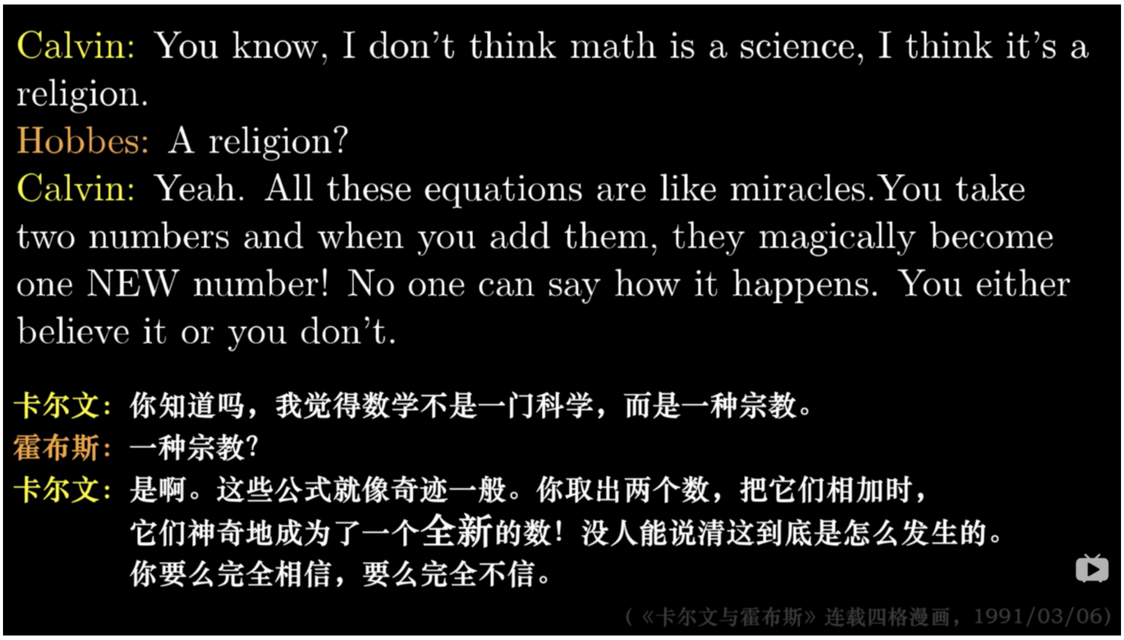
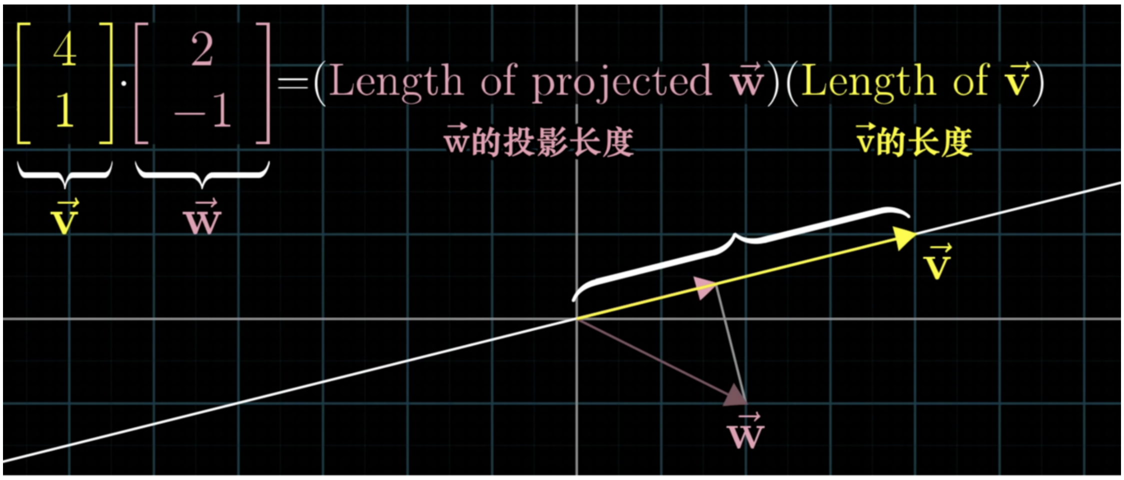
