴含的几何变换的意义。

多个矩阵的加法比较简单，即使不用给出几何意义，我们也能轻松掌握它。不过画出矩阵加法的几何图形，可以帮助你对多个向量所组成的几何图形的叠加有个形象的认知。



上图显示了三组向量同时连加，每组有三个向量的分量连加。把上述用矩阵表述出来就是：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ c_{11} & c_{12} & c_{13} \end{bmatrix} + \begin{bmatrix} a_{21} & a_{22} & a_{23} \\ b_{21} & b_{22} & b_{23} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} + \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ b_{31} & b_{32} & b_{33} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

我们把上述每个矩阵都分解为三个行向量来给出图形的，其实因为矩阵的加法是对每个元素分别对应相加，因此对于列向量同样等效。

5.3. 矩阵与向量的乘法的几何意义

矩阵与向量乘积比如 $\mathbf{Ax}$ 表现为矩阵 $\mathbf{A}$ 对一个向量 $\mathbf{x}$ 作用的结果。其作用的主要过程是对一个向量进行旋转和缩放的综合过程（即线性变换的过程），一个向量就变换为另外一个向量。一个 m 行 n 列的实矩阵 $\mathbf{A}_{m\times n}$ 就是一个 $R^n \to R^m$ 上的线性变换，或者地说，矩阵 $\mathbf{A}_{m\times n}$ 把一个 n 维空间的 n 维向量变换为一个 m 维空间的 m 维向量。

矩阵与向量的乘积的概念

矩阵 $\mathbf{A}$ 与向量 $\mathbf{c}$ 的乘积（矩阵左乘向量，记为 $\mathbf{Ac}$）是一个向量，这个向量的每个分量是以矩阵 A 的每个行向量分别与列向量 $\mathbf{c}$ 的数量积作为元素的。乘式如下：

$$\mathbf{A}\cdot\mathbf{c}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}=\begin{pmatrix} (a_1\ \ a_2\ \ a_3)\cdot\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \\ (b_1\ \ b_2\ \ b_3)\cdot\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \end{pmatrix}=\begin{pmatrix} a_1c_1+a_2c_2+a_3c_3 \\ b_1c_1+b_2c_2+b_3c_3 \end{pmatrix}$$

上式 $\mathbf{Ac}$ 的乘积是把矩阵 $\mathbf{A}$ 看作两个行向量，实质上还是向量与向量的点乘积。

类似的，向量 $\mathbf{c}$ 与矩阵 $\mathbf{A}$ 的乘积（矩阵右乘向量，记为 $\mathbf{cA}$）也是一个向量，这个 $\mathbf{cA}$ 向量的每个分量是行向量 $\mathbf{c}$ 与矩阵 $\mathbf{A}$ 的每个列向量的点乘积。乘法公式如下：

$$\mathbf{c}\cdot\mathbf{A}=(c_1\ \ c_2\ \ c_3)\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}=\left((c_1\ \ c_2\ \ c_3)\cdot\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}\quad (c_1\ \ c_2\ \ c_3)\cdot\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\right)=(c_1a_1+c_2a_2+c_3a_3\quad c_1b_1+c_2b_2+c_3b_3)$$

以上的乘积运算都是“行向量·列向量”的形式。下面我们换一下思维方式，把乘积运算看成“列向量·行向量”的形式是否说得通？先从 $\mathbf{Ac}$ 的乘积开始：

如果把矩阵 A 分解为 3 个列向量的话，我们可以这样展开上式：

$$\mathbf{A}\cdot\mathbf{c}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}=\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}c_1+\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}c_2+\begin{pmatrix} a_3 \\ b_3 \end{pmatrix}c_3=\begin{pmatrix} a_1c_1+a_2c_2+a_3c_3 \\ b_1c_1+b_2c_2+b_3c_3 \end{pmatrix}。$$

上式的含义是 $\mathbf{Ac}$ 的乘积可以理解为矩阵 $\mathbf{A}$ 的列向量的线性组合，组合系数是向量 $\mathbf{c}$ 的三个分量。这个展开的实质仍然是向量点乘积的乘法。再看 $\mathbf{c}\cdot\mathbf{A}$ 的展开：

$$\mathbf{c}\cdot\mathbf{A}=(c_1\ c_2\ c_3)\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix}=c_1(a_1\ \ b_1)+c_2(a_2\ \ b_2)+c_3(a_3\ \ b_3)=(c_1a_1+c_2a_2+c_3a_3\quad c_1b_1+c_2b_2+c_3b_3)$$

这个式子可以理解为矩阵 A 的行向量的线性组合，组合系数是向量 c 的三个分量。

实际上，在以上的各种乘法中，我们使用了矩阵和向量的分块技术（全部是“行向量·列向量”的形式），每一个分块都要看成是矩阵最基本的元素“数”进行运算。至此，我们较全面地理解了向量与矩阵乘积的展开实质。显然，左乘与右乘的结果是不同的。为什么不同，答案就在随后的章节里。

矩阵与向量乘积的几何意义

为了更具体的观察矩阵和向量乘积的几何意义，我们下面先考察一个矩阵与欧式空间的单位坐标向量的乘积的过程，然后再看一个矩阵与任意向量的乘积的几何意义。

矩阵与单位坐标向量的乘积的几何解释

三维的单位坐标向量就是 $\mathbf{i}=(1,0,0)$，$\mathbf{j}=(0,1,0)$，$\mathbf{k}=(0,0,1)$。我们取 x 坐标轴上的单位向量 $\mathbf{i}$ 与 $\mathbf{A}=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$ 相乘，得乘式如下：

$$\mathbf{iA}=(1,0,0)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(a_1\ \ a_2\ \ a_3),$$

$$\mathbf{Ai}^T=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$$

$\mathbf{iA}$ 的结果是 $(a_1\ \ a_2\ \ a_3)$，这恰是矩阵 $\mathbf{A}$ 的第一行。$\mathbf{Ai}^T$ 的结果是 $\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$，这恰是矩阵 $\mathbf{A}$ 的第一列。

为了更明了，下面把 $\mathbf{j}$ 和 $\mathbf{k}$ 与矩阵 $\mathbf{A}$ 的乘积也一并列出来：

$$\mathbf{jA}=(0,1,0)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(b_1\ \ b_2\ \ b_3)$$

$$\mathbf{kA}=(0,0,1)\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}=(c_1\ \ c_2\ \ c_3)$$

$$\mathbf{Aj}^T=\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}=\begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}$$

$$\mathbf{A}\mathbf{k}^T = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$$

从向量对矩阵的作用方面上，我们可以这样理解上述的乘式给出的操作意义上的内涵：

1、 x 单位坐标向量 $\mathbf{i}$ 左乘一个矩阵就是把矩阵 x 轴行向量（第一行）给取出来；类似的，y 轴单位坐标向量 $\mathbf{j}$ 左乘一个矩阵就是把矩阵 y 轴行向量（第二行）给取出来；z 坐标单位向量 $\mathbf{k}$ 左乘一个矩阵就是把矩阵 z 轴行向量（第三行）给取出来;

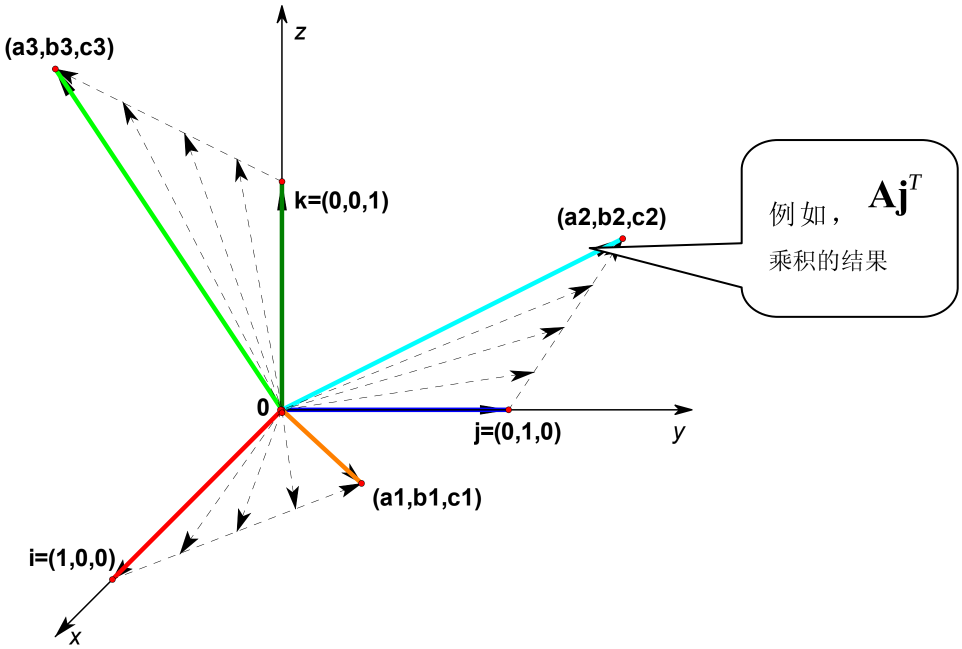
2、 x 坐标单位向量 $\mathbf{i}$ 右乘一个矩阵就是把矩阵 x 轴列向量（第一列）给取出来；类似的，y 坐标单位向量 $\mathbf{j}$ 右乘一个矩阵就是把矩阵 y 轴列向量（第二列）给取出来；z 坐标单位向量 $\mathbf{k}$ 右乘一个矩阵就是把矩阵 z 轴列向量（第三列）给取出来;

另外，从矩阵对向量的作用上，我们又可以从几何图形上这样理解其作意义上的内涵:

3、 一个矩阵右乘 x 坐标单位向量 $\mathbf{i}$，就是把向量 $\mathbf{i}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 x 轴上的行向量（第一行）；类似的，一个矩阵右乘 y 坐标单位向量 $\mathbf{j}$，就是把向量 $\mathbf{j}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 y 轴上的行向量（第二行）；一个矩阵右乘 z 坐标单位向量 $\mathbf{k}$，就是把向量 $\mathbf{k}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 z 上的行向量（第三行）;

4、 一个矩阵左乘 x 坐标单位向量 $\mathbf{i}$，就是把向量 $\mathbf{i}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 x 轴上的列向量（第一列）；类似的，一个矩阵左乘 y 坐标单位向量 $\mathbf{j}$，就是把向量 $\mathbf{j}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 y 轴上的列向量（第二列）；一个矩阵左乘 z 坐标单位向量 $\mathbf{k}$，就是把向量 $\mathbf{k}$ 的图形缩放旋转变换，变换后的向量就是这个矩阵的 z 轴上的列向量（第三列）;

任何一个矩阵和单位向量 $\mathbf{i}$ 相乘得到矩阵的第一列向量。相类似的，任何一个向量和单位向量 $\mathbf{i}$ 相乘（内积），是这个向量的第一个分量。



上图中，我们给出了三阶矩阵 $\mathbf{A}$ 分别右乘 x, y, z 坐标单位向量 $\mathbf{i}^T$ ， $\mathbf{j}^T$ ， $\mathbf{k}^T$ 后的变化几何图形，单位向量 $\mathbf{i}^T$ ， $\mathbf{j}^T$ ， $\mathbf{k}^T$ 分别**缩放旋转**变化成了向量 $\begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}$ 、向量 $\begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix}$ 和向量 $\begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$ 。虚线向量表示一种变化的过程。

矩阵与任意向量的乘积的几何解释

在前面的章节中我们讲过，一个向量可以拆分为单位坐标向量的线性表示，或者讲是单位坐标向量的伸缩变换后的和。我们再次列出这个表达式如下：

$$\mathbf{d} = (d_1, d_2, d_3) = d_1\mathbf{i} + d_2\mathbf{j} + d_3\mathbf{k} = d_1(1,\ 0,\ 0) + d_2(0,\ 1,\ 0) + d_3(0,\ 0,\ 1)$$

或

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = d_1\mathbf{i}^T + d_2\mathbf{j}^T + d_3\mathbf{k}^T = d_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + d_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + d_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

所以，矩阵 $\mathbf{A}$ 对任意向量 $\mathbf{d}$ 的乘积式就是

$$\mathbf{A}\cdot\mathbf{d}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}(d_1\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}+d_2\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}+d_3\begin{pmatrix}0\\ 0\\ 1\end{pmatrix})$$

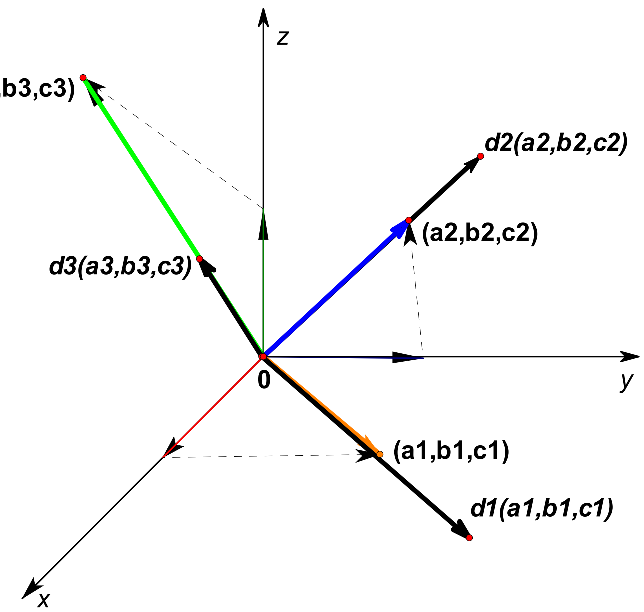
$$=d_1\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}+d_2\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}+d_3\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\cdot\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$$

$$=d_1\mathbf{A}\cdot\mathbf{i}^T+d_2\mathbf{A}\cdot\mathbf{j}^T+d_3\mathbf{A}\cdot\mathbf{k}^T$$

