线性代数

1. 向量

物理　　统计学　　计算机

有方向的量　　空间中的点　　数表

数学里，我们可以将以上三种向量的概念合并到一起。
只要我们默认向量的起点在原点。如下页所示。

向量一般用小写字母表示，如 α, β， u, v。

（注：向量的准确定义是：向量空间中的元素。这比我们此章要研究的向量含义更广。）

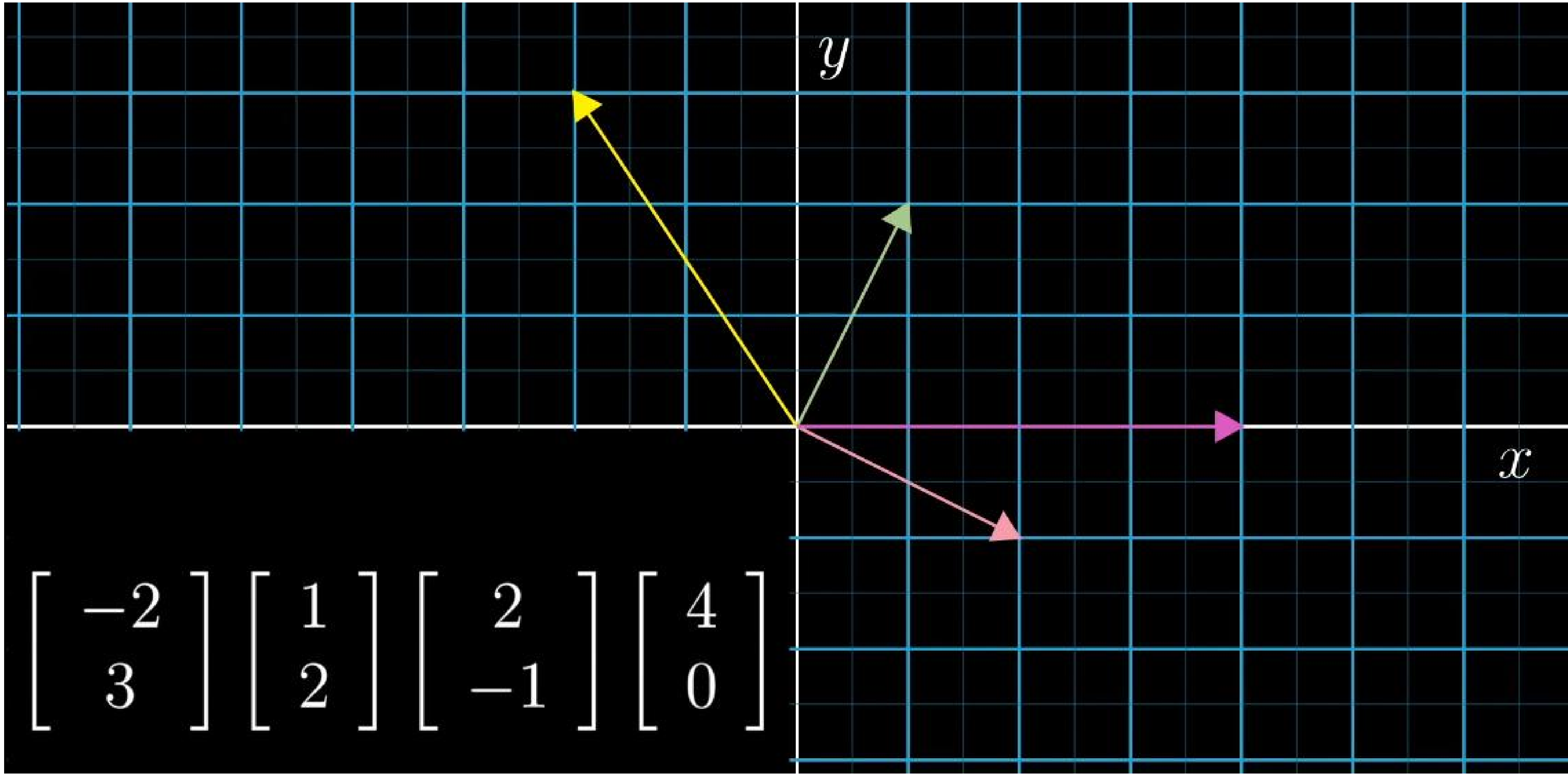
1. 向量

向量的坐标

数表中，第一个数（标量）告诉我们从原点沿X轴的方向如何移动。

第二个数告诉我们从原点沿Y轴如何移动。

此为2维空间的向量的坐标，本PPT之后都主要以2维向量举例。



CDA数据分析师 CERTIFIED DATA ANALYST

向量运算

向量的加法：2维空间内，就是求给定2个向量所围成的平行四边形的对角线。

$$\binom{a}{b} + \binom{c}{d} = \binom{a+c}{b+d}$$

向量的数乘：将给定向量按比例缩放（拉伸），负数表示反向拉伸。

$$\mathrm{a}\binom{x}{y} = \binom{ax}{ay}$$

2. 线性组合与向量空间

线性组合

将一个向量组中的向量做数乘后相加，即得到该向量组的一个所谓的线性组合。定义如下：

空间V中的一组向量$v_1, \ldots, v_n$的线性组合是指形如

$$a_1 v_1 + \ldots + a_n v_n$$

的向量，其中$a_1, \ldots, a_n$为常数。

例：
在R^2空间中，(6,1)是(2,3), (2,-1)的线性组合，因为

$$\binom{2}{3} + 2\binom{2}{-1} = \binom{6}{1}$$

向量空间

向量空间就是对加法和数乘封闭的集合

2个常用的特殊向量：

单位向量：数轴上的一单位长度的向量，例如$\binom{1}{0}$。

零向量：原点，即向量的所有元素都为零。

在线性空间中，所有其他向量都是以上两种向量的线性组合。
或者说，用单位向量和零向量，可以通过加法和数乘组合出任意一个其他向量。

例：

$$\binom{3}{-2} = 3i + (-2)j = 3\binom{1}{0} + (-2)\binom{0}{1}$$

忽视一些数学细节的话，我们可以简单地认为：每个单位向量都是坐标系的一个基向量。所有基向量的集合叫一组基，基组定义了该向量空间。

（注：基也可以不由数组型的向量组成，例如 $1, x, x^2, \dots, x^{n-1}$ 也是一组基。）

张成空间

以二维空间为例：在线性组合中，给定两个非零向量。
一个向量固定，另一个向量自由变化，其线性组合可得到一条直线。

但是，令两个向量都自由变化，可以得到一个平面吗？

大多数情况下确实如此，除非两向量共线。

张成空间：所有可以表示为给定向量的线性组合的向量集合，被称为给定向量张成(span)的空间。

若给定多个向量，移除其中一部分而不减小张成空间，是为线性相关。
如果所有向量都给张成空间增加了维度，是为线性无关。

基与坐标

基的定义：向量空间的一组基是张成该空间的一个线性无关的向量组。

例如，在R^2空间中，最常用的一组基就是自然基（i,j），其中$i=\begin{pmatrix}1\\0\end{pmatrix}, j=\begin{pmatrix}0\\1\end{pmatrix}$

显然，R^2中所有向量都可以通过他们的线性组合线性表出，且这两个向量线性无关。

坐标的定义：向量α可以被基$\epsilon_1, \epsilon_2, \ldots, \epsilon_n$唯一地线性表出：

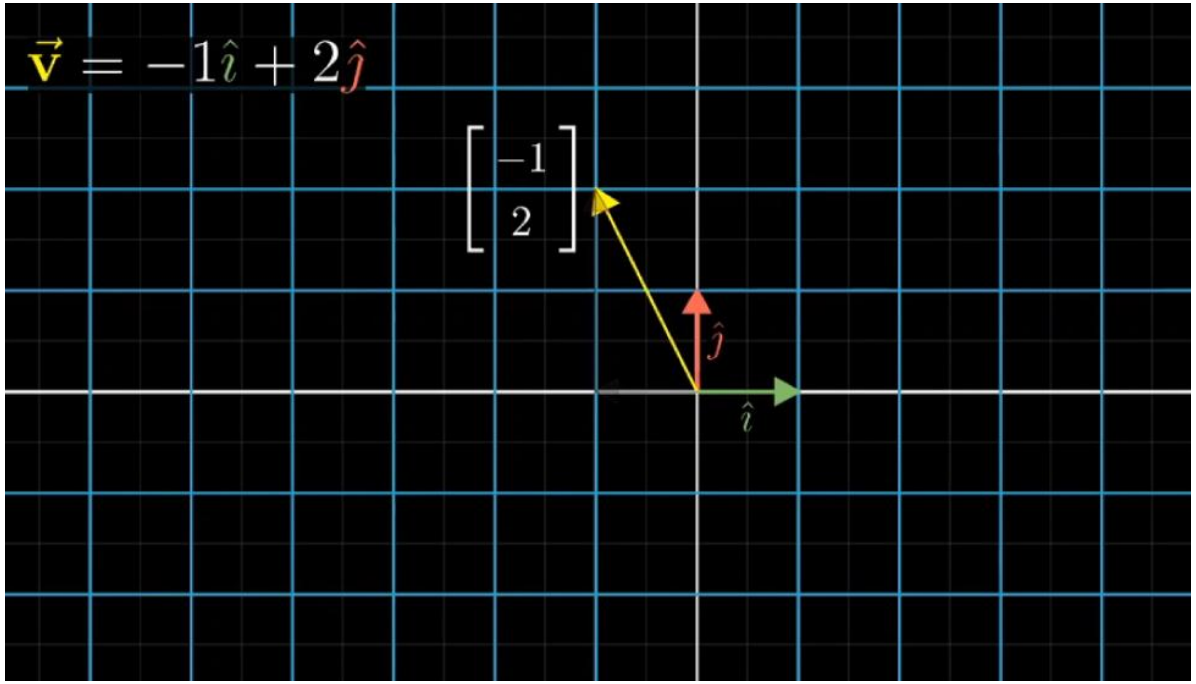
$$\alpha = a_1\epsilon_1 + a_2\epsilon_2 + \ldots + a_n\epsilon_n$$

