线性代数基础知识

刘闯 (武汉大学计算机学院)
1659608216@qq.com

2019 年 9 月 26 日

摘要

本文为 3blue1brown 线性代数系列视频课程笔记，所有内容均来自 3blue1brown。

link：https://space.bilibili.com/88461692?from=search&seid=6764604302374459112

1 什么是向量

1.1 视角

1. **Physics**:

物理学眼中的向量，看作空间的一个箭头。使用长度和方向表征这个向量。这个向量和起点无关。如图 1(a) 所示，起点不同，但是都是向量 $\vec{V}$.

2. **Computer science**:

计算机科学眼中的向量是一个数字列表，用来表征特征。例如下面的矩阵，第一行表示房子的面积，第二行表示房子的价格

$$\begin{bmatrix} 100 \\ 300,000 \end{bmatrix}$$

1.2 基本运算

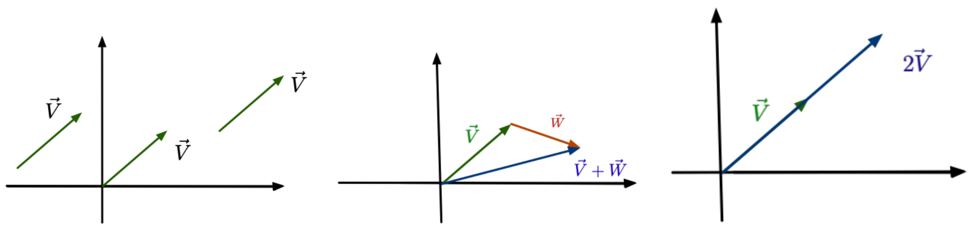
向量的加法: 如图 1 (b) 所示，我们可以将其看作一个移动。点先按照向量 $\vec{V}$ 进行移动，之后按照 $\vec{W}$ 进行移动

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \end{bmatrix} \tag{1}$$

向量数乘：如图 1 所示，对向量进行缩放。

$$2\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 2v_1 \\ 2v_2 \end{bmatrix} \tag{2}$$

这两种运算虽然很基础，但是是向量的基本，很重要，后文的很多理解都基于此。



(a) 向量　(b) 向量加法　(c) 向量数乘

图 1: 向量运算

2　线性组合，空间和基

2.1　基

定义: A set of linearly independent vectors that span the full space.

简单来说，我们通过这组基向量的运算（加法，数乘）可以表示空间中任何一个点。

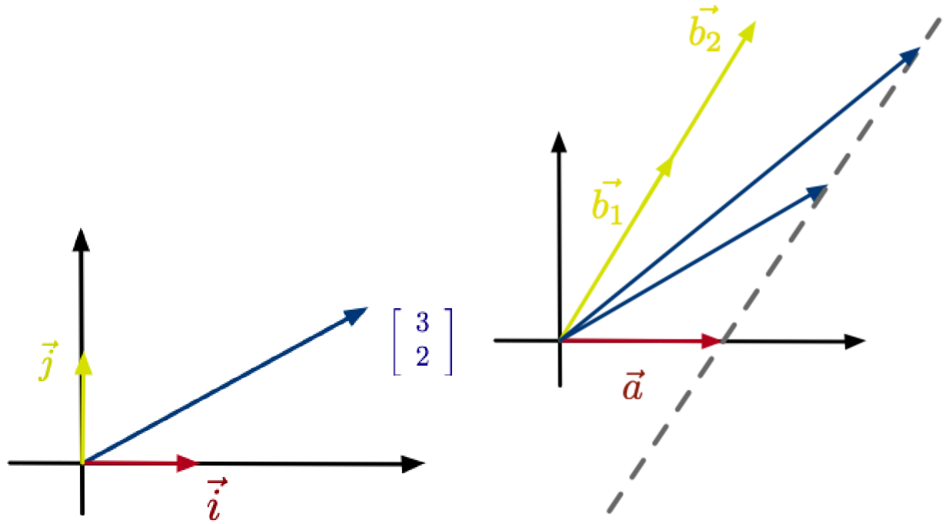
实际上在上面我们使用矩阵表示向量的时候，默认使用了一组基向量。

从另外一个角度，我们可以理解为，向量是基向量的缩放并相加。如图 2(a) 所示，向量 $\vec{i}$ 放大 3 倍，$\vec{j}$ 放大两倍，之后相加，缩放的大小使用矩阵进行表示。

2.2　线性组合

$$a\vec{V} + b\vec{W} \tag{3}$$

线性: 简单直观的理解，如果固定一个向量 $(\vec{a})$，另一个向 $(\vec{b})$ 的标量任意变化，产生的点是一条直线。如图2(b)



(a) 基　　(b) 线性组合

图 2: 向量线性组合

严格定义:

$$L(\vec{V}+\vec{W}) = L(\vec{V}) + L(\vec{W}) \quad L(c\vec{V}) = cL(\vec{V})$$

2.3 张成的空间 span

定义 The span of $\vec{V}$ and $\vec{W}$is the set of all their linear combination.