上式揭示了矩阵 $\mathbf{A}$ 对任意向量 $\mathbf{d}$ 的乘积的几何解释可以分解为几个操作过程：首先，矩阵 $\mathbf{A}$ 分别对单位向量 $\mathbf{i}^T$， $\mathbf{j}^T$， $\mathbf{k}^T$ 进行伸缩旋转变换后得到三个向量（列向量，上节的内容），然后对这三个列向量 $\mathbf{Ai}^T$， $\mathbf{Aj}^T$， $\mathbf{Ak}^T$ 分别进行伸缩变换得到了 $d_1\mathbf{Ai}^T$， $d_2\mathbf{Aj}^T$， $d_3\mathbf{Ak}^T$，再把此三向量相加，得到的和向量就是 $\mathbf{A}\cdot\mathbf{d}$ 的乘积。

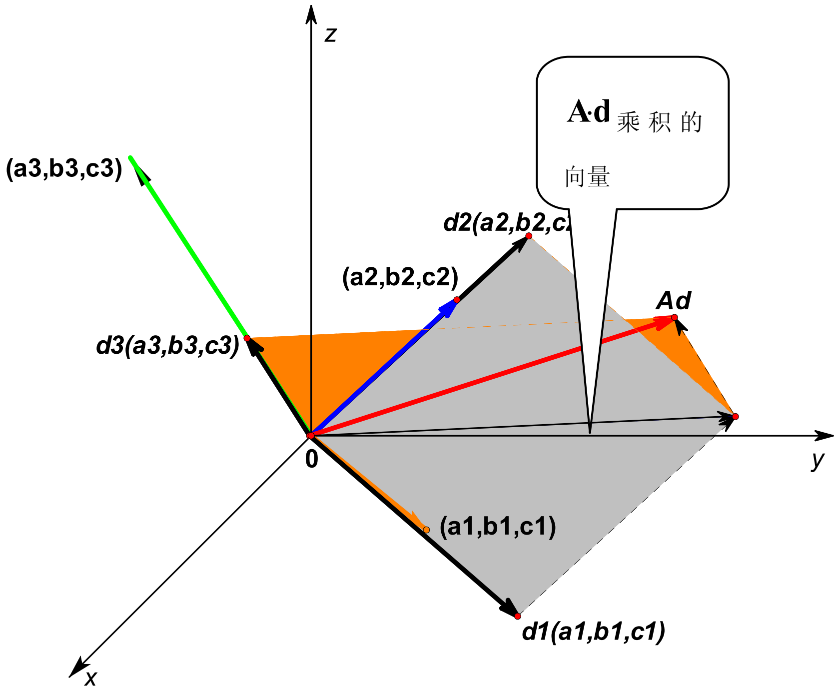
绘出上式的几何图像如下。

下图得到了中间结果三个向量 $d_1\mathbf{Ai}^T=d_1(a_1,b_1,c_1)$， $d_2\mathbf{Aj}^T=d_2(a_2,b_2,c_2)$，$d_3\mathbf{Ak}^T=d_3(a_3,b_3,c_3)$ 的几何图形。



下图把中间结果三个向量 $d_1\mathbf{Ai}^T=d_1(a_1,b_1,c_1)$， $d_2\mathbf{Aj}^T=d_2(a_2,b_2,c_2)$ 和 $d_3\mathbf{Ak}^T=d_3(a_3,b_3,c_3)$，依照平行四边形法则或多边形法则连加，得到最终结果

$\mathbf{A}\cdot\mathbf{d}=(a_1d_1+a_2d_2+a_3d_3,b_1d_1+b_2d_2+b_3d_3,c_1d_1+c_2d_2+c_3d_3)$。



呵，不太直观是吧。好，把上式简化一下，利用线性表示的概念（在矩阵与向量乘积概念一节中讲过）我们可能得到一个更直接的易于理解的几何图像，调整乘式如下：

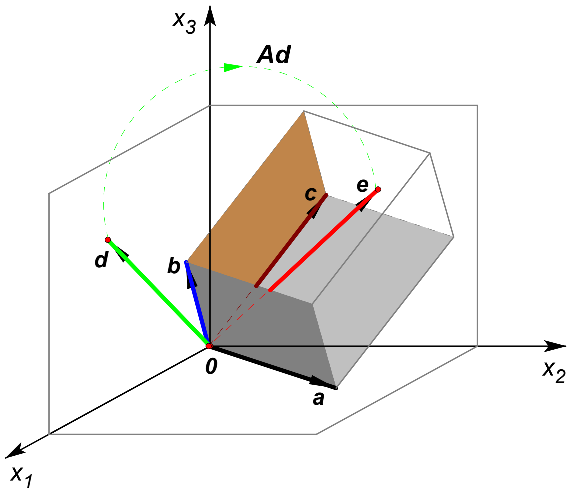
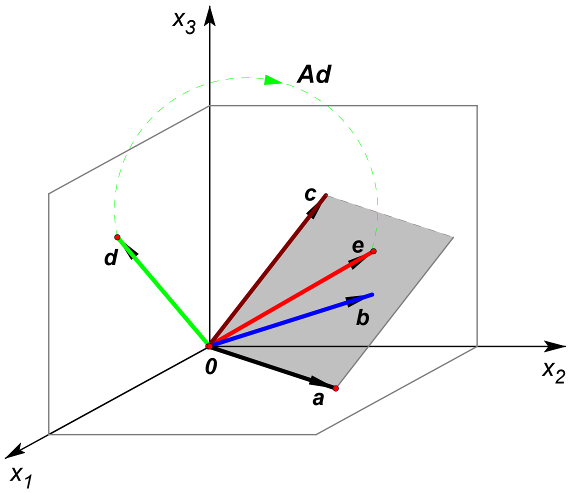
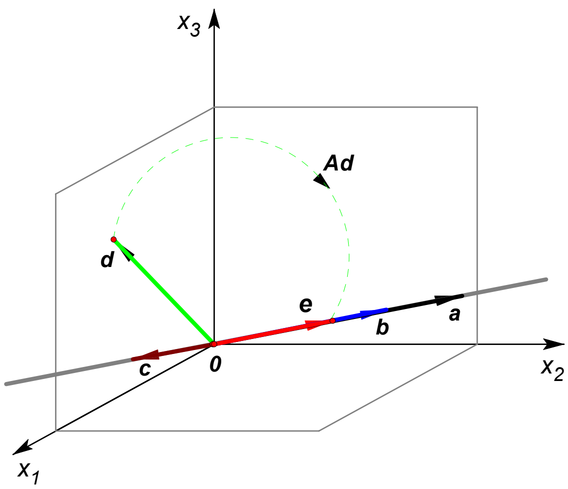
$$\mathbf{A}\cdot\mathbf{d}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}\begin{pmatrix}d_1\\ d_2\\ d_3\end{pmatrix}=\begin{bmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{bmatrix}(d_1\begin{pmatrix}1\\0\\0\end{pmatrix}+d_2\begin{pmatrix}0\\1\\0\end{pmatrix}+d_3\begin{pmatrix}0\\0\\1\end{pmatrix})$$

$$=d_1\begin{pmatrix}a_1\\ b_1\\ c_1\end{pmatrix}+d_2\begin{pmatrix}a_2\\ b_2\\ c_2\end{pmatrix}+d_3\begin{pmatrix}a_3\\ b_3\\ c_3\end{pmatrix}$$

上式的几何意义就是向量组的线性表示的几何意义（后面的章节中还要讨论），几何解释就是把矩阵 $\mathbf{A}$ 的列向量进行伸缩变换（比例变换，也可能改变方向）后首尾相连（即向量的和）得到了一个新向量，这个向量就是 $\mathbf{A}\cdot\mathbf{d}$ 。这个几何解释校前面的解释更简洁一点，但没有大的差别。如果我们知道矩阵的列空间的概念，我们可以得到一个新的几何意义上的理解：

矩阵 $\mathbf{A}$ 的列向量空间是 $\mathbf{A}$ 的所有列向量所张成 R^n 中的一个子空间。尽管这些列向量可能是线性相关的，那么 $\mathbf{A}\cdot\mathbf{d}$ 的表达式（见上式）则说明，矩阵 $\mathbf{A}$ 把 R^n 中的向量 $\mathbf{d}$ 映射到列向量子空间里的一个向量上去了。一个 m 行 n 列的实矩阵 $\mathbf{A}_{m\times n}$ 就是一个 $R^n\to R^m$ 上的线性变换，或者地说，矩阵 $\mathbf{A}_{m\times n}$ 把一个 n 维空间的 n 维向量变换为一个 m 维空间的 m 维向量。

看一个综合性的表格如下：

顺序	图形解释	对应的几何图形
1	$$\mathbf{A}_1\cdot\mathbf{d}_1=\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}=d_1\begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}+d_2\begin{pmatrix}b_1\\b_2\\b_3\end{pmatrix}+d_3\begin{pmatrix}c_1\\c_2\\c_3\end{pmatrix}=\begin{pmatrix}e_1\\e_2\\e_3\end{pmatrix}$$ 3×3 矩阵 $\mathbf{A}$ 把一个三维向量 $\mathbf{d}$ 映射到一个三维向量 $\mathbf{e}$ 。	
2	$$\mathbf{A}_2\cdot\mathbf{d}_2=\begin{bmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}=d_1\begin{pmatrix}a_1\\a_2\end{pmatrix}+d_2\begin{pmatrix}b_1\\b_2\end{pmatrix}+d_3\begin{pmatrix}c_1\\c_2\end{pmatrix}=\begin{pmatrix}e_1\\e_2\end{pmatrix}$$ 2×3 矩阵 $\mathbf{A}$ 把一个三维向量 $\mathbf{d}$ 映射到一个二维向量 $\mathbf{e}$ 。	
3	$$\mathbf{A}_3\cdot\mathbf{d}_3=\begin{bmatrix}a_1&b_1&c_1\end{bmatrix}\begin{pmatrix}d_1\\d_2\\d_3\end{pmatrix}=d_1a_1+d_2b_1+d_3c_1=e$$ 1×3 矩阵 $\mathbf{A}$ 把一个三维向量 $\mathbf{d}$ 映射到一个一维向量 $\mathbf{e}$ （或实数 e ）。	

下节我们讨论一个特殊的矩阵---旋转矩阵。因为通过进一步了解旋转矩阵会深化大家对矩阵几何意义的理解。

旋转矩阵对向量的乘积的几何解释

大家已经知道，一个矩阵乘以一个向量，一般将会对向量的几何图形进行旋转和伸缩变化。在教科书中，我们常见的一个例子就是旋转矩阵，旋转矩阵只对向量进行旋转变化而没有伸缩变化。例如二阶旋转矩阵 **A**:

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

显然，二阶的所有的某一特定角度的旋转矩阵都分布在单位园上。对θ取几个不同的弧度，就会得到几个旋转矩阵：

θ	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$
$\mathbf{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$	$\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$	$\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$	$\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}$	$\begin{pmatrix}-\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$

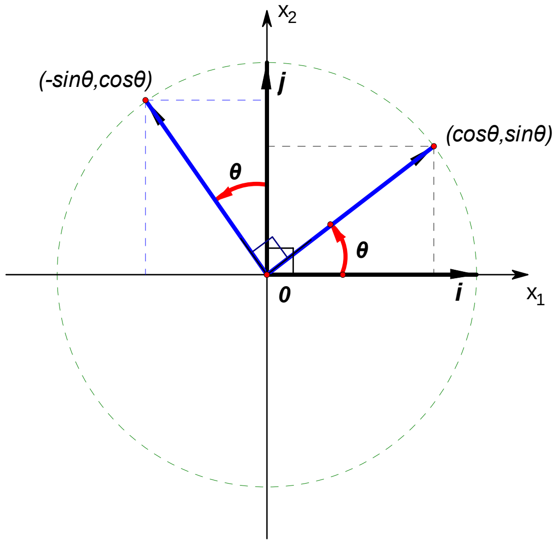
比如单位矩阵 $\mathbf{A}_1=\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$ 实际上就是对一个向量 **c** 旋转的角度是 0，也就是 $\mathbf{Ac}=\mathbf{c}$；而矩阵 $\mathbf{A}_2=\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$ 就是对一个向量 **c** 旋转的角度是 $\frac{\pi}{4}$。

是不是这样的呢？首先看一下旋转矩阵 **A** 对单位向量 $\mathbf{i},\mathbf{j}$ 的作用效果。

$$\mathbf{Ai}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1\\ 0\end{pmatrix}=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},$$

$$\mathbf{Aj}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}0\\ 1\end{pmatrix}=\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}=\begin{pmatrix}\cos(\theta+\pi/2)\\ \sin(\theta+\pi/2)\end{pmatrix}$$

从上式和下图看出，旋转矩阵对单位向量 $\mathbf{i},\mathbf{j}$ 确实分别逆时针旋转了一个θ角度。旋转后的两个向量 **Ai** 和 **Aj** 保持长度不变和夹角不变。或者说向量 $\mathbf{i}\rightarrow\mathbf{Ai}$ 长度不变，角度逆时针增加了θ度；向量 $\mathbf{j}\rightarrow\mathbf{Aj}$ 长度不变，同时同向角度增加了θ度。显然，向量的和 $\mathbf{i}+\mathbf{j}\rightarrow\mathbf{Ai}+\mathbf{Aj}$ 长度不变，角度逆时针增加了θ度。