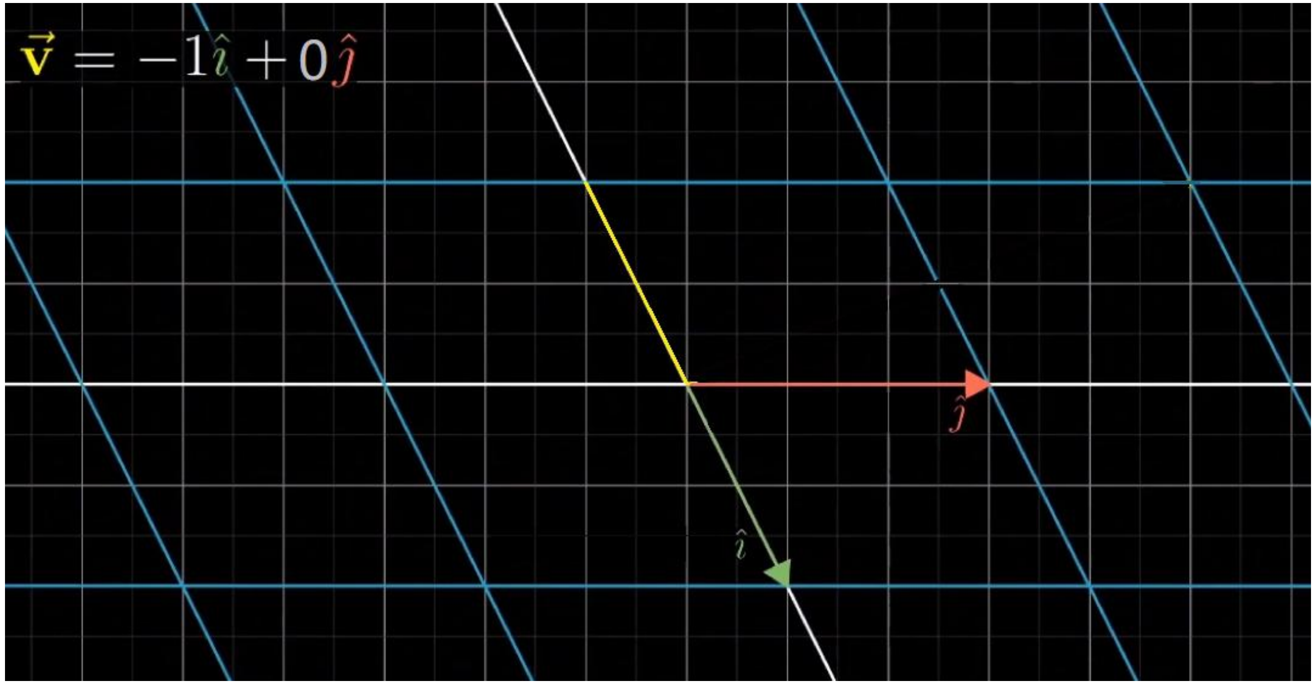
则系数$a_1, a_2, \ldots, a_n$称为向量α在基$\epsilon_1, \epsilon_2, \ldots, \epsilon_n$下的坐标，记为$(a_1, a_2, \ldots, a_n)$

一个向量 在不同的基下 坐标一般也不同。

例如，我们取向量$\binom{-1}{2}$，准确地说此向量在基$\binom{1}{0}, \binom{0}{1}$下的坐标是$\binom{-1}{2}$，因为

$\binom{-1}{2}$= -1 $\binom{1}{0} + 2\binom{0}{1}$。而如果把基换成$\binom{1}{-2}, \binom{3}{0}$，则坐标就变成了$\binom{-1}{0}$。


$\vec{v} = -1\hat{\imath} + 2\hat{\jmath}$
$\begin{bmatrix} -1 \\ 2 \end{bmatrix}$
$\hat{\jmath}$
$\hat{\imath}$


$\vec{v} = -1\hat{\imath} + 0\hat{\jmath}$
$\hat{\jmath}$
$\hat{\imath}$

3. 线性变换与矩阵

线性变换（也称线性映射）：
从空间V到V的线性变换是对加法和数乘封闭的函数 $T: V \to V$
也就是说，空间V中的任意一个元素，都可以通过变换T从V中找到另一个元素与之一一对应。

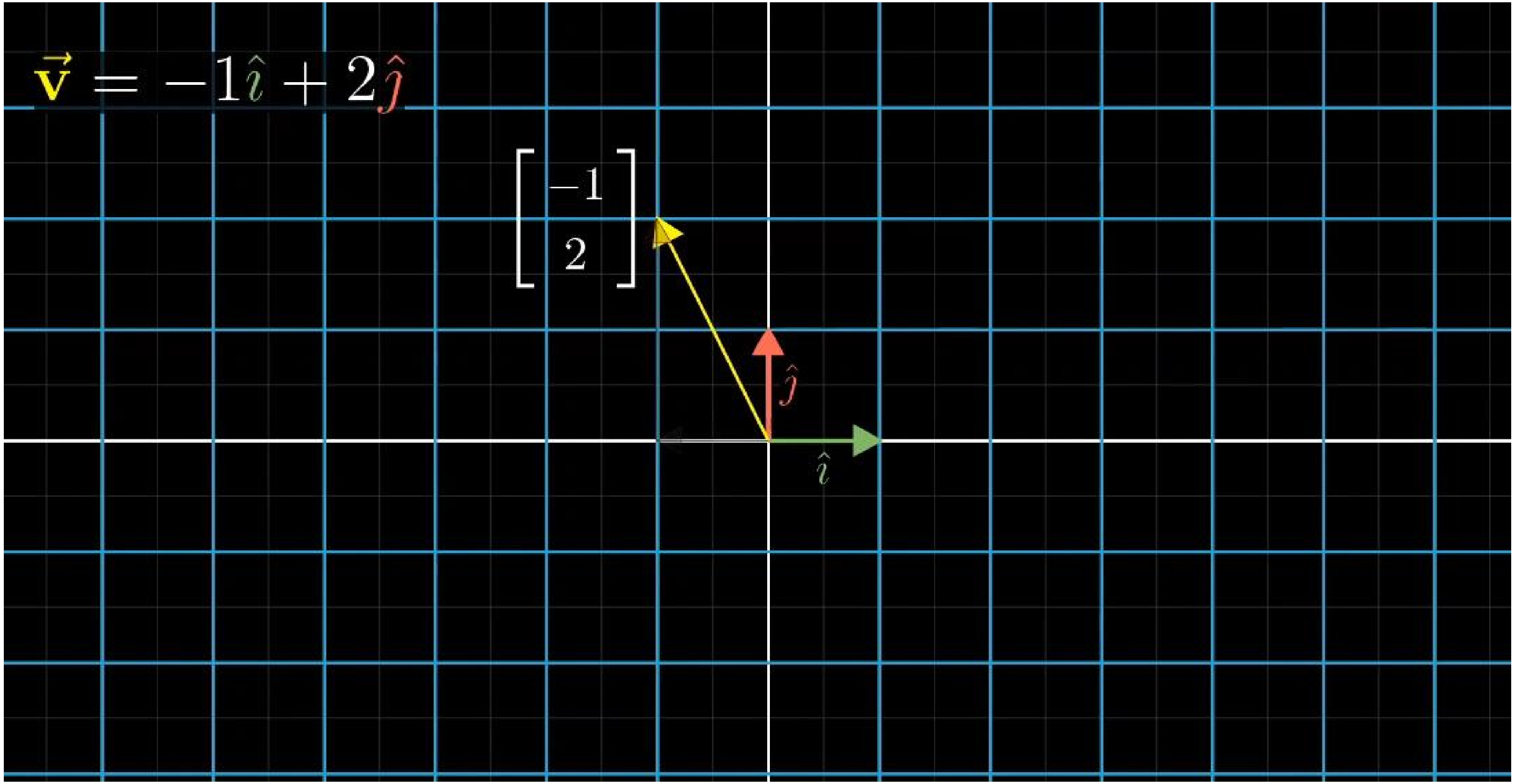
变换一个向量有两种方式：
1. 将该向量旋转拉伸。
2. 改变整个坐标系，这是操纵空间的手段。