考虑只使用向量加法和数乘，对于 $\vec{V}$ *and* $\vec{W}$ 能够得到的向量集合是什么？

1. $\vec{V}$ and $\vec{W}$ 不共线，可以得到整个二维空间

2. 如果共线，只能得到一条线

对于 $\vec{V}$ ，$\vec{W}$ ，$\vec{U}$ 能够得到的向量集合是什么。

$$a\vec{V} + b\vec{W} + c\vec{U}$$

1. 如果 $\vec{U}$ 不在 $\vec{V}$ and $\vec{W}$ 的平面内，可以得到一个三维空间
2. 如果 $\vec{U}$ 在 $\vec{V}$ and $\vec{W}$ 的平面内，只能得到一个平面

$$\vec{U} = a\vec{V} + b\vec{W}$$

3 矩阵和线性变换

3.1 线性变换

3.1.1 基本理解

我们可以将线性变换看作一个函数，输入一个向量，输出一个向量，如图 3 所示。

但是，这个函数有两个很**重要的性质**：

1. 直线变换后还是直线，如果将空间看作网格线构成，变换之后网格线仍然平行且等距。

2. 原点保持固定，变换只是对空间的挤压和伸展

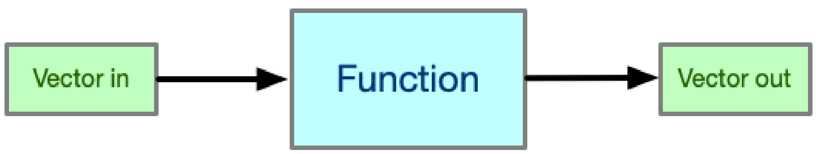


图 3: 线性变换

3.1.2 数值表示

如何使用数值表示这种变换？

根据变换的性质，网格线平行而且等距，而且，我们知道向量可以看作对基的缩放之后加和。所以我们只需要**关注基的变换**

例如：

$$\vec{V} = -1\vec{i} + 2\vec{j}$$

经过变换之后:

$$Transformed\vec{V} = -1(Transformed\vec{i}) + 2(Transformed\vec{j})$$

线性变换只和和变换后的基相关:

$$Transformed\vec{i} = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

$$Transformed\vec{j} = \begin{bmatrix} x_j \\ y_j \end{bmatrix}$$

所以这个变换可以使用一个矩阵进行表示:

$$\begin{bmatrix} x_i & x_j \\ y_i & y_j \end{bmatrix}$$

给定任意一个向量 $[5,3]$，计算其变换之后的

$$5\begin{bmatrix} x_i \\ y_i \end{bmatrix} + 3\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} 5x_i + 3x_j \\ 5y_i + 3y_j \end{bmatrix}$$

由此我们可以推出 **矩阵乘法**

$$\begin{bmatrix} x_i & x_j \\ y_i & y_j \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} ax_i + bx_j \\ ay_i + by_j \end{bmatrix} \tag{4}$$

矩阵的每一列都可以看作变换之后的基向量

例如: *将空间旋转 90 度*

如图 4 所示，变换后的基坐标为:

$$\vec{i} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \vec{j} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

所以变换矩阵我们可以写为:

$$Trans = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

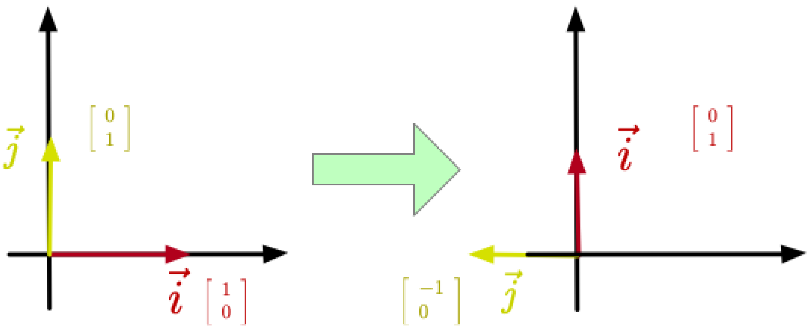


图 4: 空间旋转 90 度

反过来，给定矩阵考虑变换

例如：给定一个变换矩阵

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

其实相当于将 $\vec{i}$, $\vec{j}$ 移动到相应的坐标位置 $[1,2]$, $[3,1]$, 空间内其他的所有线都一起移动，保持平行等距。

假设变换矩阵是

$$Trans = \begin{bmatrix} 1 & -1 \\ 2 & -2 \end{bmatrix}$$

变换后的基 $\vec{i}$, $\vec{j}$ 是线性相关，或者说是共线的，两个基共线，说明**线性变换将二维空间压缩为一条直线**

4 矩阵乘法和线性变换

4.1 线性变换

不同于上一节的单一变换，这一节我们考虑连续进行 2 个或者多个变换，称为 **复合变换**

和上一节同样的想法，我们关注的是两次变换的基向量

例如，我们先进行一个旋转变换，再进行一个剪切变换

剪切变换如图 5 所示，X 轴向量保持不变，y 轴向量旋转到 [1,1] 位置。

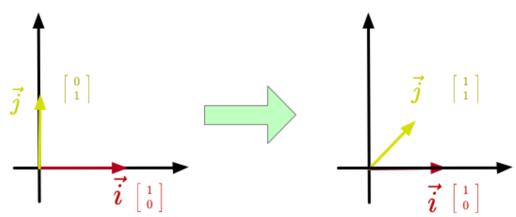


图 5: 剪切变换

复合操作如图 6 所示：(先进行一个旋转变换，再进行一个剪切变换)

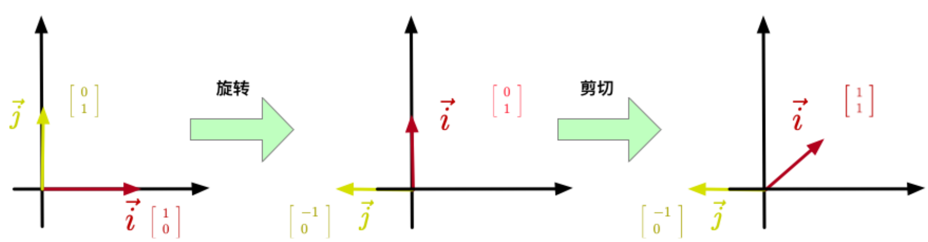


图 6: 复合变换

利用矩阵计算复合操作的复合矩

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

可以发现，矩阵计算的结果的列正是我们复合操作后的基向量的坐标。
所以：**矩阵相乘的几何意义可以理解为两个线性变换相继进行**
Note：两个矩阵的顺序也是类似于函数 $f(g(x))$，后进行的写在外面