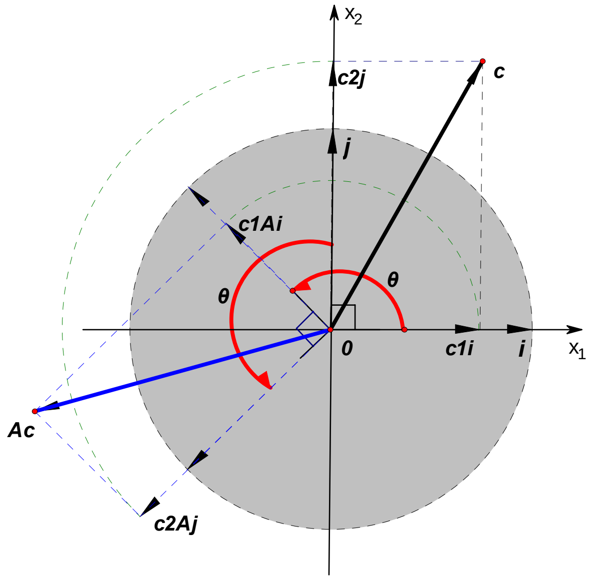
对于任意向量 $\mathbf{c}$，我们知道，可以分解为单位向量的线性表示：

$$\mathbf{c}=\begin{pmatrix}c_1\\c_2\end{pmatrix}=c_1\begin{pmatrix}1\\0\end{pmatrix}+c_2\begin{pmatrix}0\\1\end{pmatrix}=c_1\mathbf{i}+c_2\mathbf{j}$$

那么，旋转矩阵作用于向量 $\mathbf{c}$ 的式子为：

$$\mathbf{A}\begin{pmatrix}c_1\\c_2\end{pmatrix}=\mathbf{A}(c_1\begin{pmatrix}1\\0\end{pmatrix}+c_2\begin{pmatrix}0\\1\end{pmatrix})\ =c_1\mathbf{A}\begin{pmatrix}1\\0\end{pmatrix}+c_2\mathbf{A}\begin{pmatrix}0\\1\end{pmatrix}=c_1\mathbf{Ai}+c_2\mathbf{Aj}$$

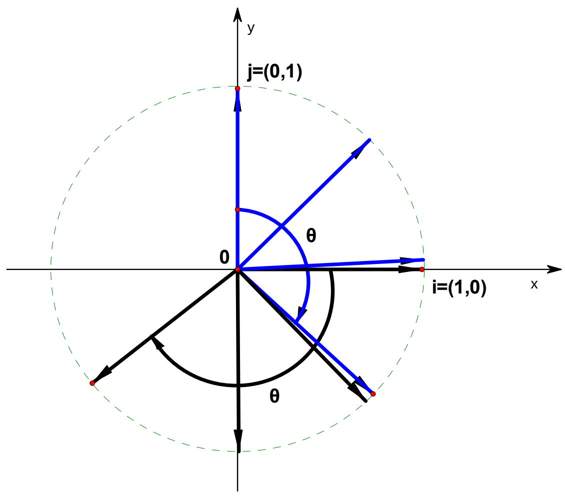
类似的，我们考察旋转矩阵 $\mathbf{A}$ 把任意向量 $\mathbf{c}$ 变换到 $\mathbf{Ac}$ 时，对比上述两式，向量 $c_1\mathbf{i}\to c_1\mathbf{Ai}$ 长度不变，角度逆时针增加了 θ 度；向量 $c_2\mathbf{j}\to c_2\mathbf{Aj}$ 长度不变，同时同向角度增加了 θ 度。那么，向量的和 $c_1\mathbf{i}+c_2\mathbf{j}\to c_1\mathbf{Ai}+c_2\mathbf{Aj}$ 长度不变，角度逆时针增加了 θ 度，即 $\mathbf{c}\to\mathbf{Ac}$ 长度不变，角度逆时针增加了 θ 度。如下图所示。



另外，对于不同的旋转矩阵之间的比较，也有一个小小的规律。

θ	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$
$A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$	$\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$	$\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$	$\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}$	$\begin{pmatrix}-\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$

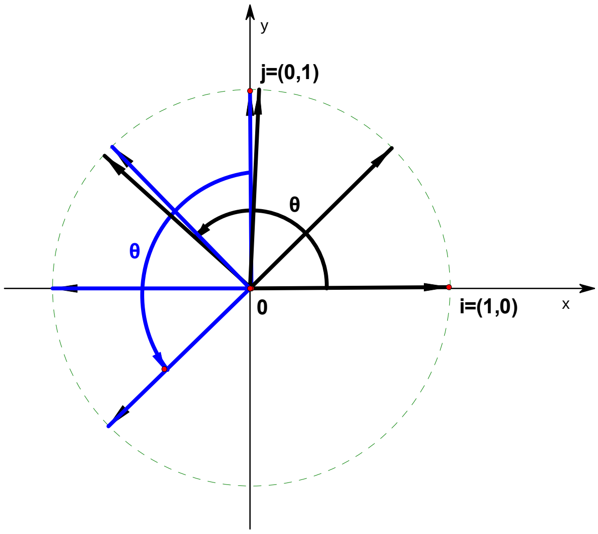
我们给出矩阵几组具体的数值来观察其**行向量**的图形的变化。



由上表格和图形知道，当 θ =0 时，旋转矩阵的行向量就是 x 和 y 轴的单位向量 **i** 和 **j**，随着 θ 角度的增大，单位向量在单位园上同时进行顺时针等角度旋转。

如果我们从列向量的角度看旋转矩阵，又会得到什么图形呢？把上述矩阵的几组具体的数值重新列在下面，来观察其**列向量**的图形的变化。

θ	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$
$\mathbf{A}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$	$\begin{pmatrix}1, & 0\\ 0, & 1\end{pmatrix}$	$\begin{pmatrix}\frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & \frac{\sqrt{2}}{2}\end{pmatrix}$	$\begin{pmatrix}0, & -1\\ 1, & 0\end{pmatrix}$	$\begin{pmatrix}-\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2}\\ \frac{\sqrt{2}}{2}, & -\frac{\sqrt{2}}{2}\end{pmatrix}$



显然，从上图看出列向量在随着 θ 的角度的增加而逆时针旋转，刚好与行向量同时反向旋转！这正是正交矩阵的特征。也就是说，旋转矩阵就是正交矩阵。

5.4. 矩阵与矩阵的乘法几何意义

两个矩阵相乘如 **AB** 的几何意义可以从多个角度来了解。如果把矩阵 **A** 看作一个几何图形，那么乘以 **B** 就是把 **A** 的图形进行了有规律的变换，这个规律的变换就是线性变换（这里实际上把矩阵 **A** 看成了多个向量的组合）；如果把两个矩阵看作等同的，那么 **AB** 的结果是把两个线性变换进行了叠加或复合（嗨，至于矩阵乘法有什么代数上的意义，你可以先看看 6.9 节“方程组和矩阵、向量组的关系”会得到合理的解释）。下面我们具体的讨论一下。

矩阵与矩阵的乘法的意义

矩阵与矩阵的乘法可以从矩阵与向量的乘法得到，因为一个矩阵与多个向量相乘，这多个向量就可以组成一个矩阵（会有些限制）。或者说，矩阵本身就是一个有排列顺序要求的向量组，所以矩阵与矩阵相乘可以看作矩阵乘以列向量（或者行向量乘以矩阵的）的组合。例如：

$$\mathbf{Ac}=\begin{bmatrix}a_1 & a_2\\ b_1 & b_2\end{bmatrix}\begin{pmatrix}c_1\\ c_2\end{pmatrix}\quad,\quad \mathbf{Ad}=\begin{bmatrix}a_1 & a_2\\ b_1 & b_2\end{bmatrix}\begin{pmatrix}d_1\\ d_2\end{pmatrix}$$

如果把列向量 $\mathbf{c}$ 和 $\mathbf{d}$ 可以组合成一个矩阵 $\mathbf{B}=(\mathbf{c},\mathbf{d})$，那么上述的乘积可以用两个矩阵的乘积来表示：

$$\mathbf{AB}=\mathbf{A}(\mathbf{c},\mathbf{d})=\begin{bmatrix}a_1 & a_2\\ b_1 & b_2\end{bmatrix}\begin{bmatrix}c_1 & d_1\\ c_2 & d_2\end{bmatrix}$$

当然，如果有更多的向量组合起来，可以形成这样的矩阵乘式：

$$\mathbf{AC}=\mathbf{A}(\mathbf{c},\mathbf{d},\mathbf{e},\mathbf{f},\mathbf{g})=\begin{bmatrix}a_1 & a_2\\ b_1 & b_2\end{bmatrix}\begin{bmatrix}c_1 & d_1 & e_1 & f_1 & g_1\\ c_2 & d_2 & e_2 & f_2 & g_2\end{bmatrix}$$

类似的，通过向量乘以矩阵的定义，我们同样可以定义反顺序的矩阵乘式，这里向量为行向量：

$$\mathbf{CA}=\begin{pmatrix}\mathbf{c}\\ \mathbf{d}\\ \mathbf{e}\\ \mathbf{f}\\ \mathbf{g}\end{pmatrix}\mathbf{A}=\begin{bmatrix}c_1 & c_2\\ d_1 & d_2\\ e_1 & e_2\\ f_1 & f_2\\ g_1 & g_2\end{bmatrix}\begin{bmatrix}a_1 & b_1\\ a_2 & b_2\end{bmatrix}$$

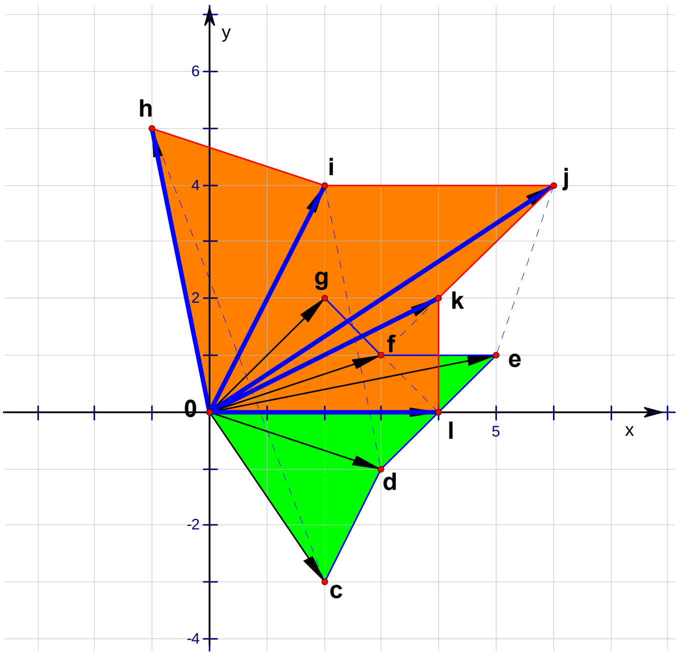
因此，矩阵与矩阵的相乘的几何意义，可以从矩阵与多个向量相乘的几何意义得到，只是多个向量被按照顺序组合成了另一个矩阵。

矩阵乘以矩阵如 $\mathbf{AB}=\mathbf{C}$ 的一般几何意义就是把其中一个矩阵如 $\mathbf{B}$ 的数个行向量或列向量构成的几何图形进行旋转、缩放、镜像等变换（另外一个矩阵 $\mathbf{A}$ 起到的作用）得到了数个新向量，这些新向量作为行向量或者列向量组成一个新的矩阵 $\mathbf{C}$，这个新矩阵 $\mathbf{C}$ 会构成新的几何图形。对于下面的乘式：

$$\mathbf{C}=\mathbf{AB}=\mathbf{A}\begin{pmatrix}\mathbf{c} & \mathbf{d} & \mathbf{e} & \mathbf{f} & \mathbf{g}\end{pmatrix}=\begin{bmatrix}a_1 & a_2\\ b_1 & b_2\end{bmatrix}\begin{bmatrix}c_1 & d_1 & e_1 & f_1 & g_1\\ c_2 & d_2 & e_2 & f_2 & g_2\end{bmatrix}=\begin{bmatrix}h_1 & i_1 & j_1 & k_1 & l_1\\ h_2 & i_2 & j_2 & k_2 & l_2\end{bmatrix}=\begin{pmatrix}\mathbf{h} & \mathbf{i} & \mathbf{j} & \mathbf{k} & \mathbf{l}\end{pmatrix}$$

我们给出具体的数据例子并画出这个变换的图形：

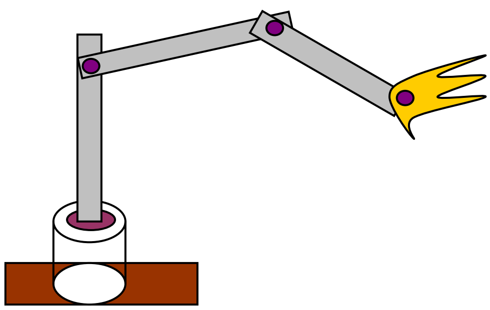
$$\mathbf{C}=\mathbf{AB}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}2 & 3 & 5 & 3 & 2\\ -3 & -1 & 1 & 2 & 2\end{bmatrix}=\begin{bmatrix}-1 & 2 & 6 & 5 & 4\\ 5 & 4 & 4 & 1 & 0\end{bmatrix}$$



一片绿色的小枫叶，经过夏天的风雨洗礼的自然变换，终于成了一片红彤彤的大枫叶。

其实，两个矩阵相乘，我们可以有两种理解。一种理解主要是考察一个矩阵对另一个矩阵所起的变换作用。其作用的矩阵看作是动作矩阵，被作用的矩阵可以看作是由行或列向量构成的几何图形。这个理解就是上面给出的几何解释。