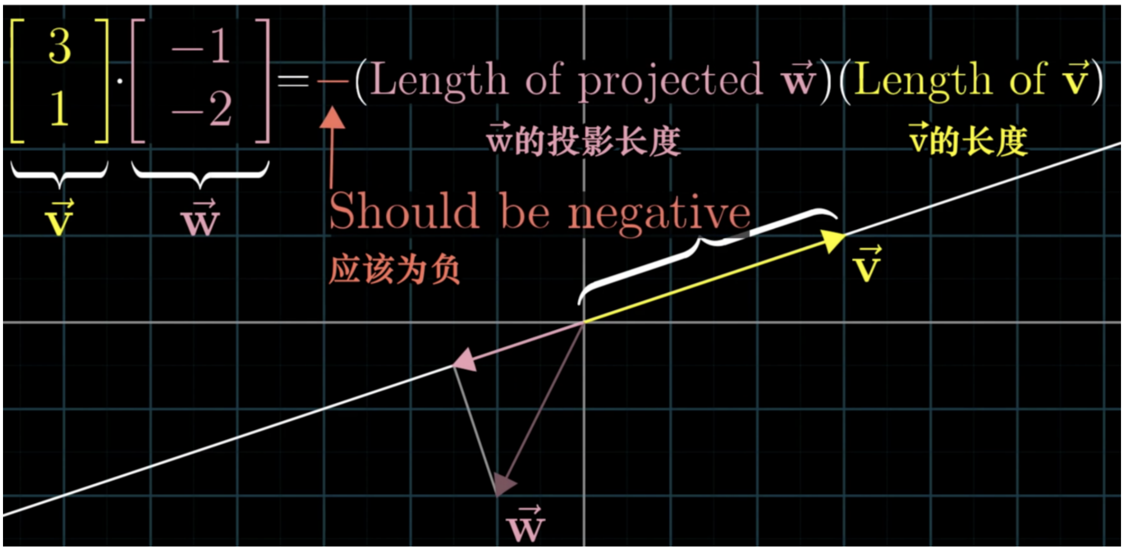
9.点积与对偶性