线性空间中，改变空间只需改变基，也就是改变基向量的方向和长度。

线性变换通常用大写花体字母表示，例如 �，�

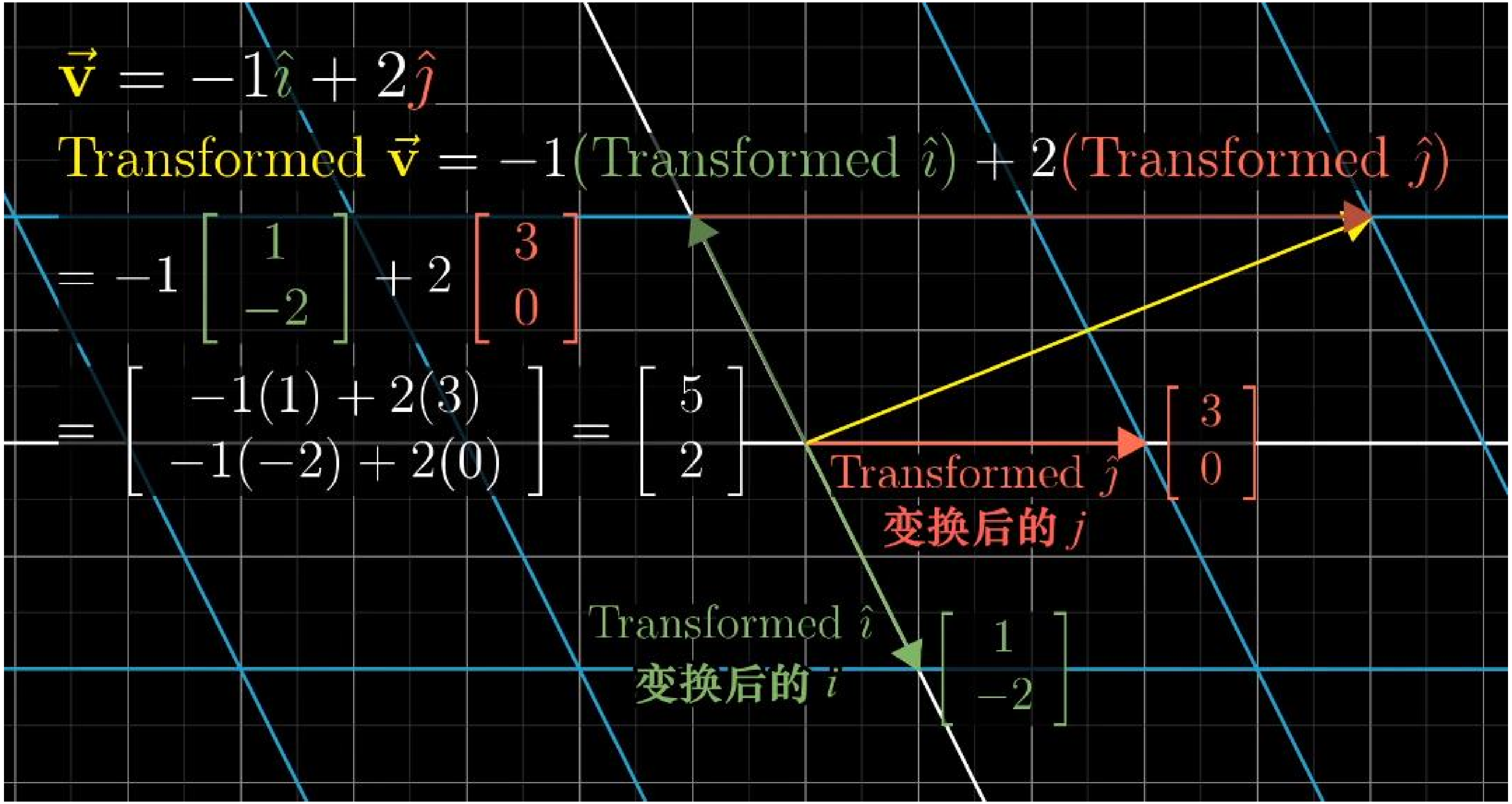
CDA数据分析师 CERTIFIED DATA ANALYST

若我们想将给定的黄色向量变换到$\binom{5}{2}$



变换的是基组。因此变换后，可通过新的基组求出新的向量。

换句话说，只要知道基组的新落脚点，我们就可以推断出任意向量的新落脚点。



上例中的两个基向量的新落脚点$\begin{pmatrix}1\\-2\end{pmatrix}$和$\begin{pmatrix}3\\0\end{pmatrix}$，可以拼成一个矩阵。

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

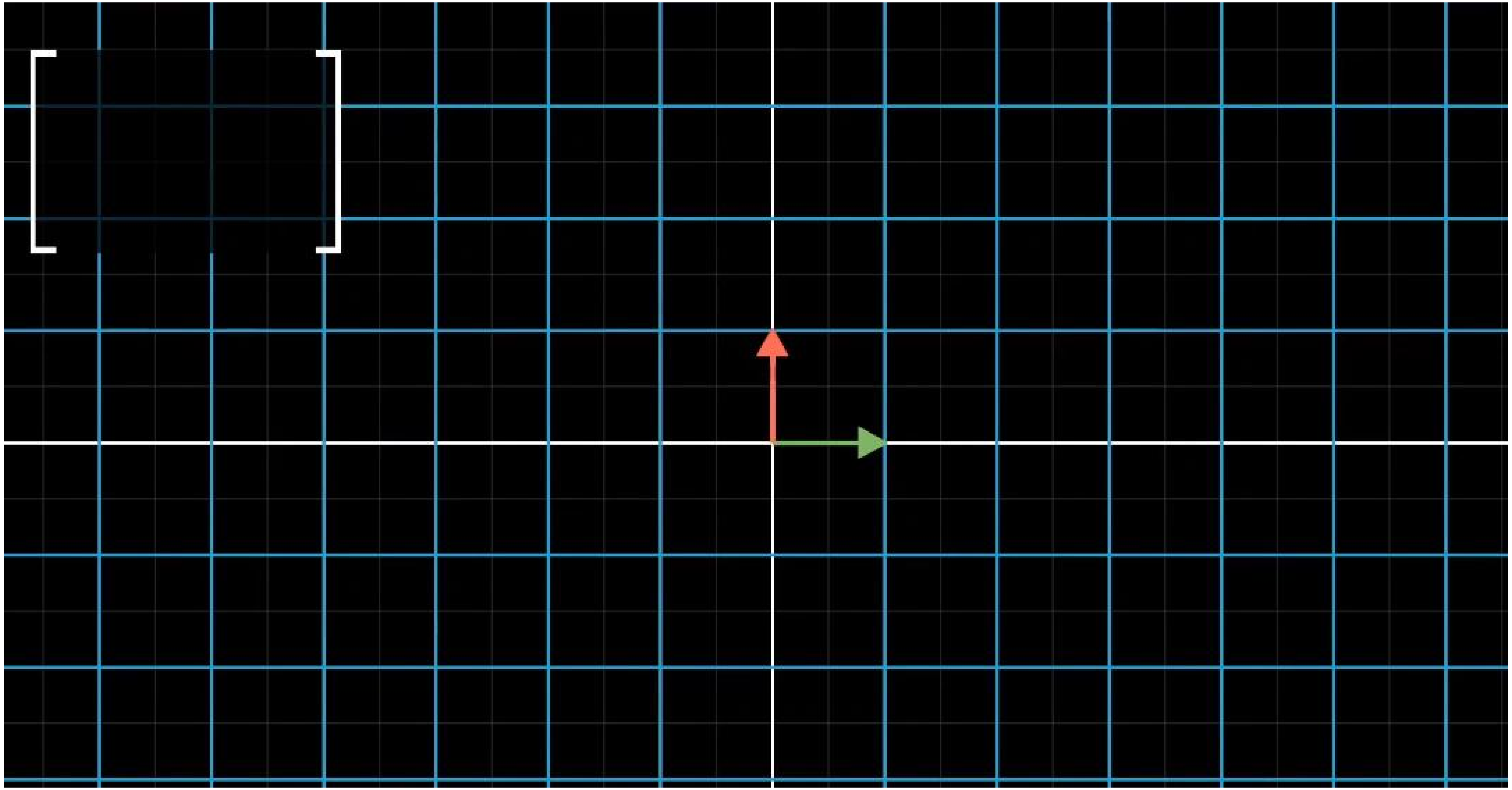
$$\begin{pmatrix}1 & 3\\-2 & 0\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=x\begin{pmatrix}1\\-2\end{pmatrix}+y\begin{pmatrix}3\\0\end{pmatrix}$$