另外一个理解就是两个矩阵都被看作是作用矩阵，两个矩阵的乘积被看作是两个矩阵的和作用。一连串的矩阵相乘如$\mathbf{S}=\mathbf{A}\cdot\mathbf{B}\cdot\mathbf{C}\cdot\mathbf{D}\cdot\mathbf{E}\cdot\mathbf{F}$，会把所有的矩阵线性变化的作用力传递并积累下去，最终得到一个和作用力 $\mathbf{S}$。工业上的例子就是机器人的手臂，机械臂上的每个关节就是一个旋转矩阵（比如可以是一个4×4矩阵），机械臂末端的位置或动作是所有关节运动的综合效果。这个综合效果可以用旋转矩阵的乘法得到。在下面的矩阵乘法运算律一节里面，我们将给个几何图形的例子。



矩阵左乘与右乘的不同

一个矩阵左乘或右乘另一个矩阵的几何图形变化不会相同，因此，矩阵相乘不能满足交换律。也即

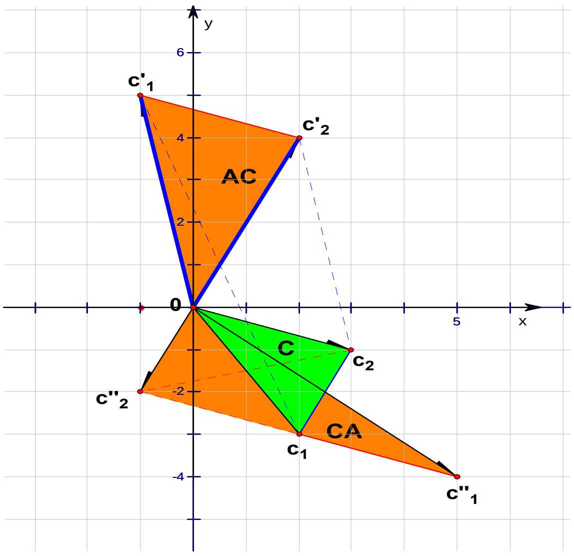
$$\mathbf{AC}\neq\mathbf{CA}$$

下面大家看看矩阵的乘积交换后他们的列向量所构成的三角图形的变化：

$$\mathbf{AC}=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}2 & 3\\ -3 & -1\end{bmatrix}=\begin{bmatrix}-1 & 2\\ 5 & 4\end{bmatrix}$$

$$\mathbf{CA}=\begin{bmatrix}2 & 3\\ -3 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}=\begin{bmatrix}5 & -1\\ -4 & -2\end{bmatrix}$$

矩阵 **C** 的图形标示为三角形 **C**，矩阵 **C** 在矩阵 **A** 的左乘和右乘的作用下分别得到了三角形 **AC** 和 **CA** 两个图形，显然，**AC** 和 **CA** 是不同的。图形如下：



：矩阵相乘 $\mathbf{AB}\neq\mathbf{BA}$ 不能满足交换律，是由矩阵相乘的定义决定的，**A** 乘以 **B** 定义为 **A** 的行向量逐个点乘 **B** 的列向量； **B** 乘以 **A** 定义为 **B** 的行向量逐个点乘 **A** 的列向量；一个矩阵的行向量与其列向量是不同的，因此 $\mathbf{AB}\neq\mathbf{BA}$ 。

一般情况下，矩阵乘法不能交换，但有两个例外要注意：一是单位矩阵可以交换，即 $\mathbf{IA}=\mathbf{AI}=\mathbf{A}$；另一个是逆矩阵可以交换，即 $\mathbf{AA}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}$ 。可以交换的主要原因之一是单位矩阵的行向量等于列向量。

下面我们给个旋转矩阵相乘积的例子。

矩阵乘幂的几何及物理解释

矩阵乘幂运算就是对一个矩阵 **A** 进行一定次数的联乘运算 $\mathbf{A}^n$ 。有实际物理意义的可乘幂运算的矩阵常见的有：Markov(马尔科夫)链中的转移概率矩阵、旋转矩阵和图的邻接矩阵等。在天气的马尔科夫链中，天气转移概率矩阵的 n 次幂 $\mathbf{A}^n$ 表示第 n 天后天气是晴、阴或雨的概率；逆时针旋转 θ 角的矩阵的 n 次幂 $\mathbf{A}^n$ 表示对一个对象旋转了 n 个 θ 角度。邻接矩阵可以表明一个电路网络、电话网络、道路网络以及机构

网络的连接关系，而邻接矩阵的乘幂可以得到任意两个网络节点之间的连接长度及是否有连接的情况。比如邻接矩阵的 2 次幂 $\mathbf{A}^2$ 的结果告诉我们任意两个节点之间是否存在长度为 2 的道路，而且还告诉了我们每一对联结长度为 2 的道路有多少。

对于三个矩阵乘幂的例子中，下面我们较详细的给出了旋转矩阵乘幂的几何解释。

前面讲过，二阶逆时针旋转矩阵表示为

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

那么，两个及至 n 个旋转矩阵的乘积表达式可以由三角公式得到：

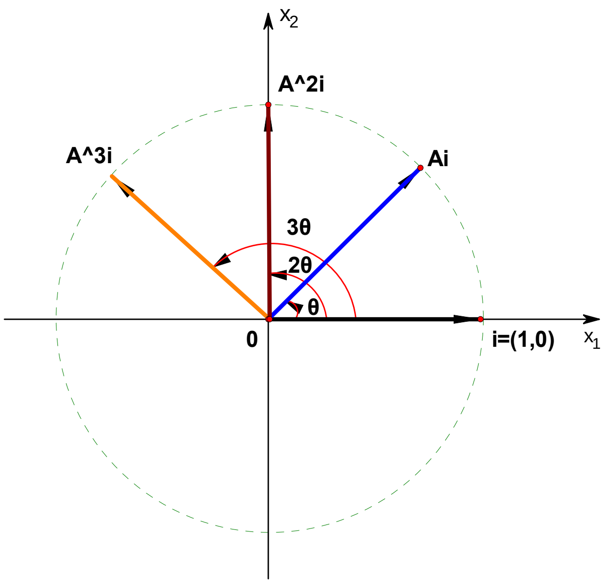
$$\mathbf{A}^2=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}=\begin{pmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{pmatrix}$$

$$\mathbf{A}^3=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos 2\theta & -\sin 2\theta\\ \sin 2\theta & \cos 2\theta\end{pmatrix}=\begin{pmatrix}\cos 3\theta & -\sin 3\theta\\ \sin 3\theta & \cos 3\theta\end{pmatrix}$$

....

$$\mathbf{A}^n=\begin{pmatrix}\cos n\theta & -\sin n\theta\\ \sin n\theta & \cos n\theta\end{pmatrix}$$

由上式看到，旋转矩阵的连乘积等效于旋转角的连加和。下图给出了旋转矩阵 $\mathbf{A},\mathbf{A}^2,\mathbf{A}^3$ 对单位向量 $\mathbf{i}$ 的旋转效果。矩阵的乘法变成了角度的加法。



5.5. 矩阵与线性变换的关系的几何意义

矩阵与线性变换的关系

从前面矩阵的乘法可以知道，任意一个矩阵其本身蕴含着一个变换。这个变换我们可以称为一个矩阵变换。前面章节里我们也讲过线性变换的概念：数域 **F** 上线性空间 **V** 中的变换 **T** 称为 **V** 中的**线性变换**，那么当且仅当满足条件：

$$T(\boldsymbol{\alpha}+\boldsymbol{\beta})=T(\boldsymbol{\alpha})+T(\boldsymbol{\beta})\quad(\boldsymbol{\alpha},\boldsymbol{\beta}\in V)$$
$$T(k\boldsymbol{\alpha})=kT(\boldsymbol{\alpha})\quad(k\in F,\boldsymbol{\alpha}\in V)$$

这里，我们把变换 **T** 的符号替换为矩阵 **A**，线性变换就变成了矩阵变换。

实际上，从 $R^n \to R^m$ 上的线性变换都可以表述为一个矩阵变换；反过来，**一个矩阵变换也必然是一个线性变换。两者具有一一对应的关系**。这个对应关系笼统地表述如下：

- 线性变换的和对应着矩阵的和；
- 线性变换的乘积对应着矩阵的乘积；
- 线性变换的数量乘积对应着矩阵的数量乘积；
- 线性变换的逆对应着矩阵的逆；

因此，对于一个变换我们需要弄懂两个问题：这个变换是否线性变换？如果是线性变换如何求对应的矩阵？

判断是否线性变换只要确定这个变换是否满足加法和数乘法则就可以了。线性变换一般可以用文字描述出来，如“一个旋转角度为 $\pi/8$ 的旋转变换”，或者“一个关于 $x_2=-x_1$ 的镜像变换”等，但如何知道旋转变换的所对应的矩阵呢？这个问题有幸有一个定理给出了回答。作为一个例子，下面的定理给出了如何把一个 R^2 空间上的线性变换转换成一个对应的 2 阶矩阵的办法：

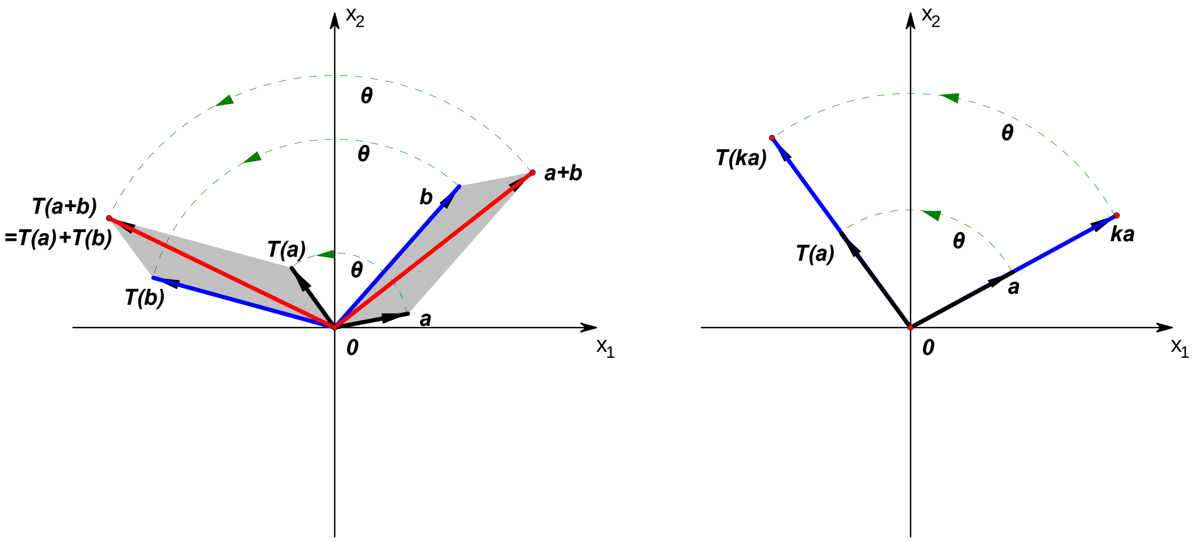
定理：设 $T:R^2\to R^2$ 是任意一个线性变换，那么 T 的矩阵的列向量为 $T(\mathbf{e}_1)$ 和 $T(\mathbf{e}_2)$。

OK，两个问题得到了解答。那我们就验证两个例子加深一下印象。

我们先看一个在平面上的旋转变换的例子（呵，这个例子举了 N 多次了）。

设变换 $T:R^2\to R^2$ 是 R^2 中把每个向量按逆时针方向绕原点 0 转动 θ 角的旋转变换。这里我们要弄清两个问题，一是旋转变换是否是线性变换？二是旋转变换对应的矩阵如何求（注：这个矩阵就是前面多次介绍的平面旋转矩阵）？

判断是否线性变换就是确认这个变换是否满足加法和数乘。我们看下图所示。

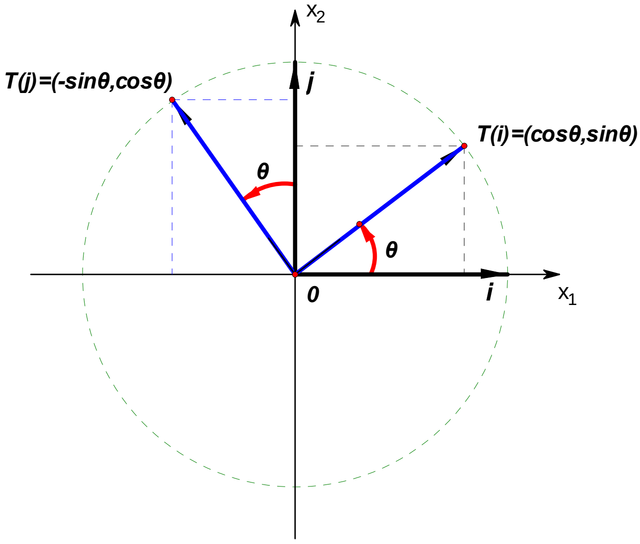


左图中向量 $\mathbf{a}$ 和 $\mathbf{b}$ 同时逆时针旋转了一个 θ 角，因此两个向量之间的夹角没变，长度也没变，显然，构成的平行四边形的外形没变；或者说，平行四边形也同时逆时针旋转了同样的角度。因此旋转前后的平行四边形的对角线长度没变同时也旋转了同样的 θ 角，换句话说，旋转前的平行四边形对角线向量 $\mathbf{a}+\mathbf{b}$ 经过旋转变换 $T()$ 后等于旋转后平行四边形的对角线 $T(\mathbf{a})+T(\mathbf{b})$，亦即 $T(\mathbf{a}+\mathbf{b})=T(\mathbf{a})+T(\mathbf{b})$。这说明满足加法原则。