9.1点积的几何意义



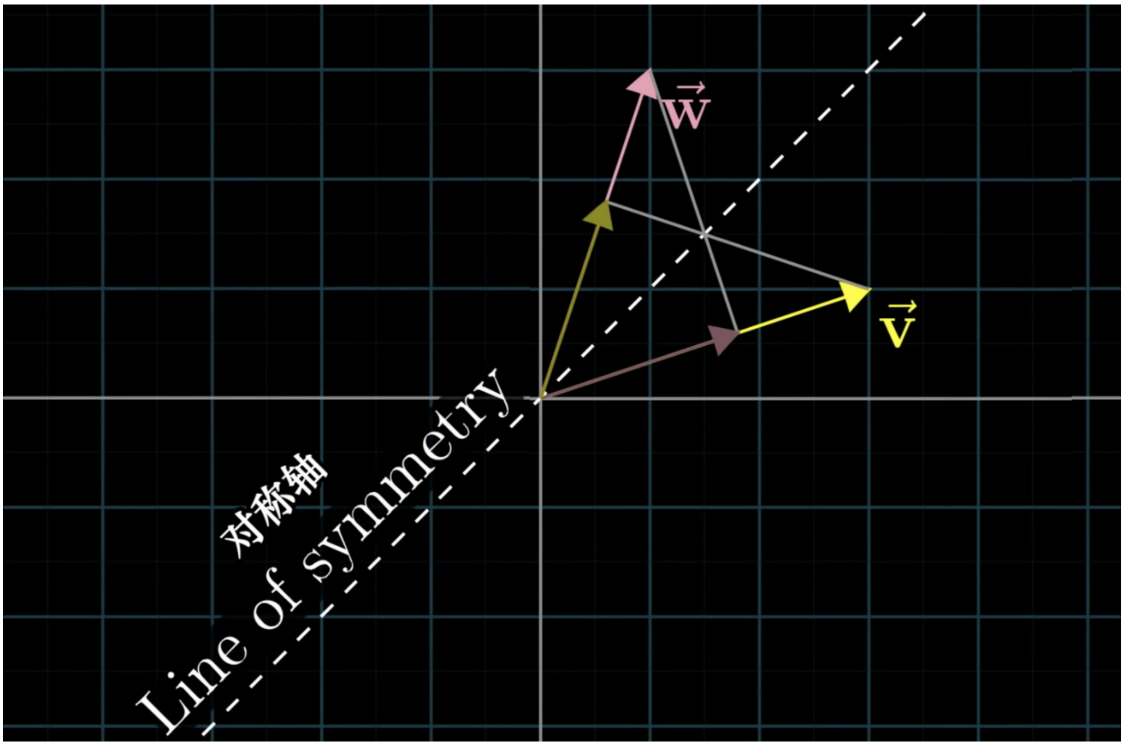


点积大于0表示它们的夹角小于90°（也就是指向相同）；点积为0表示向量互相垂直（一个向量再另一个向量上的投影为零向量）；点积小于0表示向量夹角大于90°（指向相反）

将点积看做是向量w在另一个向量v上的投影的长度乘以向量v的长度。这种解释异常不对称，它对两个向量的处理方式不同。所以，点积与顺序无关看起来很奇怪，可以将v投影到w上，也可以将w投影到v上，得到的长度与另外一个向量的长度相乘最后结果是一样的。

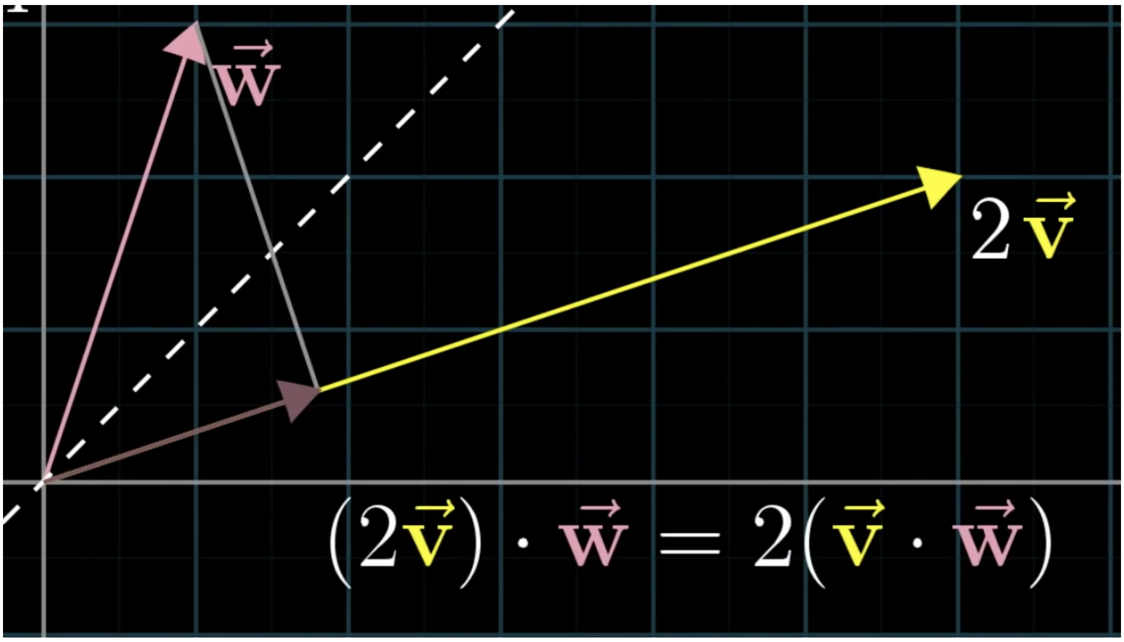
下面，我们从直观上理解点积为什么与顺序无关

如果v和w的长度相同，我们可以利用其中的对称性



w向v上投影，并将w的投影的长度与v的长度相乘，和v在w上投影，并将v的投影的长度与w的长度相乘互为镜像

如果将其中一个缩放若干倍，比如将v变成2倍，使得它们的长度不同，对称性就破坏了。但是我们可以这样解释：$2\vec{v}\cdot\vec{w}=2(\vec{v}\cdot\vec{w})$