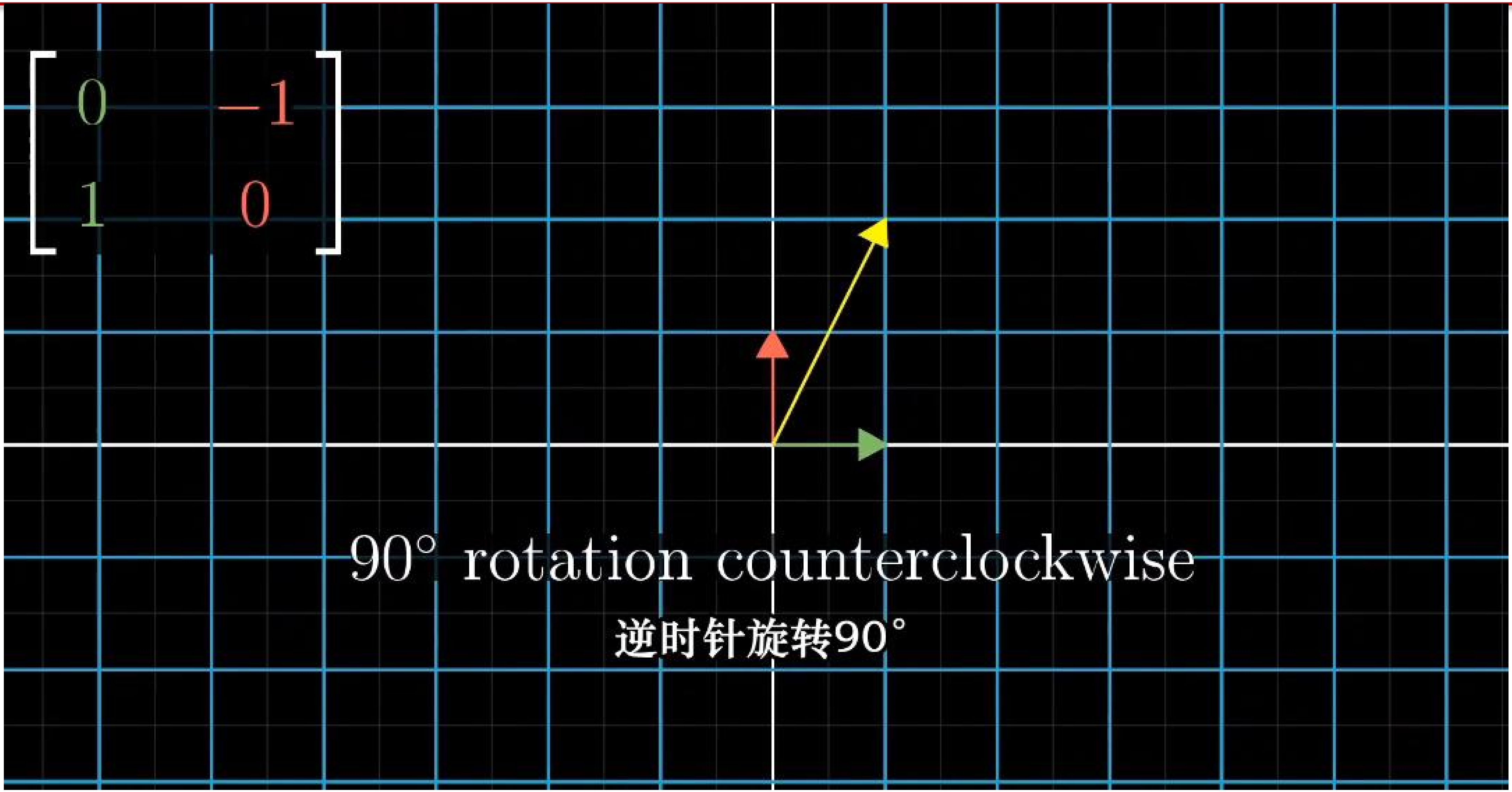
因此，矩阵代表一次变换。用一个向量左乘一个矩阵，就表示将这个向量按此矩阵所定义的变换映射到新的向量。

矩阵一般用大写英文字母表示，例如 A, B.

思考题1：

哪个矩阵可以将空间逆时针旋转90度？

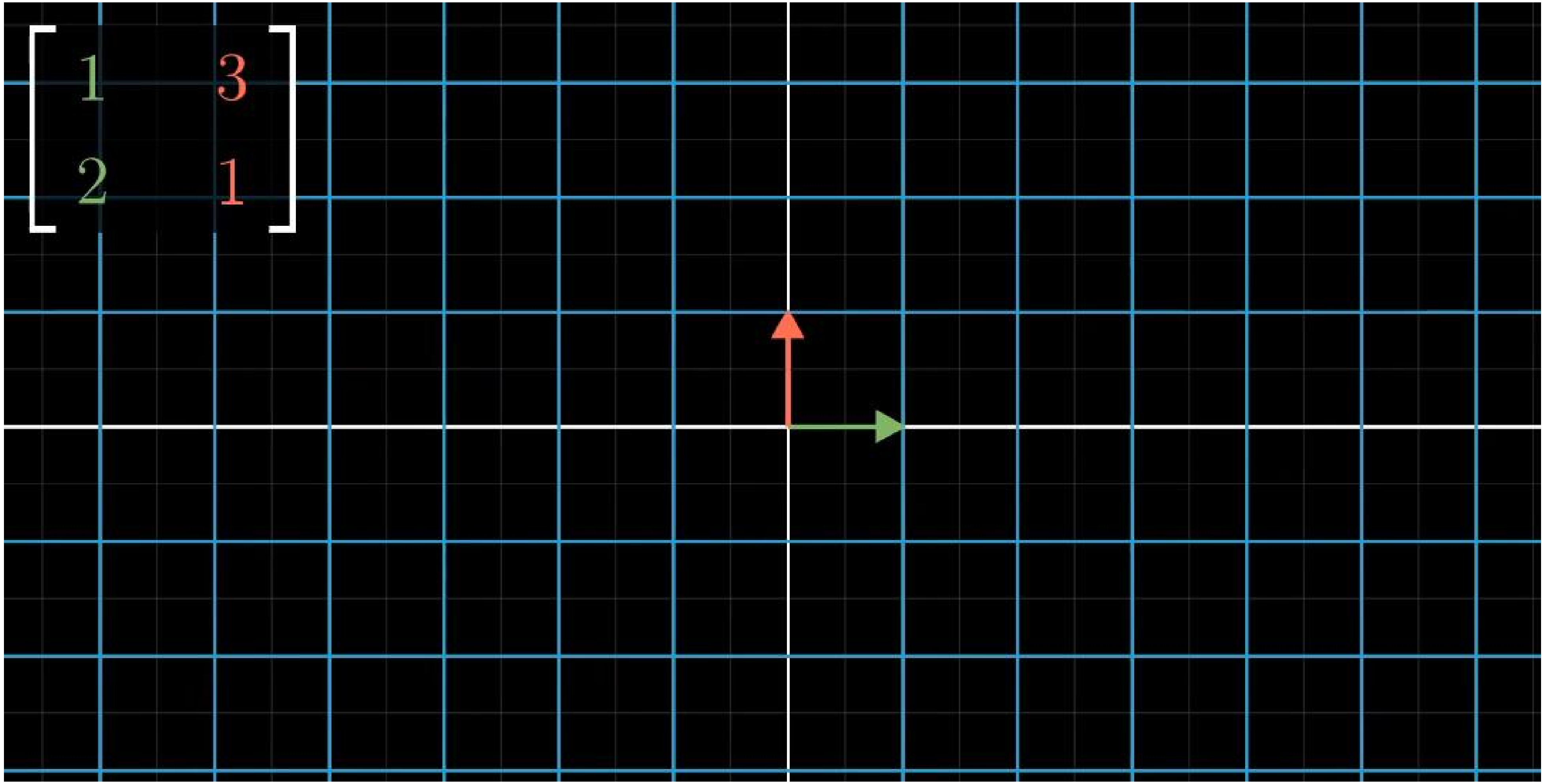




$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
90° rotation counterclockwise
逆时针旋转90°