右图是满足数乘 $T(k\mathbf{a})=kT(\mathbf{a})$ 的图示。为了不侮辱读者的智慧，解释就免了。

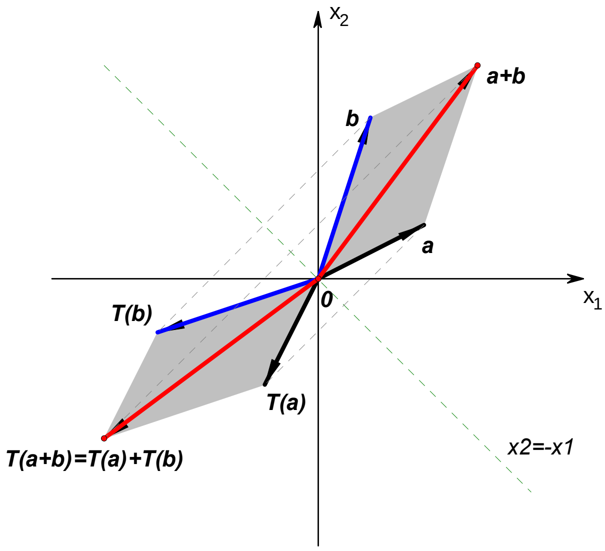
现在我们已经知道了平面空间中的逆时针旋转矩阵的通用矩阵是 $\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$，如何利用定理得到这个旋转矩阵的？实际上我们在前面的多处章节里比如“标准正交基的几何解释”和“旋转矩阵对向量的乘积的几何解释”中已经给出了回答。这里只是套一下定理，给与重新确认：

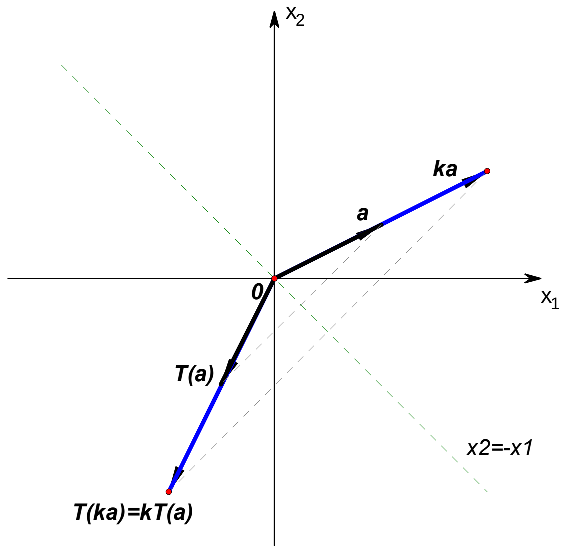
把坐标向量 $\mathbf{e}_1$（有时写作 $\mathbf{i}$）逆时针旋转 θ 角度，即 $T(\mathbf{e}_1)=T(\mathbf{i})=\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}$；同理，把坐标向量 $\mathbf{e}_2$（有时写作 $\mathbf{j}$）逆时针旋转 θ 角度，即有 $T(\mathbf{e}_2)=T(\mathbf{j})=\begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}$。这两个新向量刚好就是旋转矩阵的两个列向量，把他们按顺序组合起来就是旋转矩阵 $\begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}$。



因为这个求矩阵的定理很实用，我们再举一个镜像变换的例子：

设$T:R^2\to R^2$是把向量空间R^2中把每个向量对于直线$x_2=-x_1$作反射的变换。容易看出，这个变换是保持了向量的加法和数乘的原则（如下图，我们看到向量$\mathbf{a}$和$\mathbf{b}$作向量加法的平行四边形在镜像变换$T()$下对于直线$x_2=-x_1$的反射）。

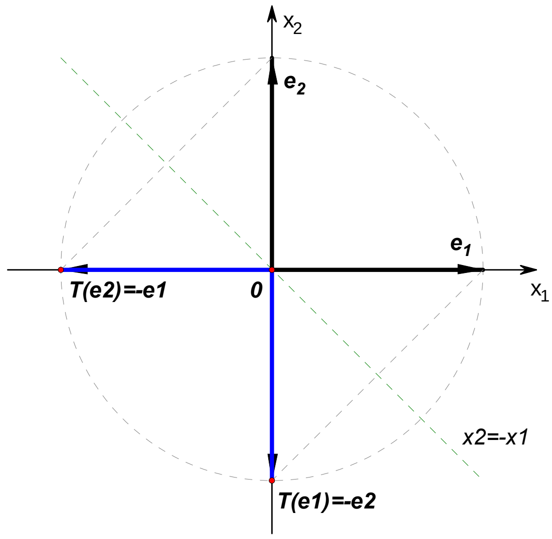




因此镜像变换是一个线性变换。

那么对应镜像变换的矩阵的第一个列向量是$T(\mathbf{e}_1)=T(\mathbf{i})=-\mathbf{j}=\begin{pmatrix}0\\-1\end{pmatrix}$,第二个列向量是

$T(\mathbf{e}_2)=T(\mathbf{j})=-\mathbf{i}=\begin{pmatrix}-1\\0\end{pmatrix}$（见下图）。



所以其镜像矩阵为$\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$。

所以啊，由于线性变换与矩阵之间有一一对应（在给定基的前提下），而且保持线性运算关系不变（线性变换的加法和数乘分别对应于在某一个基下的矩阵的加法和数乘），因此，可以用矩阵来研究线性变换，也可以用线性变换来研究矩阵。

常见的线性变换有初等变换、等价变换、相似变换、合同变换等。我们也常常听到正交变换的名字，但由于正交变换包括平移、旋转和镜像，我们知道平移变换不是线性变换，因此不是所有的正交变换是线性变换。

下面或以后的章节我们将讨论对应着各种线性变换的各种类型矩阵的几何意义。首先看看对应着初等变换的称为初等矩阵的几何意义。

基本初等矩阵/初等变换的几何意义

线性变换与矩阵之间有一一对应关系，因此就有初等矩阵所对应的所谓初等变换。一组向量的线性相关性必然会体现在这组向量的数组之间的关系上，因此我们定能通过向量之间的线性运算把这个关系揭示出来，由于线性运算是在同一分量上进行，人们把一组向量并列在一起，便于对同一分量进行运算，这就是矩阵的初等变换。由此看来，初等变换只对表示向量组的矩阵才能看到其几何意义。在数学中，数域 R 上的矩阵的基本初等变换有 3 种：

（1）交换某两行的位置；

（2）把某一行乘以一个非零数$k(k \in R)$；

（3）把某一行的$k(k \in R)$倍加到另一行上。

对于数域R上的一个单位矩阵分别实施上述3种基本初等变换，所得矩阵分别称为基本初等矩阵(1)、(2)、(3)。基本初等矩阵与矩阵的基本初等变换也是一一对应的；任何一个可逆矩阵都可以分解成基本初等矩阵（1)、(2)、(3）的乘积。同时基本初等矩阵（1）可以由（2）和（3）导出。所以，任何一个可逆矩阵都可以分解成（2）和（3）这两种基本初等矩阵的乘积。从变换的角度来说，一个可逆的线性

变换是连续实施若干次（2）和（3）两种基本初等变换的结果。

：注意：初等变换（2）就是线性变换的数乘变换；初等变换（3）就是线性变换的数乘后的加法，或者讲是线性组合。他们之所以基础就是因为他们是线性变换的数乘和加法的法则。

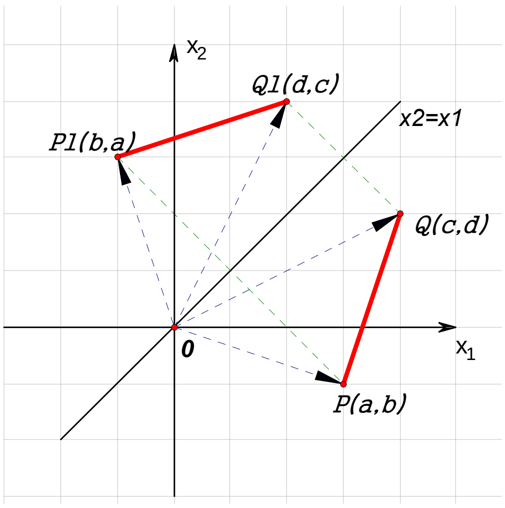
一个二阶矩阵作用在一个二维向量上得到一个新的向量。因此，一个二阶矩阵把平面上的每一个点都变成唯一的点，从而它是平面到平面的映射（即变换），可以刻画平面上的几何变换。同理，任意一个三阶矩阵可以刻画空间中的几何变换。由于任何一个可逆矩阵都可以分解成基本初等矩阵（1）、(2）、(3）的乘积，因此，基本初等矩阵（1）、(2）、(3）的几何意义是理解一般矩阵所表示的变换的几何意义的基础。从几何的观点理解矩阵，把矩阵视为一种几何变换，赋予矩阵一种直观意义。

基本初等矩阵（1）的几何意义是：关于某一“标准轴（面）”的镜像反射（对称）变换；基本初等矩阵（2）的几何意义是：在某一坐标轴方向的伸缩变换；基本初等矩阵（3）的几何意义是：在某一坐标轴方向的切变变换。

- **基本初等矩阵（1）的几何意义**

在二维和三维几何空间中，基本初等矩阵（1）所表示的几何变换是：对图形实施关于某一“标准轴（面）”的镜像反射（对称）变换。其中，标准轴是指在二维几何平面上的直线 $x_2 = x_1$；标准面是指在3维几何空间中的平面 $x_2 = x_1, x_2 = x_3, x_3 = x_1$。

例如：设 $P(a,b)$ 是二维几何平面 $x_2 0 x_1$ 上的任意一点（等同于任意向量），点 P 经基本初等矩阵 $T(1,2)=\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$ 变换后的结果为 $P_1 = T(1,2)P = (b,a)$（如下图），

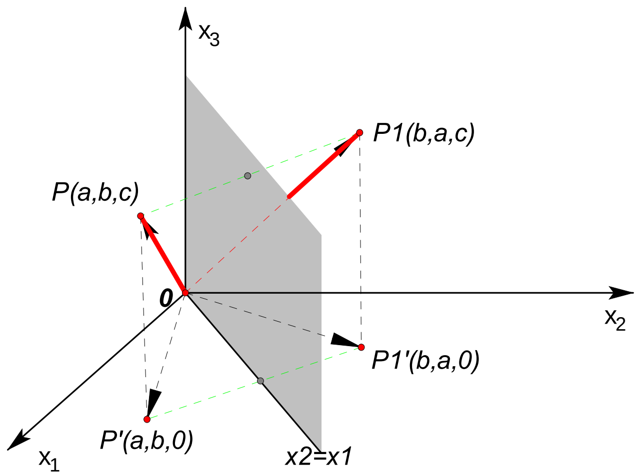


P_1 是点 P 关于直线 $x_2 = x_1$ 的对称点。同样，Q_1 是点 Q 关于直线 $x_2 = x_1$ 的对称点，由此可类推知线

段 P_1Q_1 是 PQ 关于直线 $x_2 = x_1$ 的对称图形。

设 $P(a,b,c)$ 是 3 维几何空间 $0x_1x_2x_3$ 中任意一点，点 P 经基本初等矩阵 $T(1,2),T(2,3),T(1,3)$ 分别变换后，所得结果为：$P_1 = T(1,2)P = (b,a,c)$，$P_2 = T(2,3)P = (a,c,b)$，$P_3 = T(1,3)P = (c,b,a)$。点 P_1 是点 P 关于平面 $x_2 = x_1$ 的对称点、点 P_2 是点 P 关于平面 $x_2 = x_3$ 的对称点、点 P_3 是点 P 关于平面 $x_3 = x_1$ 的对称点。

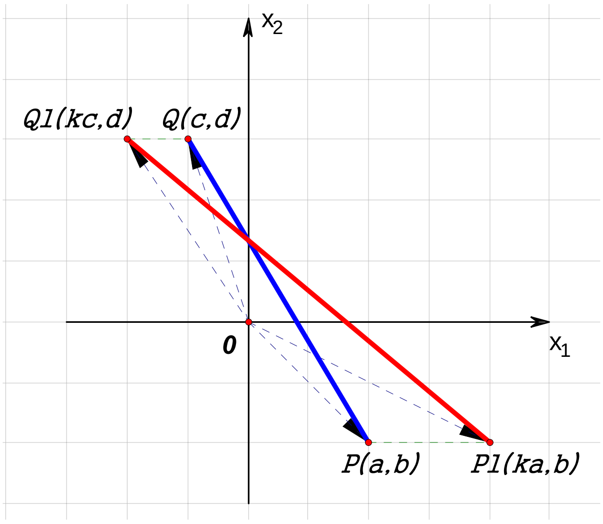
下图绘出了点 P_1 是点 P 关于平面 $x_2 = x_1$ 的对称点。



这个结论可以从二维的例子演推到，点 $P'(a,b,0)$ 和 $P_1'(b,a,0)$ 在二维平面 $x_2 0 x_1$ 上关于直线 $x_2 = x_1$ 镜像对称；扩展到三维空间后，就得到了 $P(a,b,c)$ 和 $P_1(b,a,c)$ 关于平面 $x_2 = x_1$ 镜像对称。

- **基本初等矩阵（2）的几何意义**

在二维和三维几何空间中，基本初等矩阵（2）所表示的几何变换是：对图形实施的在某一坐标轴方向的伸缩变换。例如：设 $P(a,b)$ 是平面 $x_1 0 x_2$ 上的任意一点，点 P 经基本初等矩阵 $T(1(k)) = \begin{pmatrix} k & o \\ 0 & 1 \end{pmatrix}$ 和 $T(2(k)) = \begin{pmatrix} 1 & o \\ 0 & k \end{pmatrix}$ 分别实施变换后结果为：$P_1 = T(1(k))P = (ka,b)$ 和 $P_2 = T(2(k))P = (a,kb)$。点 P_1,P_2 分别是把点 P 在 x_1 轴方向的坐标伸缩 k 倍在 x_2 轴方向的坐标不变，在 x_2 轴方向的坐标伸缩 k 倍在 x_1 轴方向的坐标不变而得到的。