这是因为，将v放大为原来的两倍并不会改变w的投影长度，只是v的长度变为了原

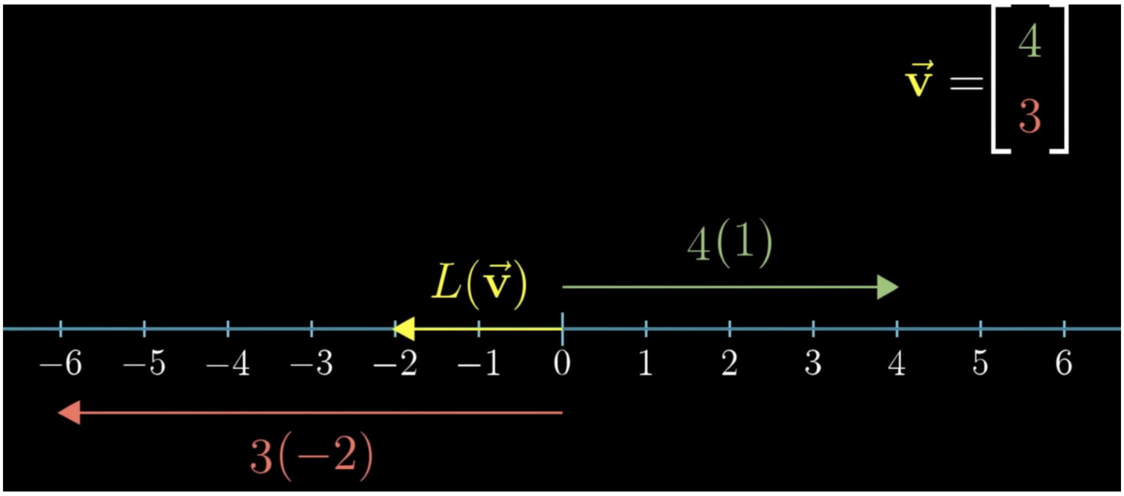
来的2倍；另一方面，假设将v投影到w上，这次是投影的长度变为原来的2倍，而向量w的长度不变，所以总体效果仍然是点击变为2倍。

所以，即使对称性被破坏了，在两种理解方式下，缩放向量对点积结果的影响是相同的。

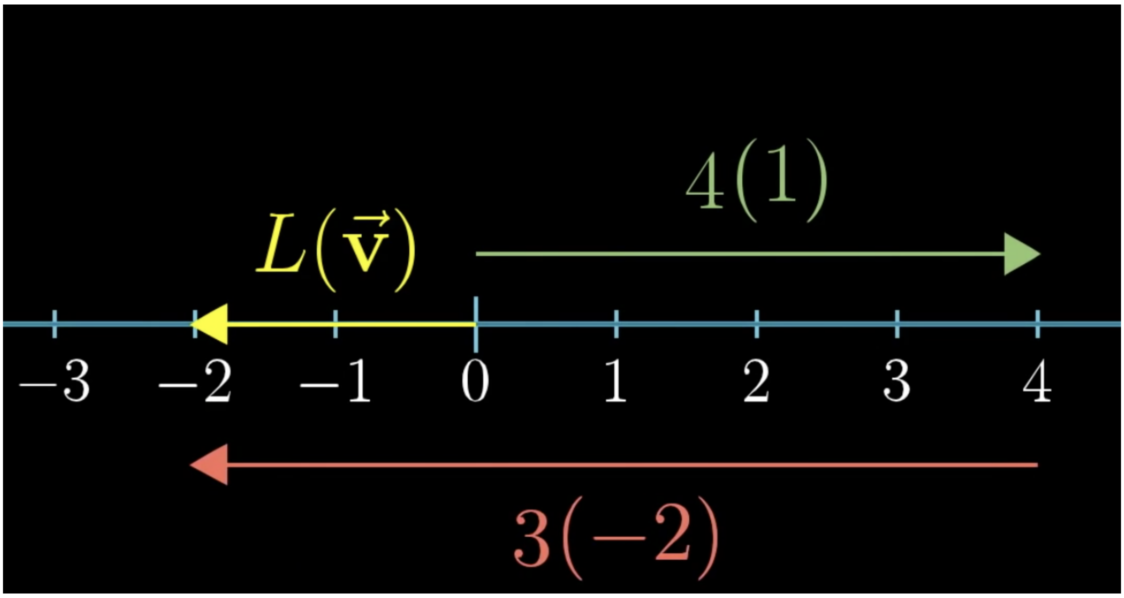
那么，还有一个疑问：点积的运算过程（对应坐标相乘并将结果相加）为什么会和投影有关系?