思考题2：

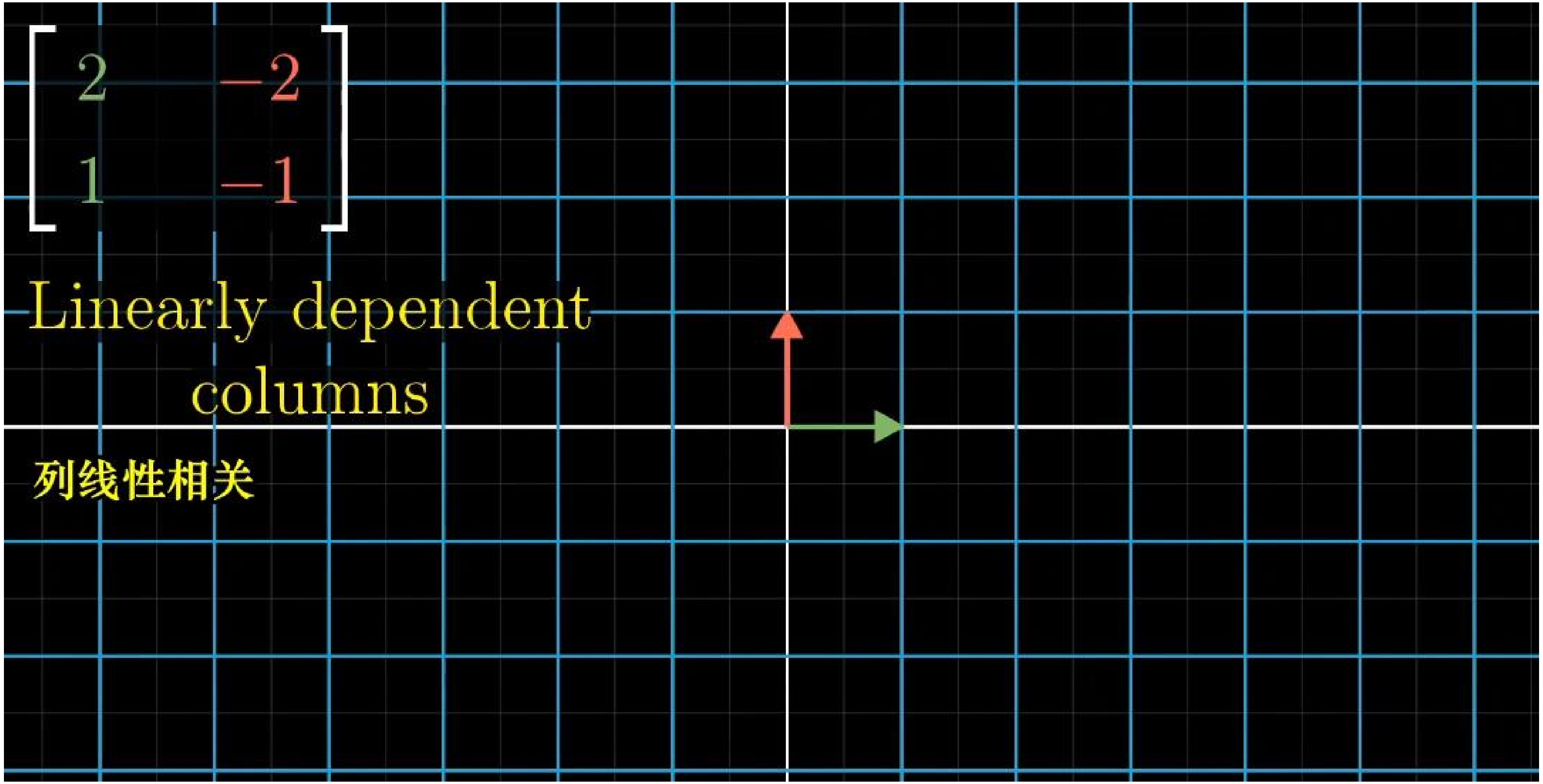
逆向思维。
给定一个矩阵，如何知道它代表什么线性变换？



还是找基组。
注意这道题的变换涉及到了空间的翻转。

思考题3：

如果矩阵的某些列向量线性相关，空间会如何变换？

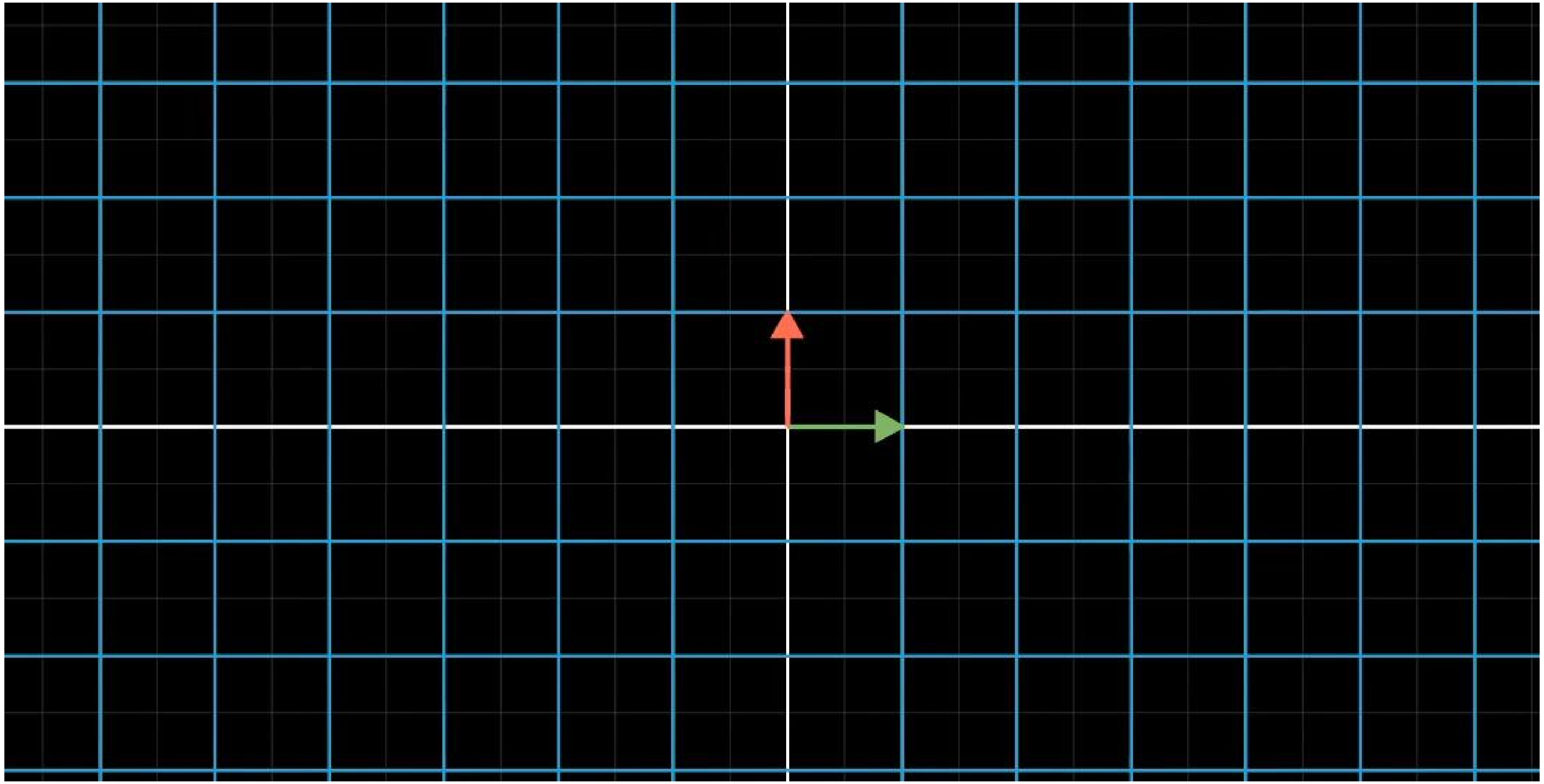


降维

4. 矩阵乘法

思考一种“复合变换”

先旋转（逆时针90度）
后剪切（斜置）



$$\underbrace{\begin{bmatrix}1&1\\0&1\end{bmatrix}}_{\substack{\text{Shear}\\\text{剪切矩阵}}}\left(\underbrace{\begin{bmatrix}0&-1\\1&0\end{bmatrix}}_{\substack{\text{Rotation}\\\text{旋转矩阵}}}\begin{bmatrix}x\\y\end{bmatrix}\right)=\underbrace{\begin{bmatrix}1&-1\\1&0\end{bmatrix}}_{\substack{\text{Composition}\\\text{复合矩阵}}}\begin{bmatrix}x\\y\end{bmatrix}$$