4.2 矩阵相乘计算

给定任意两个矩阵

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

计算矩阵相乘的结果只需要关注 **基的变换**。
基 $\vec{i}$ 变换结果:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e \\ g \end{bmatrix} = \begin{bmatrix} ae + bg \\ ce + dg \end{bmatrix} \tag{5}$$

基 $\vec{j}$ 变换结果:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} f \\ h \end{bmatrix} = \begin{bmatrix} af + bh \\ cf + dh \end{bmatrix} \tag{6}$$

最终根据 5 和6复合矩阵：

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \tag{7}$$

4.3 矩阵相乘两个性质

1. $M_1M_2 \neq M_2M_1$

   因为两次变换的顺序不同，得到的结果不同。比较图 6 和图 7。

2. $(AB)C = A(BC)$

   使用空间变换进行思考:

$(AB)C$：先进行 C 变换，再进行 B 变换，再进行 A 变换。
$A(BC)$：先进行 C 变换，再进行 B 变换，再进行 A 变换。

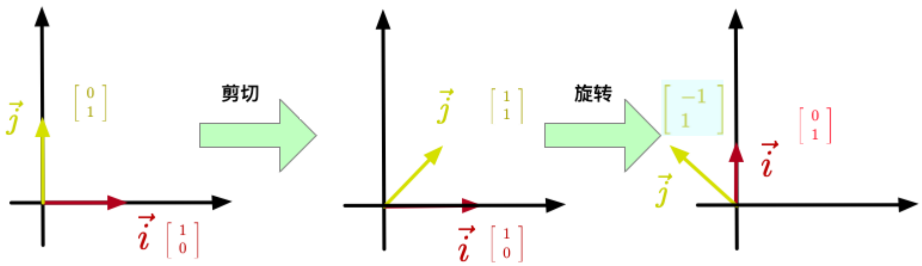


图 7: 复合变换--调整变换顺序

5 行列式

线性变换相当于对空间进行拉伸或者挤压，但是我们不知道具体拉伸和挤压了多少，即变换之后的单元格的面积。

5.1 基区域面积变化

例: 给定 2 个变换矩阵

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

观察图 8，得到其面积变换：一个基区域面积变为 6 倍，一个基区域面积不变

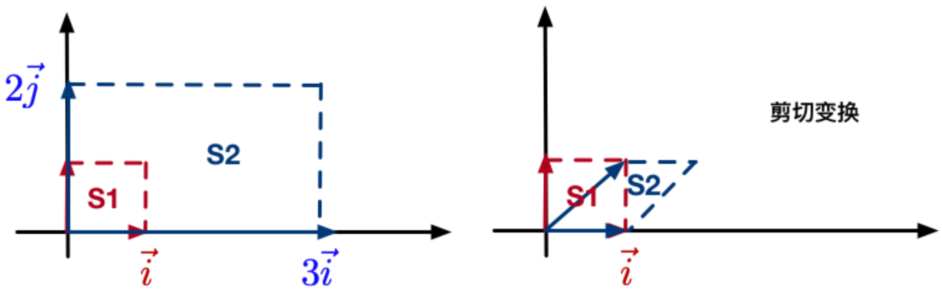
因为网格线平行等距，所以基区域面积的变化等价于所有区域.