最令人满意的答案来自对偶性（duality）。

在继续讨论之前，我们先来看下多维空间到一维空间（数轴）的线性变换。假设你有一个线性变换，它将i帽和j帽变换至1和-2。要追踪一个向量。比如(4,3)，在变换后的位置，首先要对变换后的i帽和j帽进行相应的缩放



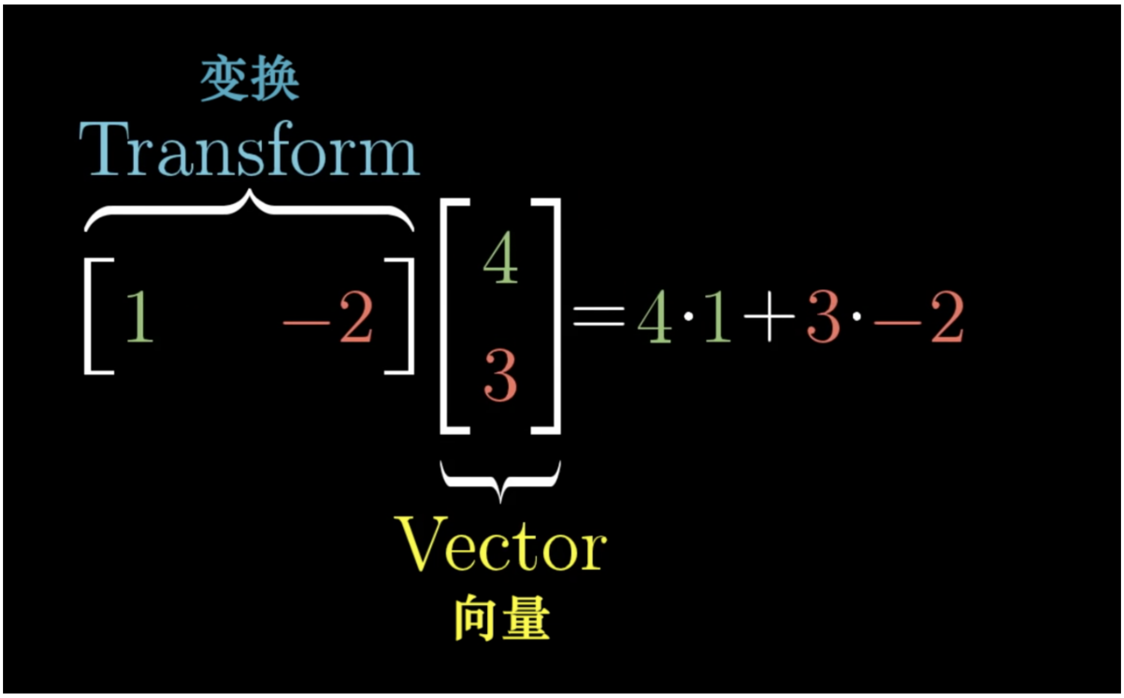
之后将它们相加：



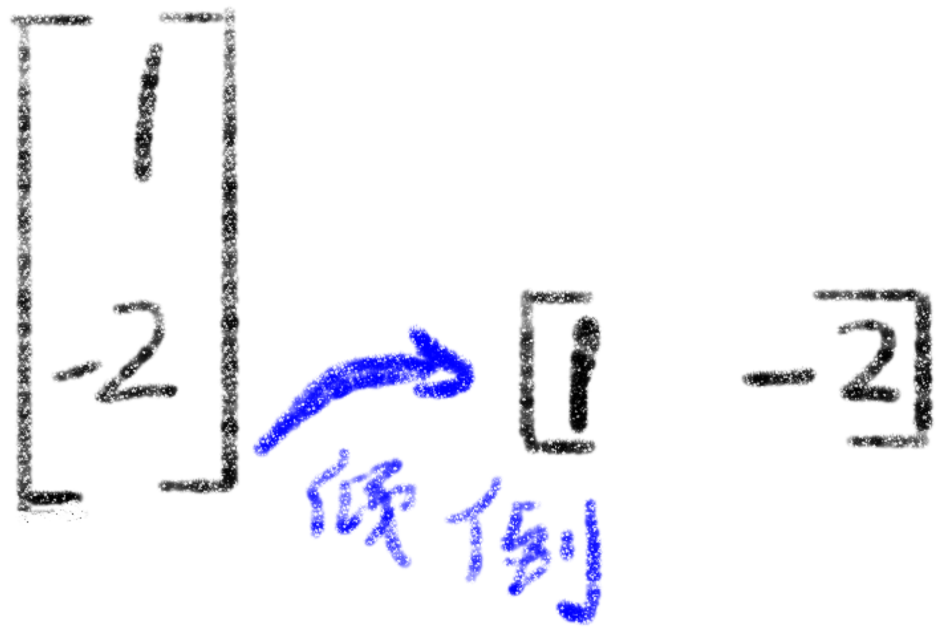
结果说明它落在-2上。

> 求长度就是将向量映射到一维空间上的过程

当我们从数值角度进行计算时，它就是矩阵向量乘法

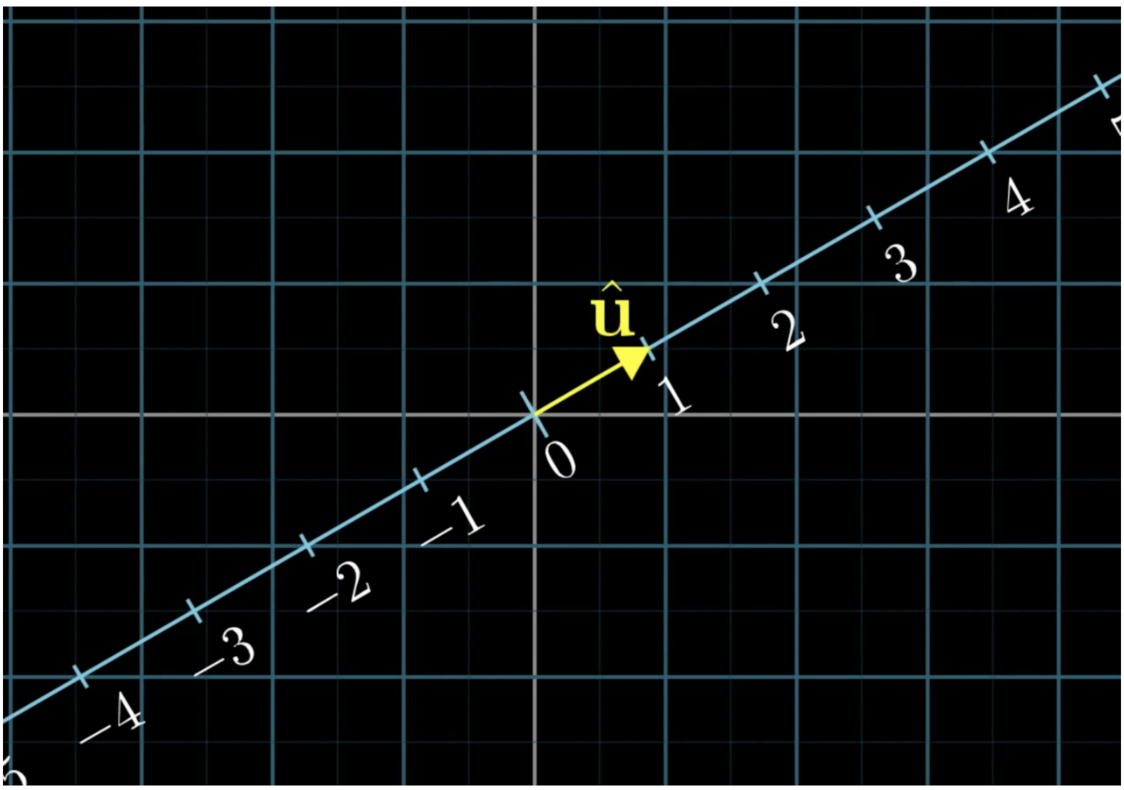


上图中，1×2矩阵的向量乘法的数值运算过程感觉上就和两个向量的点积一样。这个1×2矩阵像一个倾倒的向量。