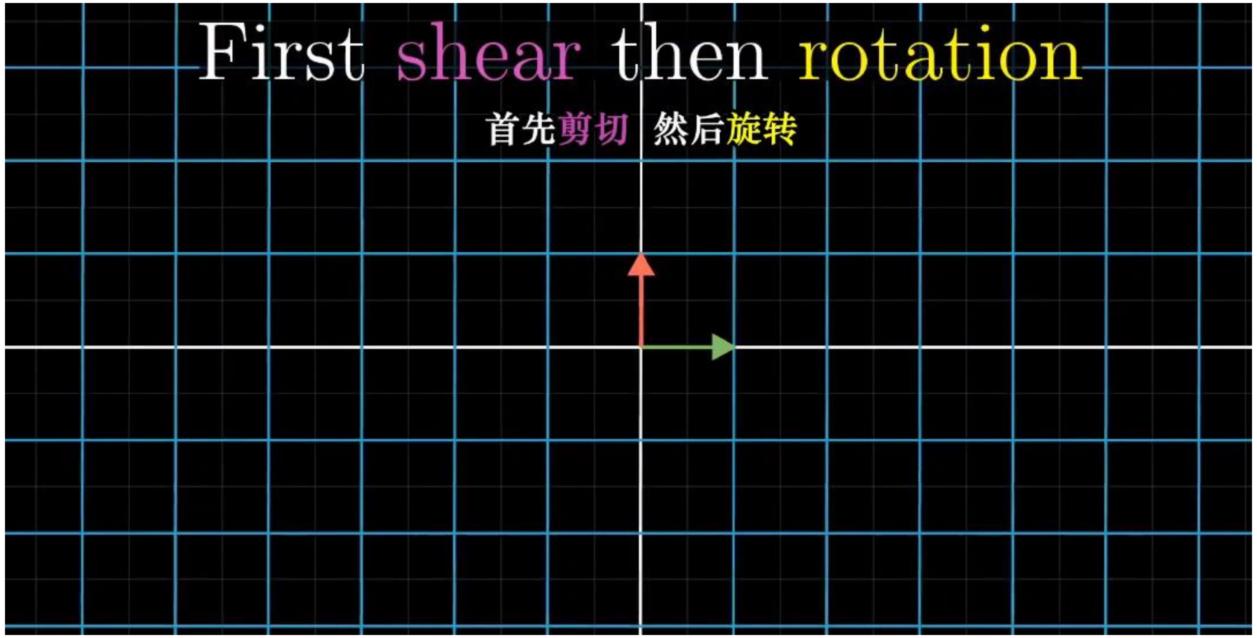
矩阵乘法就是复合变换：两个变换先后作用。

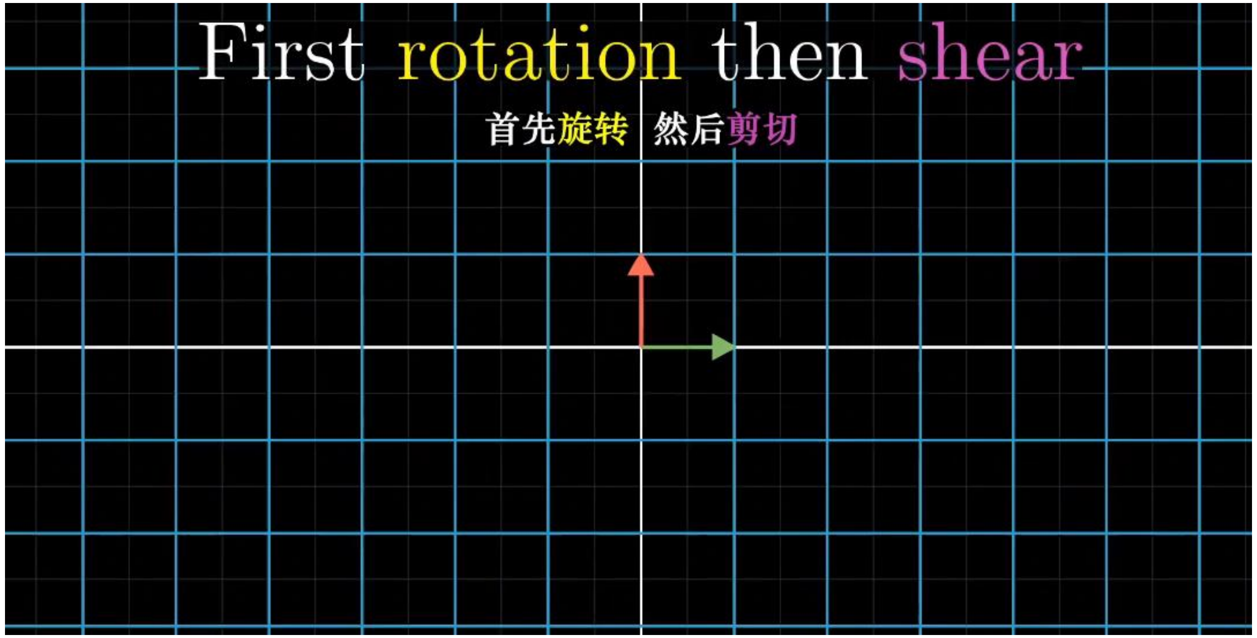
矩阵乘法的计算：行和列分别相乘。

例：

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

思考：矩阵乘法满足交换律吗？


First shear then rotation
首先剪切 然后旋转


First rotation then shear
首先旋转 然后剪切

5. 行列式

既然矩阵可以将空间做伸缩，我们不禁要问：
一次线性变换中，究竟空间被拉伸/压缩了多少？

以2维空间为例：
既然线性变换是改变基组，那么我们只需要找到一个指标来度量2个基向量围成的矩形面积 增大或缩小的比例。

该指标就是行列式的值。

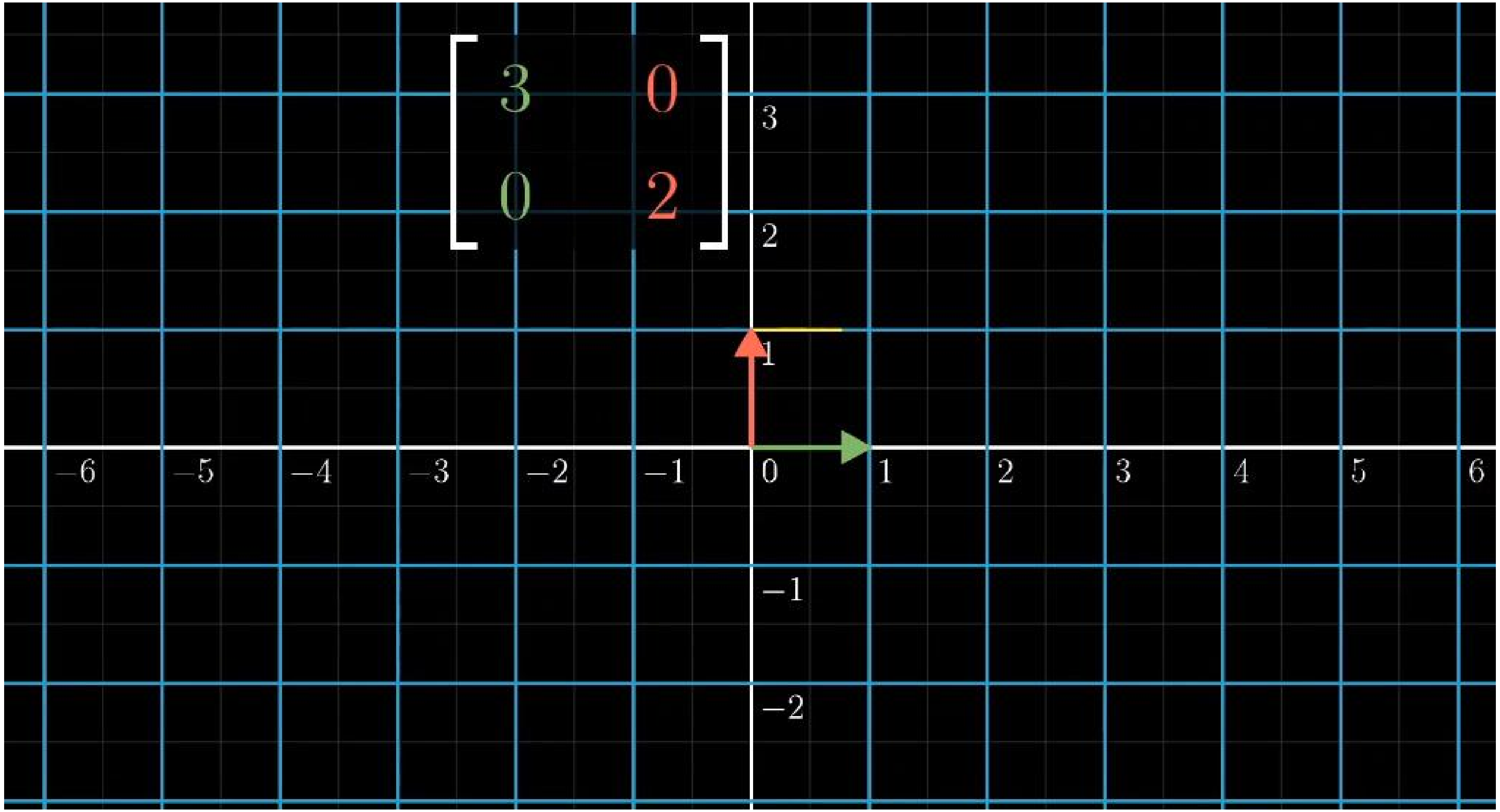
记作det()，或| |.