上图给出了两点 $P(a,b),Q(c,d)$ 在 $T(1(2))$ 的作用下沿 x_1 轴伸缩的情况。同时，线段 PQ 经在 $T(1(2))$ 的作用下变化为 P_1Q_1 。

设 $P(a,b,c)$ 是三维几何空间 $0x_1x_2x_3$ 中任意一点，点 P 经三阶基本初等矩阵 $T(1(k))$ ，$T(2(k))$ ，$T(3(k))$ （其中 $k \neq 0$ ）实施变换后，所得结果分别为： $P_1 = T(1(k))P = (ka,b,c)$ ，$P_2 = T(2(k))P = (a,kb,c)$ ，$P_3 = T(3(k))P = (a,b,kc)$ 。P_1 是把点 P 在 x_1 轴方向的坐标伸缩 k 倍，而在 x_2,x_3 轴方向的坐标不变而得到的；P_2 是把点 P 在 x_2 轴方向的坐标伸缩 k 倍，而在 x_1,x_3 轴方向的坐标不变而得到的；P_3 是把点 P 在 x_3 轴方向的坐标伸缩 k 倍，而在 x_1,x_2 轴方向的坐标不变而得到的。若 $k>0$，其伸缩方向与原坐标方向相同；若 $k<0$，其伸缩方向与原坐标方向相反。作为特例，当 $k=0$ 时，基本初等矩阵（2）表示的几何变换是：对图形实施的在某一坐标轴（面）上的投影变换；当 $k=-1$ 时，基本初等矩阵（2）表示的几何变换是：对图形实施的关于某一坐标轴（面）的镜像反射（对称）变换。

例如：矩阵 $T(1(0)),=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ 和 $T(1(-1)),=\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$，对点 $P(a,b)$ 实施变换后分别为：

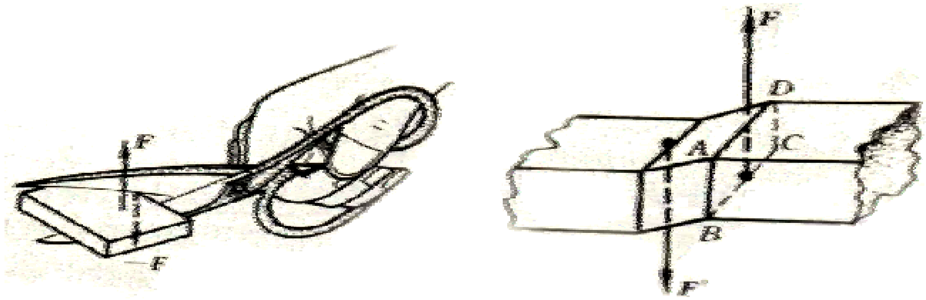
$P_1 = T(1(0))P = (0,b)$ ，是点 $P(a,b)$ 在 x_2 轴上的投影点； $P_2 = T(1(-1))P = (-a,b)$ 是点 $P(a,b)$ 关于 x_2 轴反射点。

3 阶矩阵 $T(1(0)),T(1(-1))$ 对点 $P(a,b,c)$ 实施变换后分别为： $P_1 = T(1(0))P = (0,b,c)$ ，是点 $P(a,b,c)$ 在 $x_2 0 x_3$ 平面上的投影点；$P_2 = T(1(-1))P = (-a,b,c)$ 是点 $P(a,b,c)$ 关于坐标平面 $x_2 0 x_3$ 的反射点。

- **基本初等矩阵（3）的几何意义**

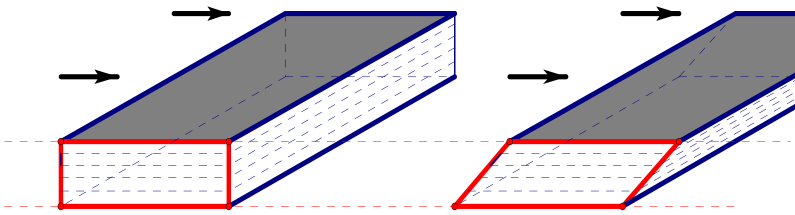
在二维和三维几何空间中，基本初等矩阵（3）所表示的几何变换是：对图形实施的在某一坐标轴方向的切变变换。切变在物理学中属于物体最基本的形变之一，在高等代数和解析几何中很少使用切变这一概念，因此，对于基本初等矩阵（3）的几何意义需要结合物理中的切变来加以说明。

切变，在物理学中指由剪应力作用造成的扭曲（或变形）。如下图 1，当物体受到力偶作用使物体两个平行截面间发生相对平行移动时，在弹性力学中把这种形变叫做剪切形变，简称切变。切变的主要特征为：（1）平行截面间相对滑移；（2）只有纯粹的形状变化，而没有大小（面积或体积）的变化。



剪切形变图

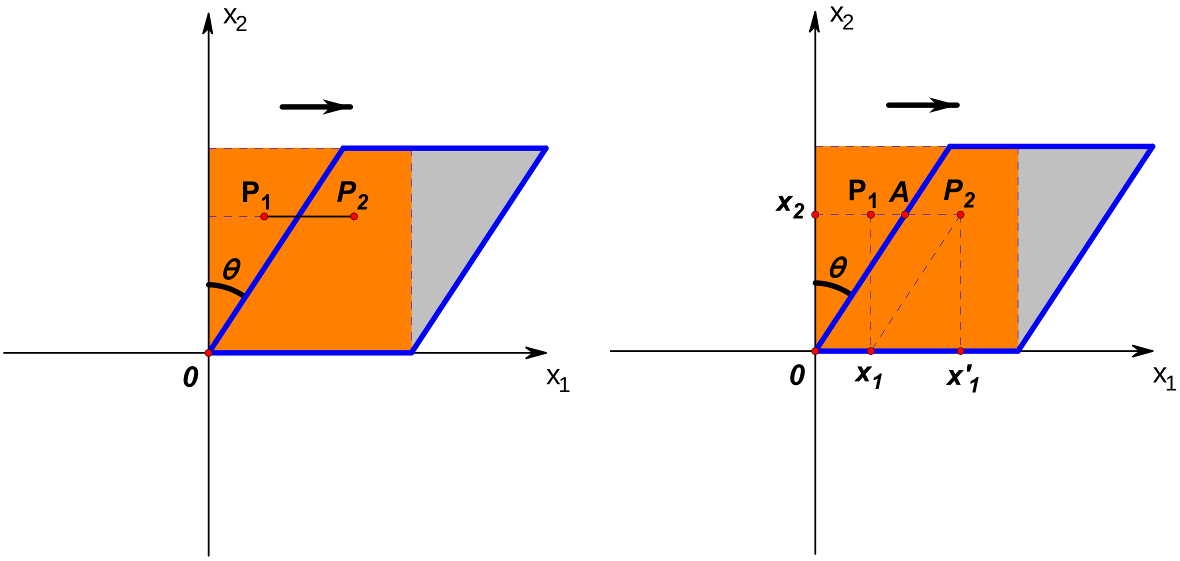
切变的事例在日常生活中有很多。如下图，将一本厚书放在桌面上，推动它的封面，使书页发生滑动，这时在书页的两头边缘上画出的一个矩形变成了平行四边形，这本书受到的就是切变。其形状发生了变化，而体积不变（这本书的宽度和厚度均没有发生变化）。



厚书切变图

二维几何空间中的切变变换矩阵

如下图左，在平面上，一个正方形受到了一个 x_1 轴方向的水平切变力，切变表现为，其力偶作用由正方形变成一个等底、等高的平行四边形。其中，点 $P_1(x_1,x_2)$ 为正方形的任意一点，点 $P_2(x_1^{'},x_2^{'})$ 是切变后 P_1 点所对应的平移点。



由切变的意义，在从点 P_1 到点 P_2 的切变过程中，纵坐标 x_2（高度）不会发生变化，即 $x_2^{'}=x_2$；横坐标 x_1 会向右（力的方向）平移一段距离 P_1P_2。如果把切变的变化量用一个变化的夹角来度量，就可以得到 $P_1P_2=x_2tg\theta$（见上图右，因为两个三角形全等：$\Delta 0Ax_2\cong\Delta x_1P_2P_1$），因此横坐标的变化关系为 $x_1^{'}=x_1+x_2tg\theta$。把 $tg\theta$ 简写为一个变量 $k\ (k\in R)$，那么有：

$$\begin{cases}x_1^{'}=x_1+kx_2\\x_2^{'}=x_2\end{cases},$$

改写成矩阵/向量的表达式为：

$$\begin{pmatrix}x_1^{'}\\x_2^{'}\end{pmatrix}=\begin{bmatrix}1&k\\0&1\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}。$$

类似的，如果正方形切变的方向是竖直 x_2 轴的方向，那么变化前后的坐标表达式：

$$\begin{cases}x_1^{'}=x_1\\x_2^{'}=kx_1+x_2\end{cases},$$

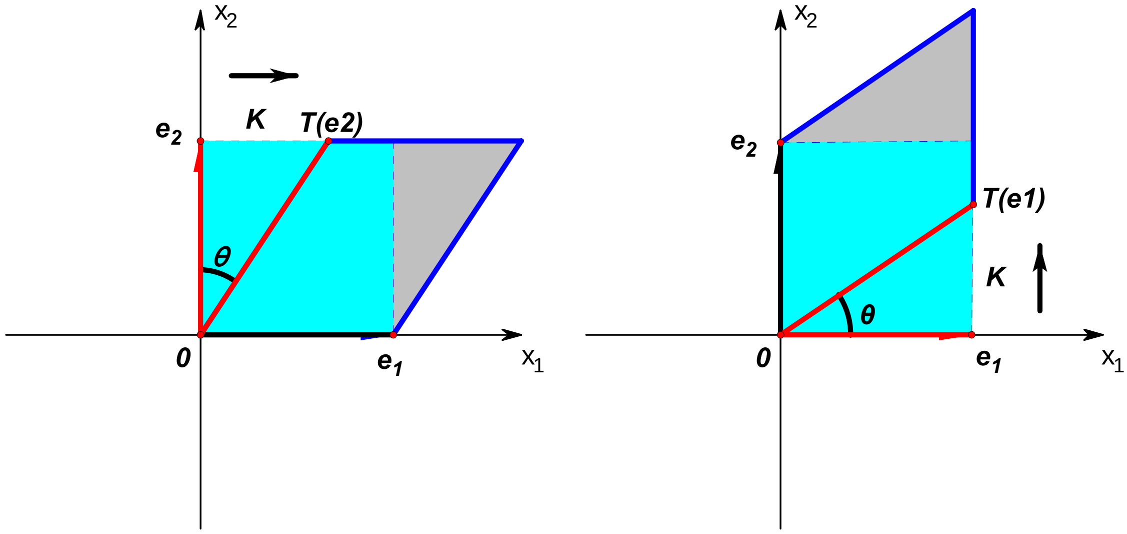
改写成矩阵/向量的表达式为：

$$\begin{pmatrix}x_1^{'}\\x_2^{'}\end{pmatrix}=\begin{bmatrix}1&0\\k&1\end{bmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}。$$

可见，矩阵 $\begin{bmatrix}1&k\\0&1\end{bmatrix}$ 和 $\begin{bmatrix}1&0\\k&1\end{bmatrix}$ 是刻画平面上切变的数学模型，我们可以称其为平面上的切变变换矩阵。

它们是把二阶单位矩阵的某一行的 k 倍加到另一行的基本初等矩阵$T(i, j(k))$。

实际上，如果我们利用上节刚刚知道的定理来快速推算这个基本的初等矩阵，可能让你更清晰简单。即使你忘掉了也可以随时推导出来。下面我们试一下：



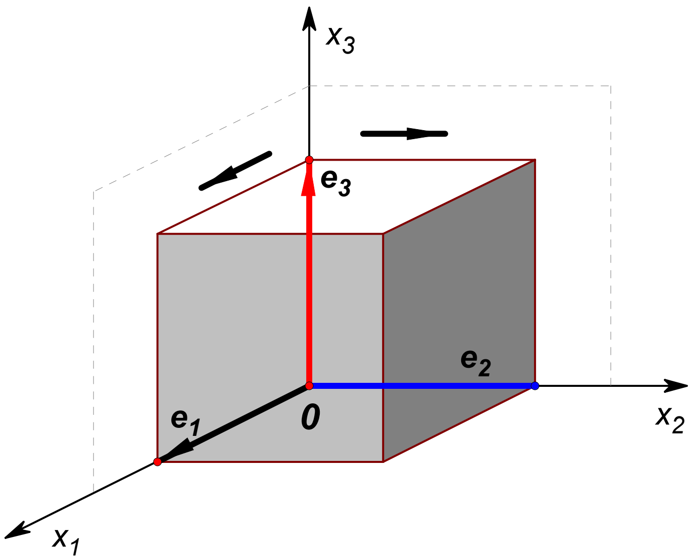
先看上图左，设切变前的正方形为单位长度的，其两个边分别是单位向量$\mathbf{e}_1,\mathbf{e}_2$。那么切变变换后，单位向量$\mathbf{e}_1,\mathbf{e}_2$分别变换为新向量$T(\mathbf{e}_1)$和$T(\mathbf{e}_2)$。由图易看出：向量$\mathbf{e}_1$没有变化，所以$T(\mathbf{e}_1)=\mathbf{e}_1=\begin{pmatrix}1\\0\end{pmatrix}$；向量$\mathbf{e}_2$的高度没变化，但在$x_1$方向移动了距离$k$，那么有$T(\mathbf{e}_2)=\begin{pmatrix}k\\1\end{pmatrix}$。因此有$x_1$方向的切变矩阵$\left(T(\mathbf{e}_1),T(\mathbf{e}_2)\right)=\begin{pmatrix}1&k\\0&1\end{pmatrix}$。