5.2 行列式的值

矩阵行列式等于经过矩阵的变换前后，每个区域面积的变化。

例如：

$$det(\begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}) = 6$$



(a) 矩阵 1 变换　　(b) 剪切变换

图 8: 基区域面积变化

1. 线性变换矩阵的行列式是 6，即这个变换将每个区域的面积变为原来的 6 倍。

$$det(\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}) = 0$$

2. 如果行列式的值是 0 ，说明变换**变换将平面压缩至 1 维**

观察矩阵，就是前文我们介绍的线性变换后基向量共线，会压缩空间。

检查行列式值是否为 0 ，优劣与帮助我们判断是否将空间压缩至更低的维度

3. 如果行列式的值出现负值，说明空间取向发生反转 (不再是右手系)，如图 9

5.3　三维空间

行列式在三维空间中等价于体积的缩放

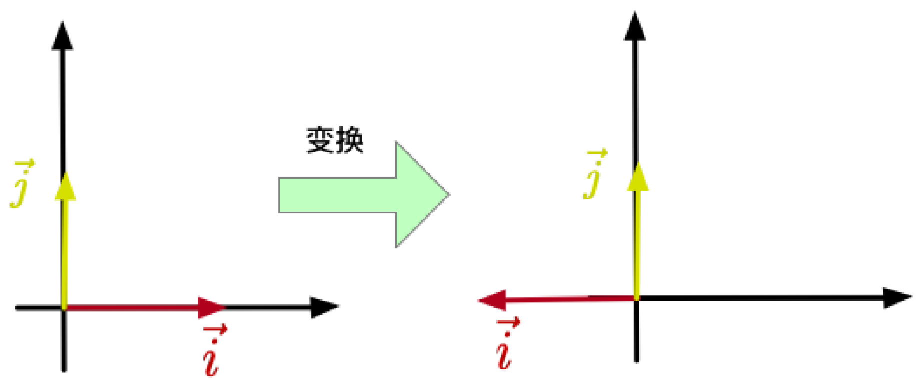


图 9: 空间反转

1. 行列式为 0 , 说明空间被压缩为一个面，一条线，或者一个点。

2. 行列式为负，说明变换后的左边系不符合右手定则

5.4 行列式计算

定义

$$det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ad - cb \tag{8}$$

为什么是这样计算？

1. 假设 $b = c = 0$, 就如前面的介绍，$s = det = ad$

2. 假设 b, c 中有一个为 0, $s = det = ad$

3. 如果 b，c 均不为 0，如图10,

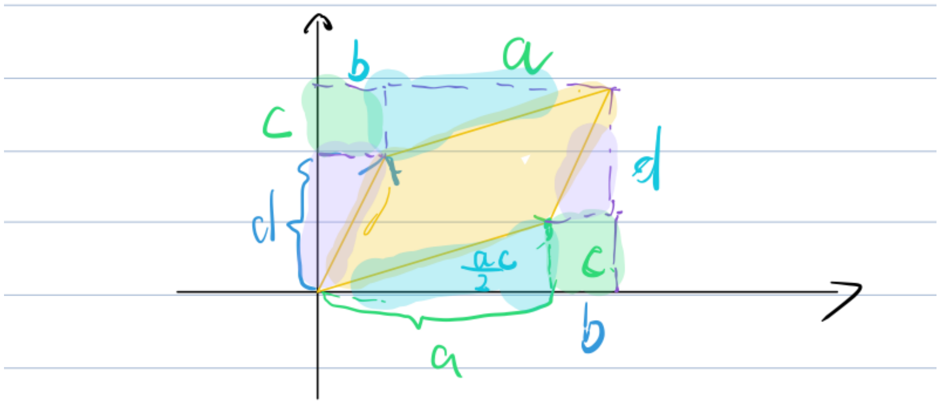


图 10: 行列式计算

$$det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = (a+b)(c+d) - ac - bd - 2bc = ad - bc$$

5.5 3 阶行列式计算

$$det(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}) = a * det(\begin{bmatrix} e & f \\ h & i \end{bmatrix}) - b * det(\begin{bmatrix} d & f \\ g & i \end{bmatrix}) + c * (\begin{bmatrix} d & e \\ g & h \end{bmatrix}) \tag{9}$$

5.6 行列式性质

$$det(M_1 M_2) = det(M_1) * det(M_2) \tag{10}$$

从空间变换的角度理解证明:

对于一个区域，先连续缩放两次求缩放总面积，和缩放第一次的面积乘缩放第二次的面积等价。

$$S \times a \times b = S \times (a \times b)$$

6 逆矩阵，秩，列空间

6.1 逆矩阵

给定一个线性方程组，我们可以表示为矩阵运算。

$$\begin{aligned}2x + 5y + 3z &= -3\\4x + 0y + 8z &= 0\\1x + 3y + 0z &= 2\end{aligned}$$

使用矩阵进行表示：

$$\begin{bmatrix}2 & 5 & 3\\4 & 0 & 8\\1 & 3 & 0\end{bmatrix} \cdot \begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}-3\\0\\2\end{bmatrix}$$

简单抽象为：

$$A\vec{X} = \vec{V}$$

寻找一个 $\vec{X}$ 经过矩阵 A 线性变换得到 $\vec{V}$

1. 考虑 $det(A) \neq 0$ 的常见情况：
我们需要在空间内进行一个逆向的变换，求解 $\vec{X}$
向量进行变换和逆变换之后，保持不变

$$(AA^{-1}) = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \tag{11}$$

所以：

$$A^{-1}A\vec{X} = A^{-1}\vec{V} \Rightarrow \vec{X} = A^{-1}\vec{V} \tag{12}$$

6.2 秩

秩: 等价于变换后空间的维度 (线性无关的列向量的个数)

变换后向量落在一条直线上，$\mathrm{Rank}=1$

变换后向量落在一个平面内，$\mathrm{Rank}=2$

6.3 列空间

列空间即是矩阵列张成的空间，矩阵所有可能的输出的向量的集合

7 点积和对偶性

7.1 点积

:

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} \cdot \begin{bmatrix} x_j \\ y_j \end{bmatrix} = x_i \times x_j + y_i \times y_j \tag{13}$$

使用向量进行表示, 如图 11：表示为一个向量在另一个向量上的投影值和被投影向量长度的乘积。

,

$$\vec{V} \cdot \vec{W} = |\vec{V}`| \cdot |\vec{W}| = |\vec{V}| \cdot |\vec{W}| \cdot sin\theta \tag{14}$$