(注：在代数学中，n阶行列式最基础的定义是一个n次多项式)

例：

矩阵$\begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}$就是将基组围成的矩形面积扩大了6倍。

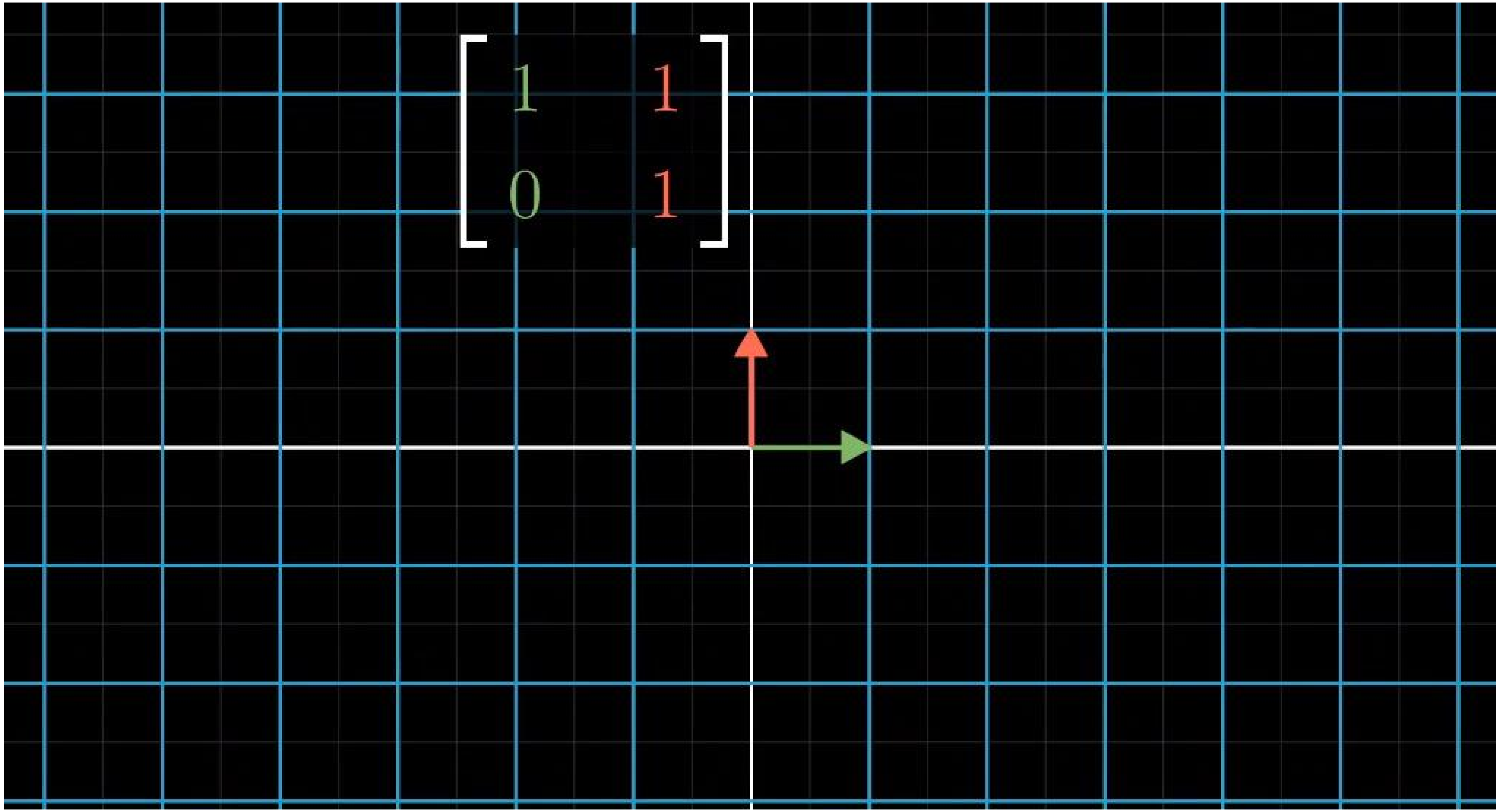
因此该矩阵的行列式$\begin{vmatrix} 3 & 0 \\ 0 & 2 \end{vmatrix}=6$



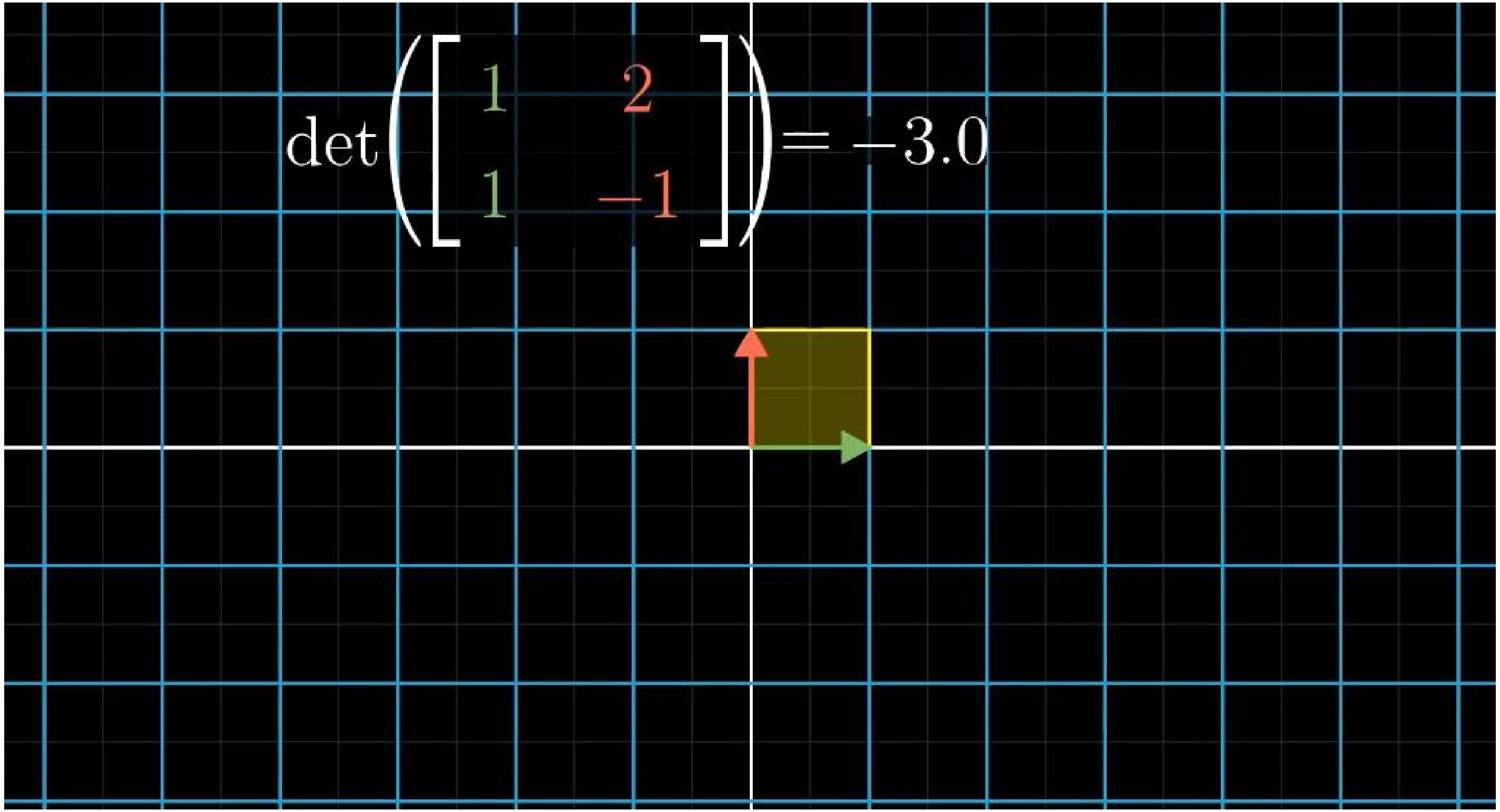
CDA数据分析师 CERTIFIED DATA ANALYST

而矩阵$\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$只是让空间倾斜，但矩形面积不变，仍为1.