类似的，由上图右可得到x_2方向的切变矩阵$\left(T(\mathbf{e}_1),T(\mathbf{e}_2)\right)=\begin{pmatrix}1&0\\k&1\end{pmatrix}$。

三维几何空间中的切变变换矩阵

在 3 维几何空间中，一个正方体的切变表现为，其受到力偶作用由正方体变成一个等底、等高的平行六面体。

在下面的例子中，我们假设正方体受到了平行于x_1或者x_2轴方向的切变作用力两种情况（见下图）。



当受到平行于 x_1 轴方向的切变作用力时（如下图左），正方体内部的任意一点的坐标中 x_2 和 x_3 不会变，只有 x_1 坐标会发生变化。 x_1 坐标的变化情况与二维平面上的切变类同。

当受到平行于 x_2 轴方向的切变作用力时（如下图右），正方体内部的任意一点的坐标中 x_1 和 x_3 不会变，只有 x_2 坐标会发生变化。 x_2 坐标的变化情况与二维平面上的切变类同。

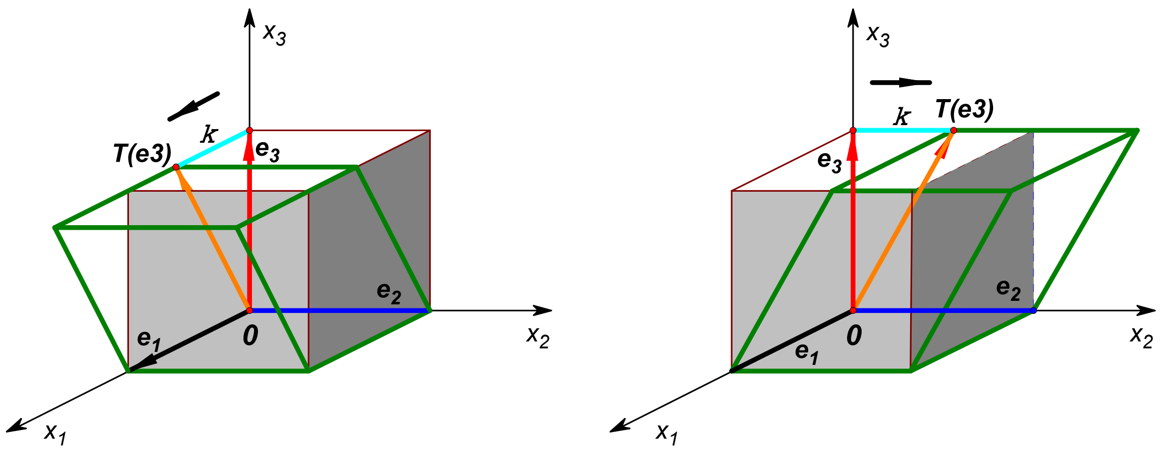


图 5 正方体的切变示意图

当我们应用求变换矩阵的定理求解其矩阵时，就会发现这两种情况都是单位向量 $\mathbf{e}_3$ 发生了变化，而 $\mathbf{e}_1,\mathbf{e}_2$ 没有变化。或者说，单位向量 $\mathbf{e}_3$ 只有两个坐标方向可以切变，要么 $\mathbf{e}_1$ 方向要么 $\mathbf{e}_2$ 方向。因此不难得到左图所示的切变换矩阵为：

$$\left[T(\mathbf{e}_1),T(\mathbf{e}_2),T(\mathbf{e}_3)\right]=\left[\mathbf{e}_1,\mathbf{e}_2,\begin{pmatrix}k\\0\\1\end{pmatrix}\right]=\begin{bmatrix}1&0&k\\0&1&0\\0&0&1\end{bmatrix},$$

其中，k 是单位向量 $\mathbf{e}_3$ 在 $x_1 0 x_3$ 平面上沿 x_1 轴方向的移动距离。

同理，右图所示的切变换矩阵为：

$$\left[T(\mathbf{e}_1),T(\mathbf{e}_2),T(\mathbf{e}_3)\right]=\left[\mathbf{e}_1,\mathbf{e}_2,\begin{pmatrix}0\\k\\1\end{pmatrix}\right]=\begin{bmatrix}1&0&0\\0&1&k\\0&0&1\end{bmatrix},$$

其中，k 是单位向量 $\mathbf{e}_3$ 在 $x_2 0 x_3$ 平面上沿 x_2 轴方向的移动距离。

至此，还有四种切变的情况没有得到矩阵：因为单位向量 $\mathbf{e}_1$ 也有两个坐标方向可以切变，要么 $\mathbf{e}_2$ 方向要么 $\mathbf{e}_3$ 方向；单位向量 $\mathbf{e}_2$ 亦有两个坐标方向可以切变，要么 $\mathbf{e}_1$ 方向要么 $\mathbf{e}_3$ 方向。如果要把其余四种切变的情况罗列出来就是：

$\mathbf{e}_1$ 沿 $\mathbf{e}_2$ 方向切变：$\left[T(\mathbf{e}_1),\mathbf{e}_2,\mathbf{e}_3\right]=\left[\begin{pmatrix}1\\k\\0\end{pmatrix},\mathbf{e}_1,\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\k&1&0\\0&0&1\end{bmatrix}$；

$\mathbf{e}_1$ 沿 $\mathbf{e}_3$ 方向切变：$\left[T(\mathbf{e}_1),\mathbf{e}_2,\mathbf{e}_3\right]=\left[\begin{pmatrix}1\\0\\k\end{pmatrix},\mathbf{e}_1,\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\0&1&0\\k&0&1\end{bmatrix}$；

$\mathbf{e}_2$ 沿 $\mathbf{e}_1$ 方向切变：$\left[\mathbf{e}_1,T(\mathbf{e}_2),\mathbf{e}_3\right]=\left[\mathbf{e}_1,\begin{pmatrix}k\\1\\0\end{pmatrix},\mathbf{e}_2\right]=\begin{bmatrix}1&k&0\\0&1&0\\0&0&1\end{bmatrix}$；

$\mathbf{e}_2$ 沿 $\mathbf{e}_3$ 方向切变：$\left[\mathbf{e}_1,T(\mathbf{e}_2),\mathbf{e}_3\right]=\left[\mathbf{e}_1,\begin{pmatrix}0\\1\\k\end{pmatrix},\mathbf{e}_2\right]=\begin{bmatrix}1&0&0\\0&1&0\\0&k&1\end{bmatrix}$；

因此，以上六个变换矩阵是把三阶单位矩阵的某一行的 k 倍加到另一行的基本初等矩阵 $T(i, j(k))$。作为特例，若 $k=0$，则基本初等矩阵（3）为单位矩阵，其几何意义是图形到其自身的恒等变换。

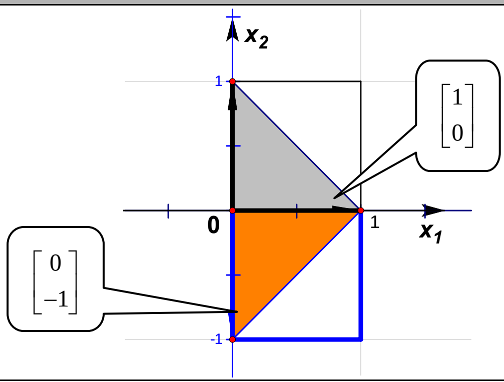
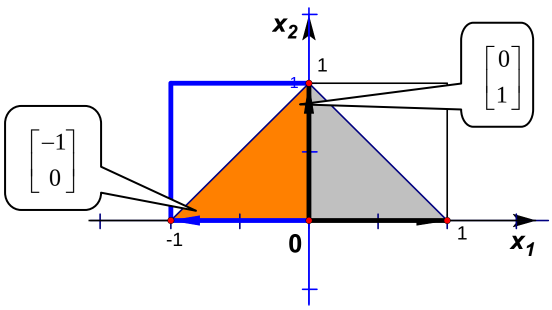
矩阵及其对应线性变换的几何图形

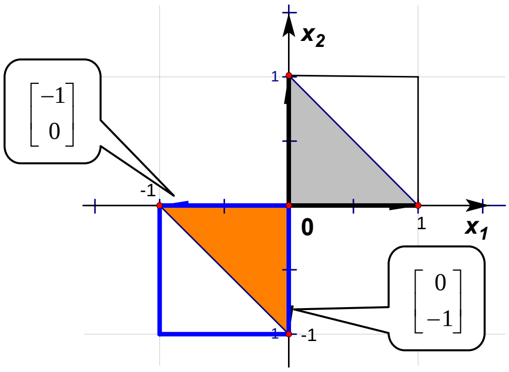
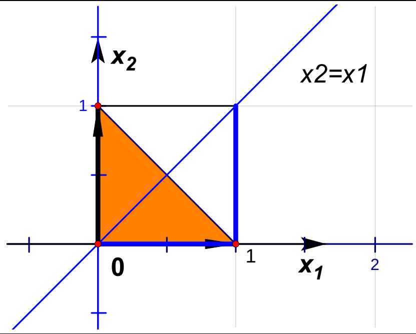
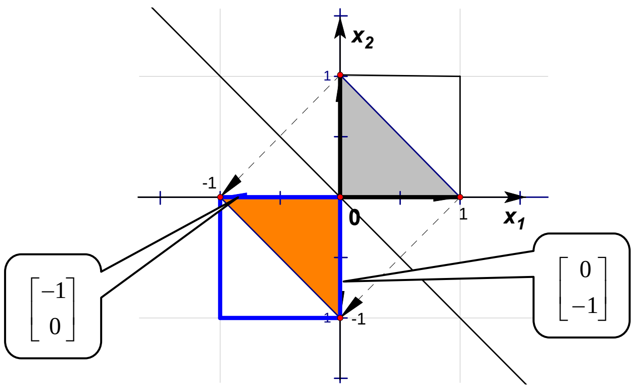
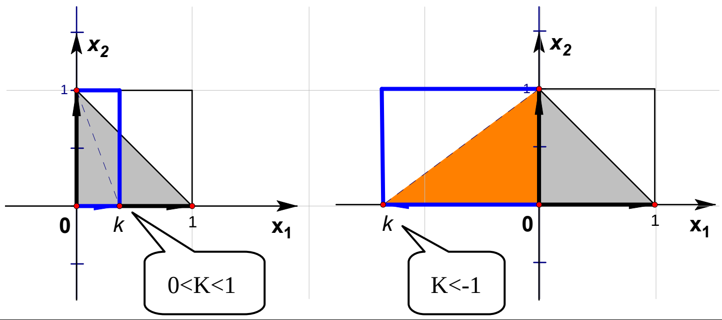
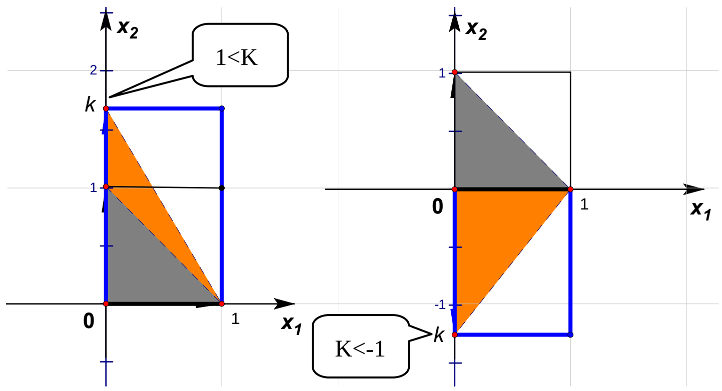
如果给你一个矩阵，如何马上理解其几何意义上的线性变换呢？其实很简单，你只要把前面已知变换求矩阵的的定理反着用就可以了。也就是说，把一个矩阵的列向量所张成平行多面体和同阶的单位矩阵的单位向量所张成的正多面体比对，就可以发现其对应的变换关系。一般情况下，我们只能对二阶和三阶矩阵的几何图形进行明显的对比。从前面的例子我们知道，二阶和三阶矩阵的正多面体就是正方形和正立方体。