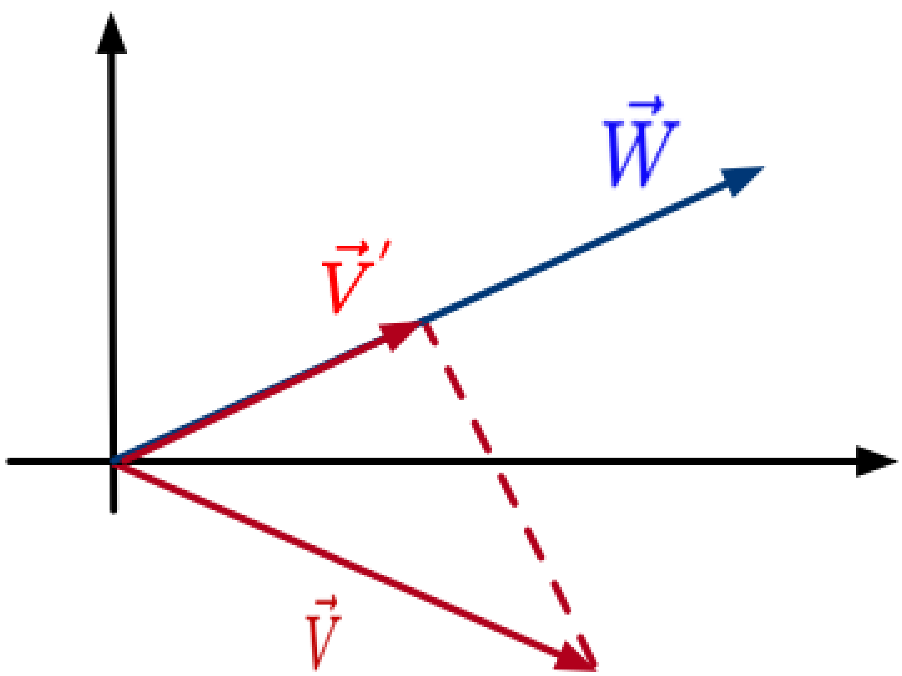


图 11: 点积计算

为什么点积运算和向量投影有关系呢？

7.2 线性变换：数轴

我们关注多维空间到一维空间 (数轴) 的线性变换。

线性变换矩阵关注变换之后的 $\vec{i}, \vec{j}$

数轴只有一个数进行表示，所以变换矩阵表示为 $\begin{bmatrix} 1 & 2 \end{bmatrix}$

给定一个变换矩阵，对向量进行作用

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

这个计算类似于上面介绍的点积计算公式 13.

7.3 投影矩阵

如图12 所示，在二维空间人为添加一条过原点的数轴，其中，$\vec{u}$ 是一个二维向量，恰好落在数轴上。

假设我们想将二维的向量投影到数轴上，我们需要定义一个二维到一维的一个线性变换。

线性变换需要一个矩阵，我们称为投影矩阵，矩阵由变换后的 i, j 决定，即

$$\begin{bmatrix} i' & j' \end{bmatrix}$$

变换后的 i, j 我们利用对称性进行推理:

$\vec{u}, \vec{i}$ 都是单位向量，存在一个对称轴，辅助观察互相投影是对称的。如图13 所示，所以 $\vec{u}$ 的投影是 u_x。所以变换矩阵是:

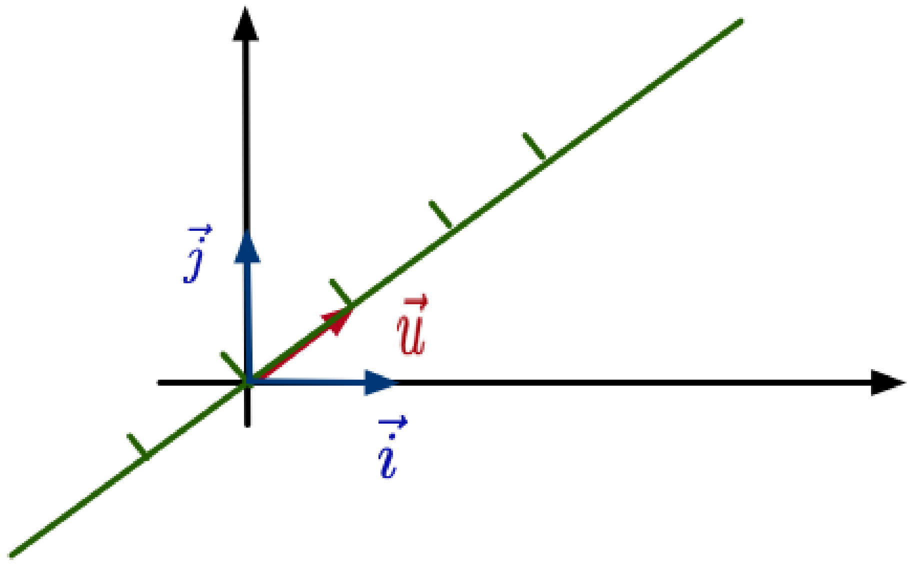


图 12: 二维向量投影到数轴

$$\begin{bmatrix} u_x & u_y \end{bmatrix}$$

对于空间中任意向量变换的结果：

1. 使用矩阵向量表示投影：

$$\begin{bmatrix} u_x & u_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

2. 使用点积表示，（和 $\vec{u}$ 的坐标进行点积）

$$\begin{bmatrix} u_x \\ u_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

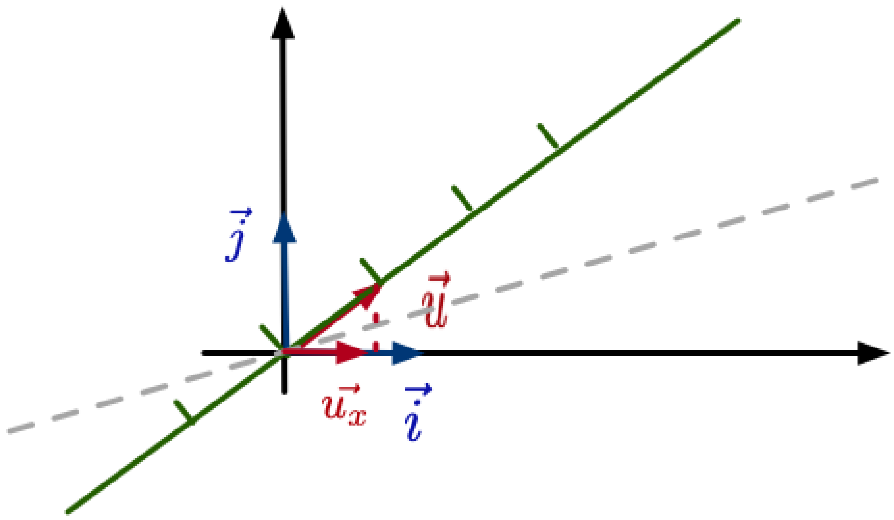


图 13: 对称性推理投影

上述二者等价，看到投影（投影变换矩阵）和点积的联系

我们可以推出一个小结论：

二维到一维的线性变换矩阵和二维空间中对应的向量相关。应用这个变换矩阵等价于和其相应的向量进行点积

7.4 对偶性

对偶性：自然而然而又出乎意料的对应关系。

一个多维到一维的线性变换的对偶是：多维空间中某个特定的向量。

8 叉积和线性变换

8.1 叉积

两个向量的叉积数值上等于向量围成的平行四边形面积, 如图 14。

$$|\vec{v} \times \vec{w}| = S \tag{15}$$

联想到我们之前学习的行列式:

$$|\vec{v} \times \vec{w}| = det(\begin{bmatrix} v_x & w_x \\ v_y & w_y \end{bmatrix}) \tag{16}$$

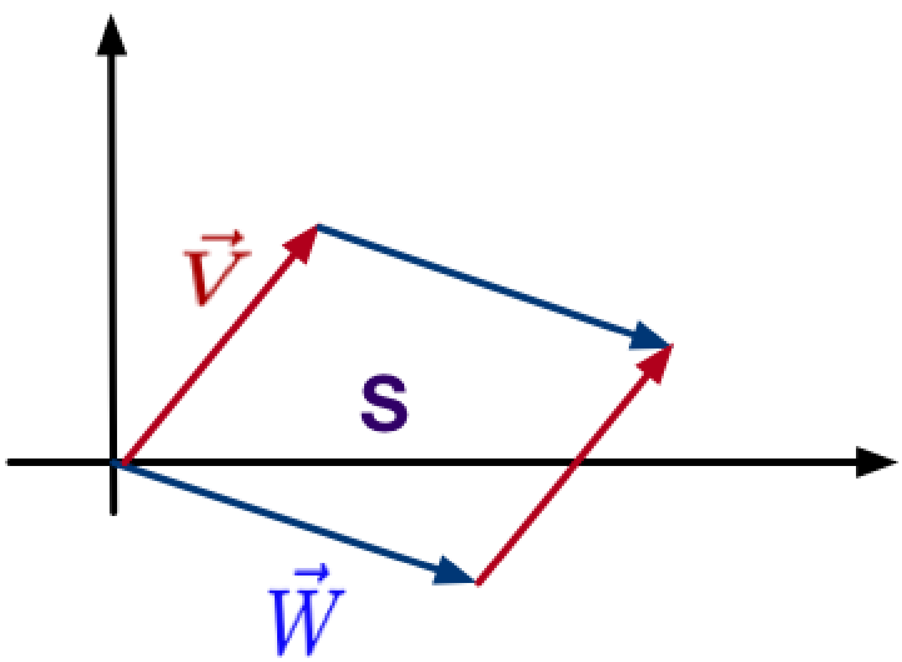


图 14: 叉积几何表示

但是叉积的结果不是一个数，而是一个向量。我们已经确定了向量的大小，接下来根据右手定则找到一个垂直于平面的方向。

8.2 计算方法理解叉积

定义一个从三维空间到一维的一个变换。

$$det(\begin{bmatrix} x & v_x & w_x \\ y & v_y & w_y \\ z & v_z & w_z \end{bmatrix}) = f(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) \tag{17}$$

上面公式意思是：给定三维空间的两个向量，$\vec{V}$, $\vec{w}$, 再随意给一个向量 $\begin{bmatrix} x & y & z \end{bmatrix}$（这里假设三个向量线性无关），我们可以通过矩阵的行列式计算三个向量组成的平行六面体的体积。即给定一个三维向量，输出一个一维的数值。

从三维到一维的变换，我们可以使用一个矩阵进行表示：

$$\begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = det(\begin{bmatrix} x & v_x & w_x \\ y & v_y & w_y \\ z & v_z & w_z \end{bmatrix}) \tag{18}$$

根据上一节 7, 矩阵变换可以看作点积运算

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = det(\begin{bmatrix} x & v_x & w_x \\ y & v_y & w_y \\ z & v_z & w_z \end{bmatrix}) \tag{19}$$

即向量 $\vec{p}$ 与任意向量的点积都等于矩阵的行列式：

$$p_1x + p_2y + p_3z = (v_yw_z - v_zw_y)x + (v_yw_x - v_xw_z)y + (v_xw_y - v_yw_x)z \tag{20}$$

假定 $\begin{bmatrix} x & y & z \end{bmatrix}$ 是 $\begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \end{bmatrix}$, 则:

$$\vec{P} = \begin{bmatrix} v_y w_z - v_z w_y & v_z w_x - v_x w_z & v_x w_y - v_y w_x \end{bmatrix} = det(\begin{bmatrix} x & v_x & w_x \\ y & v_y & w_y \\ z & v_z & w_z \end{bmatrix}) \tag{21}$$