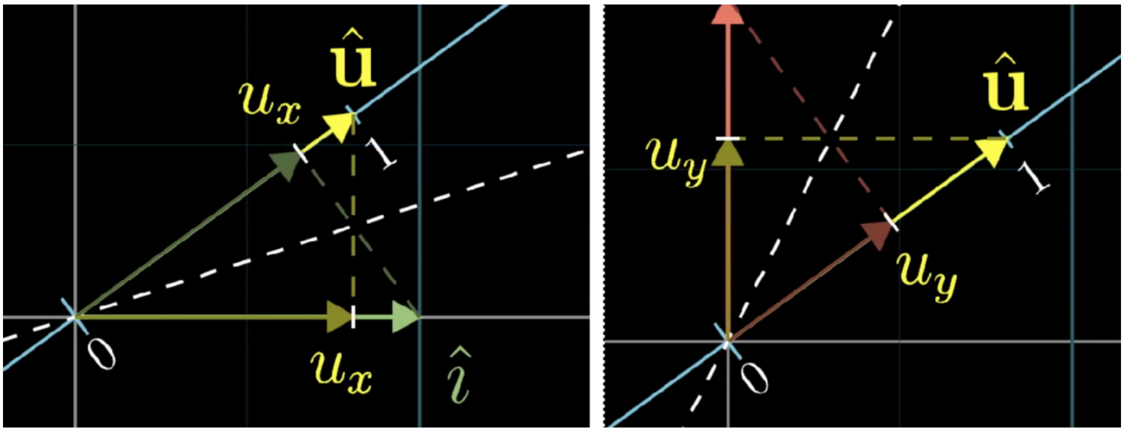
将“向量转换为数”的线性变换和这个向量本身有着某种关系。

我们现在将数轴复制一份，然后保持在原点，将它斜向放置在空间中，现在考虑这样一个二维向量，它的终点落在这条数轴的1上，给它起个名字叫u帽



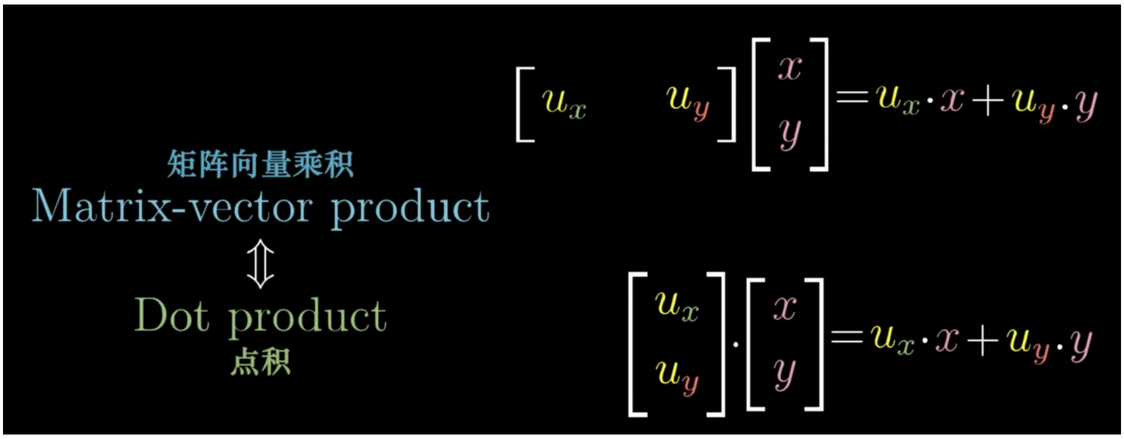
如果将向量直接投射到这条数轴上，实际上，我们就定义了一个从向量到数的函数。所以我们就能找到描述这个变换的1×2的投影矩阵。为了找到这个投影矩阵，我们把这条斜着的数轴放大看，并且需要考虑变换后i帽和j帽的位置，因为它们是矩阵的列。

我们可以通过对称性进行推理



所以变换后的1×2矩阵为$[u_x, u_y]$，即i帽和y帽变换后的坐标。其中u_x, u_y又是u帽在二维空间中的坐标

空间中任意坐标经过投影变换，也就是投影矩阵与这个向量相乘，它和这个向量与u帽的点积在计算上相同：



这就是为什么单位向量的点积可以解读为将向量投影到单位向量所在直线上所得的的投影长度。对于非单位向量，就相当于对这个单位向量进行了等比例的缩放。

思考一下这个过程：

1. 有个从二维向量到数轴的线性变换
2. 它不是由向量数值或点积运算得到的，而只是通过将空间投影到给定数值上定义。
3. 因为这个变换是线性的，可以用某个1×2的矩阵来描述

4. 1×2矩阵与向量的相乘与转置矩阵并求点积的计算过程相同，所以这个投影变换必然会与某个二维向量相关

启发： 任何时候看到一个线性变换，它的输出空间是一个一维数轴，无论它是如何定义的，空间中会存在唯一的向量v与之相关。