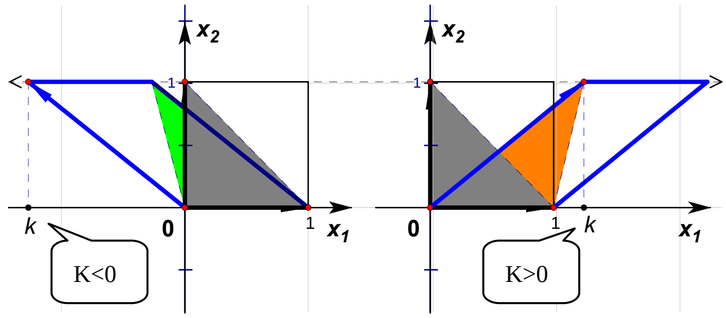
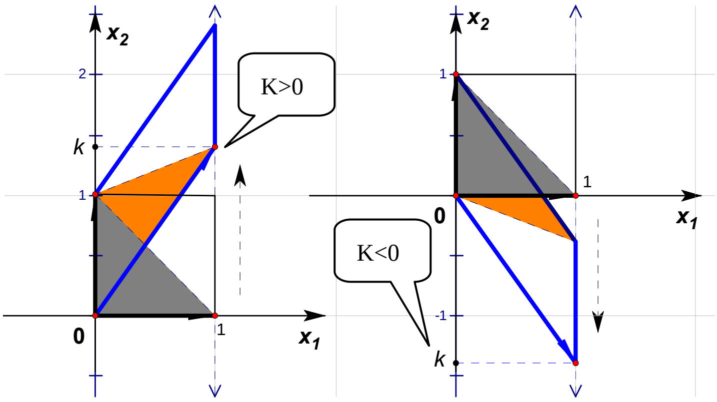
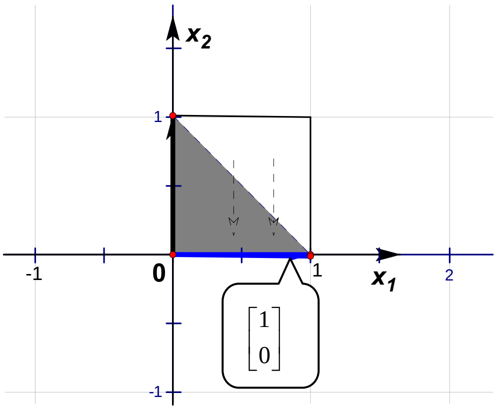
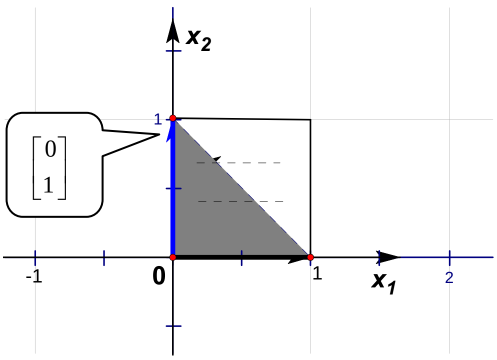
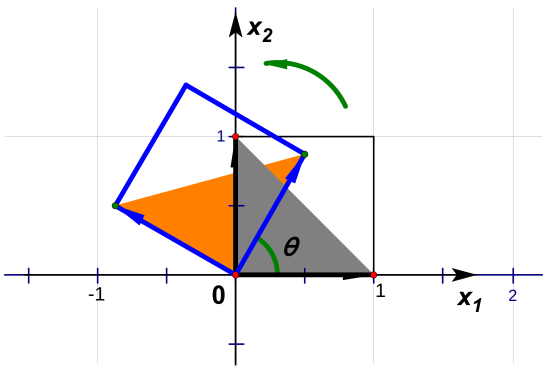
一个线性几何图形乘以一个矩阵，就会对这个图形进行一个线性变换或矩阵变换。几何图形是一个向量或者是一个三角形或者是一个四边形或者一个园，这些几何图形都被认为是一个向量集合所构成。当一个矩阵对众多不同的几何图形进行变换时就会显现不同的变换特征，同时也会加深我们对一个具体矩阵的意义的了解。比如我们常常见到二阶矩阵对一个单位正方形的图形进行变换，实际上就是把单位矩阵 **I** 的行向量或列向量所张成几何图形进行的一个变换。

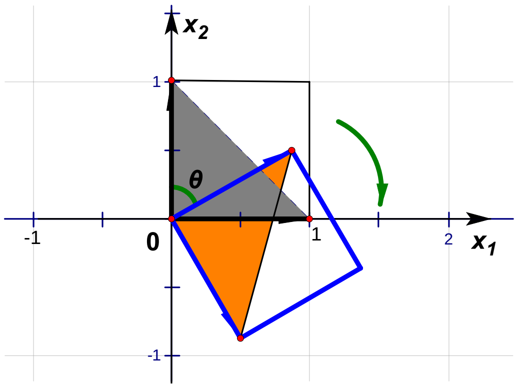
另外，如果被变换的几何图形是一个向量园，那么就会容易发现矩阵的特征值和特征向量的特质如果被变换的几何图形是一个正方形其特征向量的几何特点就不太明显。因此在下节的矩阵的特征向量和秩的几何意义的讨论中，我们总是把被变换的几何图形设为单位园或单位球。

下面我们给出各种二阶矩阵所对应的线性变换的平面几何图形，这里假设被变换的图形是一个位于第一象限内的三角形或正方形，他们由单位矩阵$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$所生成。这个汇总可帮助大家有个系统的认识。另外，在图标的最右列，给出了线性变换对应矩阵的行列式值，由行列式的值我们可以看出变换后的平面图形的面积的变化比率。

矩阵	变换前后的图像	对应的线性变换	矩阵的行列式
$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$		关于 x_1 轴的镜像对称变换	-1
$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$		关于 x_2 轴的镜像对称变换	-1

$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$		关于原点的对称变换	1
$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$		关于直线 $x_2=x_1$ 的对称变换	-1
$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$		关于直线 $x_2=-x_1$ 的对称变换	-1
$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$		x_1 轴（水平）方向的伸缩变换	k
$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$		x_2 轴（垂直）方向的伸缩变换	k

$\begin{bmatrix}1 & k\\0 & 1\end{bmatrix}$		x_1 轴（水平）方向的剪切变换	1
$\begin{bmatrix}1 & 0\\k & 1\end{bmatrix}$		x_2 轴（垂直）方向的剪切变换	1
$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$		x_1 轴（水平）方向的投影变换	0
$\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$		x_2 轴（垂直）方向的投影变换	0
$\begin{bmatrix}\cos\theta & -\sin\theta\\\sin\theta & \cos\theta\end{bmatrix}$		逆时针旋转变换	1

$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$		顺时针旋转变换	1

前面的行列式一章中知道，二阶行列式的几何意义就是两个列向量或行向量所构成的平行四边形的有向面积，实际上，二阶行列式的几何意义也可以看作是两个列向量或行向量所构成三角形的有向面积的 2 倍（如上表中的图形）。看作平行四边形或者是三角形都可以，这不会影响我们对矩阵的理解。

在上表中，我们得到一些结论：

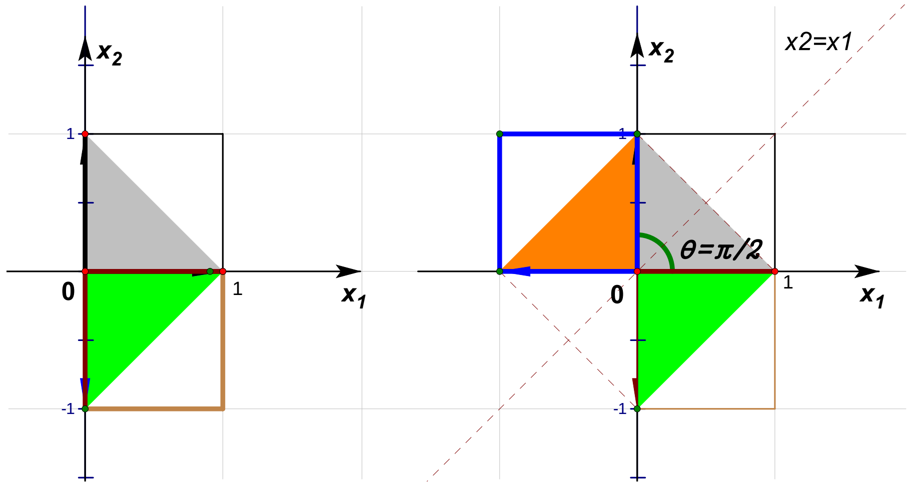
- 行列式等于 1，表示变换后图形的面积乘以 1，面积大小不会变化，图形的有向性也不会变化（通俗的讲，图面的正反面不会反转）；
- 行列式等于-1，表示变换后图形的面积乘以-1，面积大小不会变化，但图形的有向性会变化（通俗的讲，图面的正反面发生反转）；
- 行列式等于 0，表示变换后图形的面积乘以 0，面积收缩为 0，图形缩变为一维图形；
- 行列式等于 k，表示变换后图形的面积会变化 k 倍，K 大于 1，图形面积会放大，k 小于 1，图形面积会缩小，图形的有向性也会因 k 是否负值而发生变化；

矩阵的乘法运算律的几何意义

与前面的章节比较起来，本节利用了线性变换的概念来再次回顾矩阵乘积运算的几何意义，由此可更深入地认识矩阵运算性质。

- **两个矩阵相乘是两个线性变换的复合**

例如，对于平面坐标系 $x_1 0 x_2$ 上由单位矩阵所张成的第一象限的正方形实施关于 x_1 轴的反射变换 $\mathbf{A}=\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$（见下图左）再实施关于标准轴 $x_2 = x_1$ 的反射变换 $\mathbf{B}=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$，得到的图形正好是将原来的正方形逆时针旋转 90° 得到的（见下图右）。



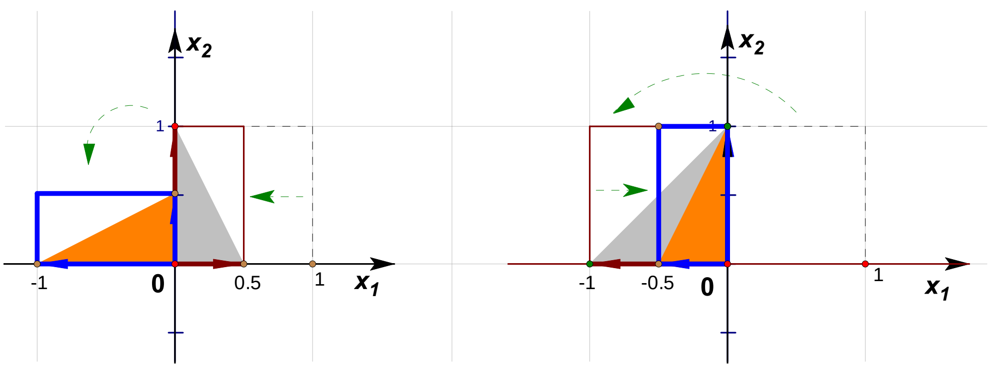
这一结果用矩阵运算表示为 $\mathbf{BA}=\begin{bmatrix}0&1\\1&0\end{bmatrix}\begin{bmatrix}1&0\\0&-1\end{bmatrix}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$，即两个矩阵相乘得到一个新矩阵，新矩阵所表示的变换正是原来两个矩阵表示的变换的复合。这正是矩阵乘法的意义，即两个矩阵相乘表示连续两次实施变换。这一特例同时也可以帮助大家认识“旋转变换可以由连续两次反射变换来实现”这一性质。

- 矩阵的乘法不满足交换律

根据上述对矩阵乘法的理解，容易直观得出矩阵运算满足结合律。是否满足交换律呢？仍然看上面的例子。

如果先对第一象限的正方形进行实施关于标准轴 $x_2=x_1$ 的反射变换 $\mathbf{B}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$，再实施关于 x_1 轴的反射变换 $\mathbf{A}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}$，得到的图形将是将原来的正方形顺时针旋转 90° 的结果。这与前面的结论不同，因为其中 $\mathbf{B}$ 的变换并没有对图形的外形造成变化，因而 $\mathbf{BA}\neq\mathbf{AB}$。

我们再看个例子。同样的，对在第一象限的由单位矩阵列向量所张成的正方形图形，先向靠近 x_2 轴的方向压缩一半，再逆时针旋转 90° 得到的结果（见下图左）与先逆时针旋转 90°，再向靠近 x_2 轴的方向压缩一半得到的结果（见下图右）的比较，大家应确信交换变换的顺序得到的结果一般来说是不同的。



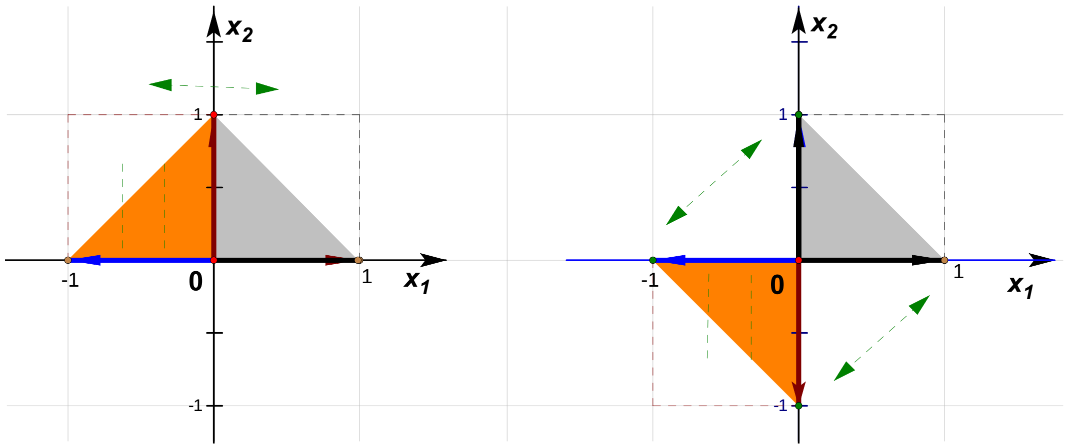
即：

$$\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 0.5\end{bmatrix}\neq\begin{bmatrix}1 & 0\\0 & 0.5\end{bmatrix}\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$$

因此，矩阵的乘法一般不满足交换律。但需要注意的是，不过有些情形矩阵的乘法可以满足交换律，例如连续多次旋转或连续两多次压缩变换是可以交换顺序的。

- **矩阵的乘法不满足消去律**

根据投影变换的特点，把一个图形先作关于 x_2 轴的反射变换，再向 x_1 轴作投影变换的结果（见图左），与先作关于坐标原点的对称变换，再向 x_1 轴作投影变换得到的结果（见图右）是一样的。



然而，虽然有：$\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$，但 $\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}\neq\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}$。因此，矩阵的乘法不满足消去律。待续。。。