8.3 几何方法理解叉积

向量 $\vec{P}$ 和 $\begin{bmatrix} x & y & z \end{bmatrix}$ 点积得到平行六面体的有向体积。

而平行六面体的体积等于向量 $\vec{V}$ 和向量 $\vec{W}$ 组成的平行四边形的面积乘向量 $\begin{bmatrix} x & y & z \end{bmatrix}$ 在垂直向量向量 $\vec{V}$ 和向量 $\vec{W}$ 方向上的分量 $\vec{S}$。

$$det(\begin{bmatrix} x & v_x & w_x \\ y & v_y & w_y \\ z & v_z & w_z \end{bmatrix}) = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} .\vec{S} \tag{22}$$

变换和点乘对偶向量等价, 对偶向量是一个方向和底面垂直 (右手), 大小等于 S 的向量。

9 基变换

9.1 基本想法

Space has no grid

前面我们都是默认使用一组最常见的基坐标构建坐标系，可是空间中是不存在坐标系的，我们可以任意构建。使用一组大小和方向不同的向量作为基向量，构建一个新的坐标系。

9.2 不同坐标系之间的转化

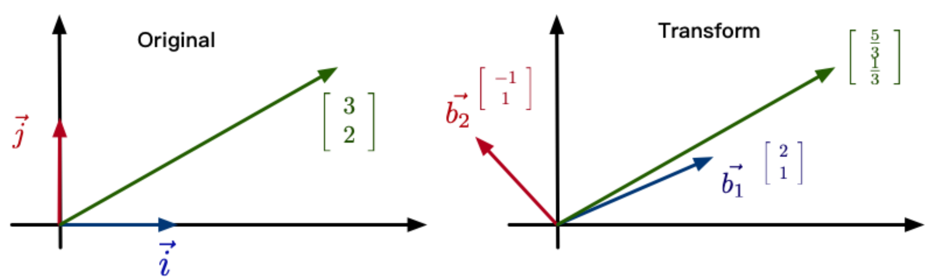


图 15: 基坐标变换

如图 15 ，我们下文将 original 叫做 O 系，将 Transform 叫做 T 系。

考虑在 T 中存在的一个向量 $\begin{bmatrix} -1 & 2 \end{bmatrix}$，*在 O 中是多少？*

$$-1\vec{b_1} + 2\vec{b_2} = -1\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

因为矩阵相乘等价于线性变换，线性变换关注基向量的坐标。所以将

坐标系 O 中变换后的基向量 ($\vec{b_1}, \vec{b_2}$) 变为变换矩阵的列向量。

1. 从几何上来说，变换矩阵将原始的 grid 变为新的（transform）grid

2. 从数值上来说，将 transform 中的向量描述转换为 Original 中的向量描述

9.3 逆运算

给定在 O 系中的坐标系的向量，计算在 T 系中应该是什么?

之前我们是从 T 系到 O 系:

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$$

现在我们要从 O 系到 T 系, 对上面的变换进行逆向执行:

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1}$$

计算得到:

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

给定一个 O 系中的向量，计算在 T 中坐标

$$\frac{1}{3}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \end{bmatrix}$$

总结:

A 是一个变换矩阵，矩阵的列向量是 T 坐标系中的基向量在 O 系中的表示。

$$A \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \end{bmatrix} \tag{23}$$

$$A^{-1} \begin{bmatrix} x_o \\ y_o \end{bmatrix} = \begin{bmatrix} x_t \\ y_t \end{bmatrix} \tag{24}$$

9.4 空间线性变换在不同坐标系中表示

例如：给定一个 90 度旋转矩阵

在 O 系中是：

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

那么这个在 T 系中的表示呢？

1. 给定 T 系中一个向量 $\vec{V}$: $\begin{bmatrix} -1 & 2 \end{bmatrix}$, 应用基变换矩阵：

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

得到向量 $\vec{V}$ 在 O 系中的表示

2. 之后应用在 O 系中的线性变换矩阵（90 度旋转矩阵）：

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

得到在 O 系中进行旋转操作之后的向量表示

3. 应用逆变换，将变换后的向量在坐标系 T 中进行表示。

$$\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} \tag{25}$$

对 T 中任意一个向量 $\vec{V}$ 进行上述操作：

基变换 -> 线性变换 -> 基变换的逆

$$A^{-1}MA$$

M 是 O 中的线性变换，$A^{-1}MA$ 则是 T 中的线性变换.

$A^{-1}MA$ 表示的是一种转移作用，可以理解为：

M 表示的是我们所见变换，

$A^{-1}MA$ 则是切换一种视角来看待这个变换 (变换还是那个变换，只是视角变了)

10 特征值和特征向量

10.1 基本概念

1. **特征向量**

有些向量在变换前后留在他们所张成的空间中，如图 16 中的绿色的虚线。这样的向量称为特征向量。

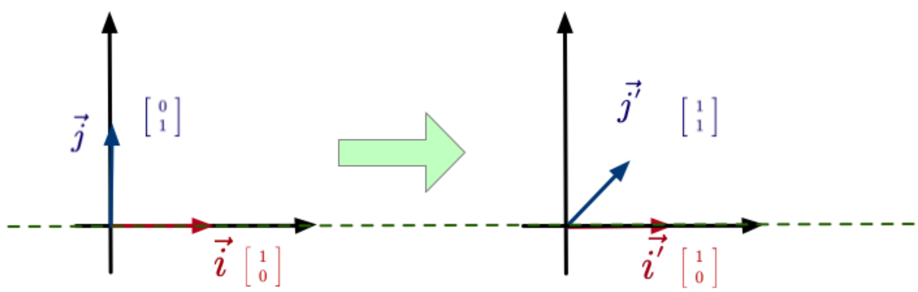


图 16: 特征向量

2. **特征值**

特征向量变换前后变化的倍数称为特征值，是一种衡量向量伸展或压缩比例的因子。

10.2 求解特征向量

$$A\vec{V} = \lambda\vec{V}$$

其中 $A\vec{V}$ 是变换矩阵，$\vec{V}$ 是特征向量，λ 是特征值。

λ 对于向量的变换等价于对向量的所有方向上乘 λ 。所以:

$$A\vec{V} = (\lambda I)\vec{V}$$

其中 I 是单位阵。

$$(A - \lambda I)\vec{V} = 0$$

矩阵变换将空间压缩到了更低的维度，所以：

$$det(A - \lambda I) = 0$$

利用这个等式进行求解。

10.3 特征基

基向量都是特征向量，表示为对角阵：

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

对角线上是特征值，其余为 0 值。

对角阵的乘法计算很容易：

$$\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}\cdots\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix} \tag{26}$$

假设有两个特征向量，但是不是基向量，我们可以使用基变换将其变为特征基向量。

例如矩阵

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

存在两个特征向量：

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

能够得到一个基变换矩阵

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

进行相应的变换，得到特征基：特征向量为基向量，一定是一个对角阵。

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

11 抽象向量空间

11.1 函数抽象成向量

根据前面线性的定义:

$$L(\vec{V}+\vec{W})=L(\vec{V})+L(\vec{W}) \quad L(c\vec{V})=cL(\vec{V})$$

函数的微分算子也是一个线性变换

$$\frac{d}{dx}f(x)$$

11.2 矩阵描述求导

选取全体多项式作为空间。选取 $1, x, x^2, \cdots, x^n$ 作为基函数。

将函数表示为基向量:

$$1x^2+3x+5=\begin{bmatrix}5\\3\\1\\0\\0\\\cdots\\0\end{bmatrix}$$

表 1

线性代数	函数
线性变换	线性算子
点积	内积
特征向量	特征函数

对应的求导矩阵为:

$$\frac{d}{dx} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \tag{27}$$

其中矩阵每一个列向量为相应的基函数进行求导得到的。

$$\frac{d}{dx}(1x^2 + 3x + 5) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 3 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 1 \\ \cdots \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \\ \cdots \end{bmatrix}$$

所以 **求导和矩阵乘法都是线性变换**

11.3 抽象向量空间

只要满足向量的相应规则，如图 17 （图片来自维基百科），就可以视为向量 (就带你玩)。

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

图 17: 向量公理化

就好像数字“3 ”可以各种三个东西，是一种抽象的集合。

12 克莱姆法则

12.1 求解线性方程组

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

可以理解为：对于一个未知向量 $\vec{V}$: $\begin{bmatrix} x & y \end{bmatrix}$，经过变换之后得到 $\vec{V}$: $\begin{bmatrix} 4 & 2 \end{bmatrix}$

根据之前我们学过的行列式，我们知道，变换前后，区域面积等比例的缩放

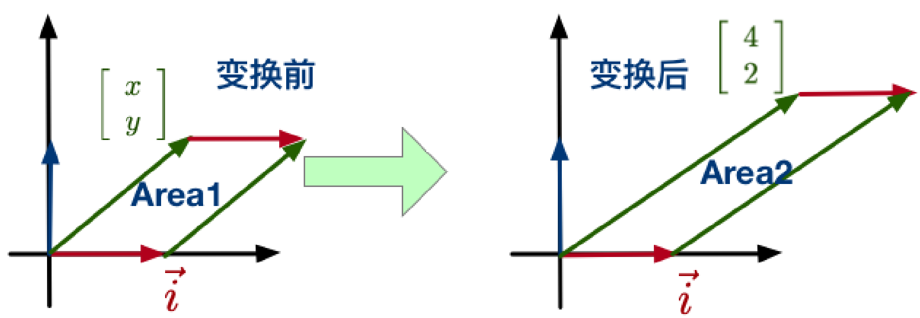


图 18: 变换前后区域面积变化

1. 变换前，如图 18所示：

$$Area_1 = det(\begin{bmatrix} 1 & x \\ 0 & y \end{bmatrix}) = y \tag{28}$$

2. 变换后，基向量 $\vec{i}$ 变为 $\begin{bmatrix} 2 & 0 \end{bmatrix}$, 所以变换后的区域面积：

$$Area_2 = det(\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix}) \tag{29}$$

3. **面积的缩放等于行列式的值**

$$\frac{Area_2}{Area_1} = \frac{Area_2}{y} = det(\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}) \tag{30}$$

4. 所以求得:

$$y = \frac{det(\begin{bmatrix} 2 & 4 \\ 0 & 2 \end{bmatrix})}{det(\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix})} \tag{31}$$

同理 x 的求解

$$y = \frac{det(\begin{bmatrix} 4 & 1 \\ 2 & -1 \end{bmatrix})}{det(\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix})} \tag{32}$$

12.2 推广：n 维情形

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ n_1 & n_2 & \cdots & n_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} \tag{33}$$

$$\begin{bmatrix} x_1 & 0 & \cdots & 0 \\ x_2 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ x_n & 0 & \cdots & 1 \end{bmatrix} = x_1 = \frac{\begin{bmatrix} x_1 & a_2 & \cdots & a_n \\ x_2 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ x_n & n_2 & \cdots & n_n \end{bmatrix}}{\begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \\ \cdots & \cdots & \cdots & \cdots \\ n_1 & n_2 & \cdots & n_n \end{bmatrix}} \tag{34}$$

13 后记

文章内容浅显，整理急促，存在很多纰漏错误。所以未完不待续。。。