因此该矩阵的行列式$\begin{vmatrix}1 & 1\\ 0 & 1\end{vmatrix}=1$



若行列式的值为负，表示空间发生了翻转，交换了坐标轴的顺/逆时针顺序。



若行列式的值为0，则是将平面压缩维数变成直线，矩形面积当然就为0.
因此检查矩阵行列式的值是否等于零，可以检验该矩阵的变换是否将空间降维。

$\det(MN)=\det(M)\det(N)$
两个矩阵 M，N，积的行列式=行列式的积

行列式行数必须等于列数。也就是说必须是方阵才有行列式。

选学

行列式的计算方法总结（由易到难）：

1. 通过初等行变换化为上（下）三角行列式
2. 行列式按行（列）展开

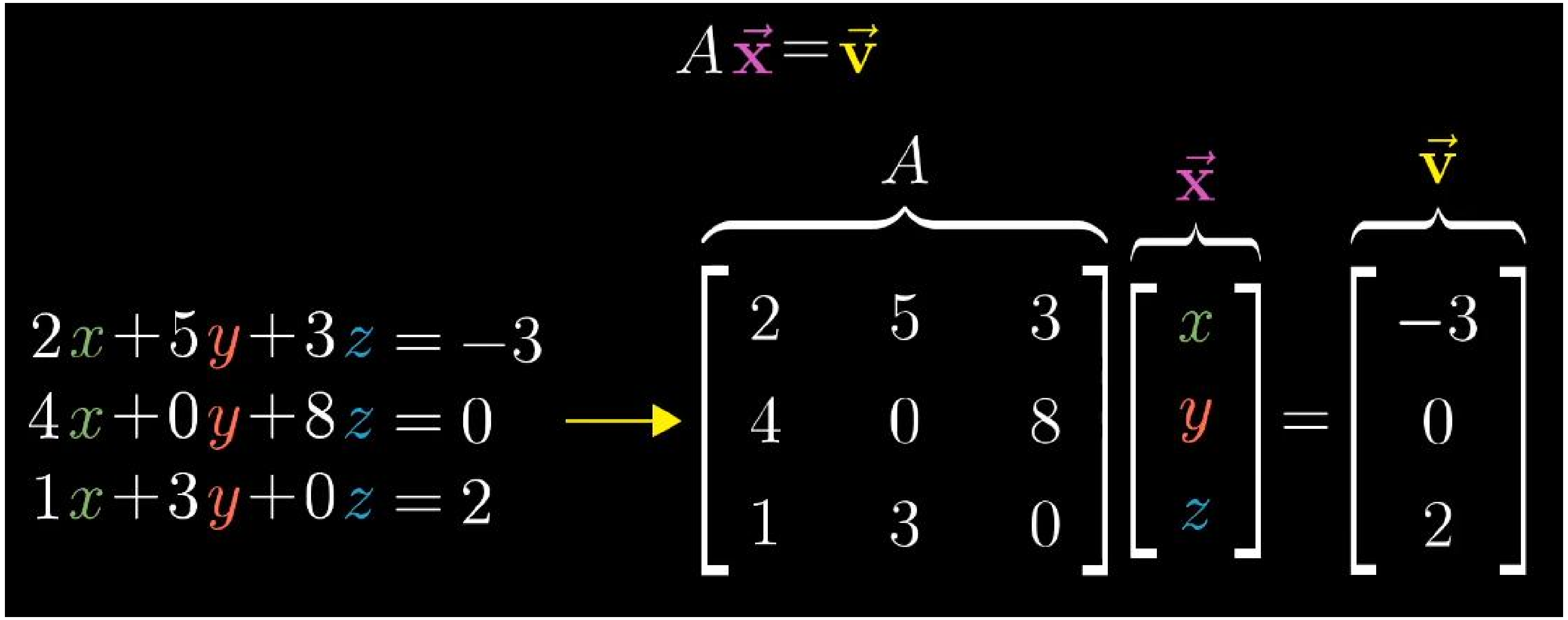
3. 应用行列式的定义计算
4. 分块

5. 范德蒙行列式
6. 拆分法

6. 逆矩阵和矩阵的秩

线性方程组可以转化成矩阵形式 $A\vec{x}=\vec{v}$.
也就是找到一个向量$\vec{x}$，使它在经过矩阵A变换后，成为$\vec{v}$ 。

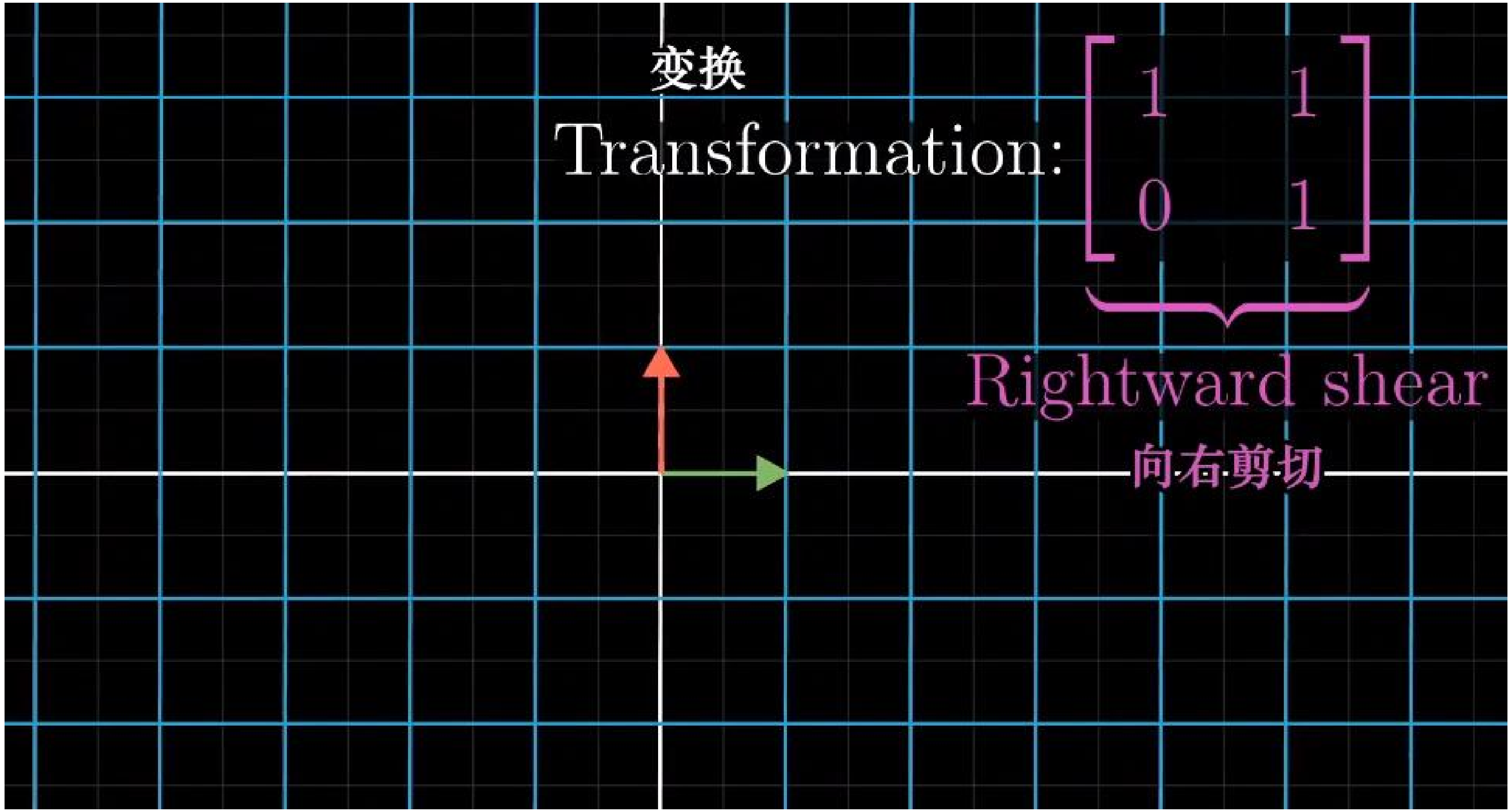
实际计算中，我们其实找的是逆变换A^{-1}，使$\vec{v}$ 经A^{-1}变换后得到$\vec{x}$。
A^{-1}叫做A的逆矩阵。



变换与逆变换

此例中出现的2个变换的矩阵即互为逆矩阵。

思考：
这两个矩阵的复合变换是什么?



$A^{-1}A$ 必须等于什么都没做。

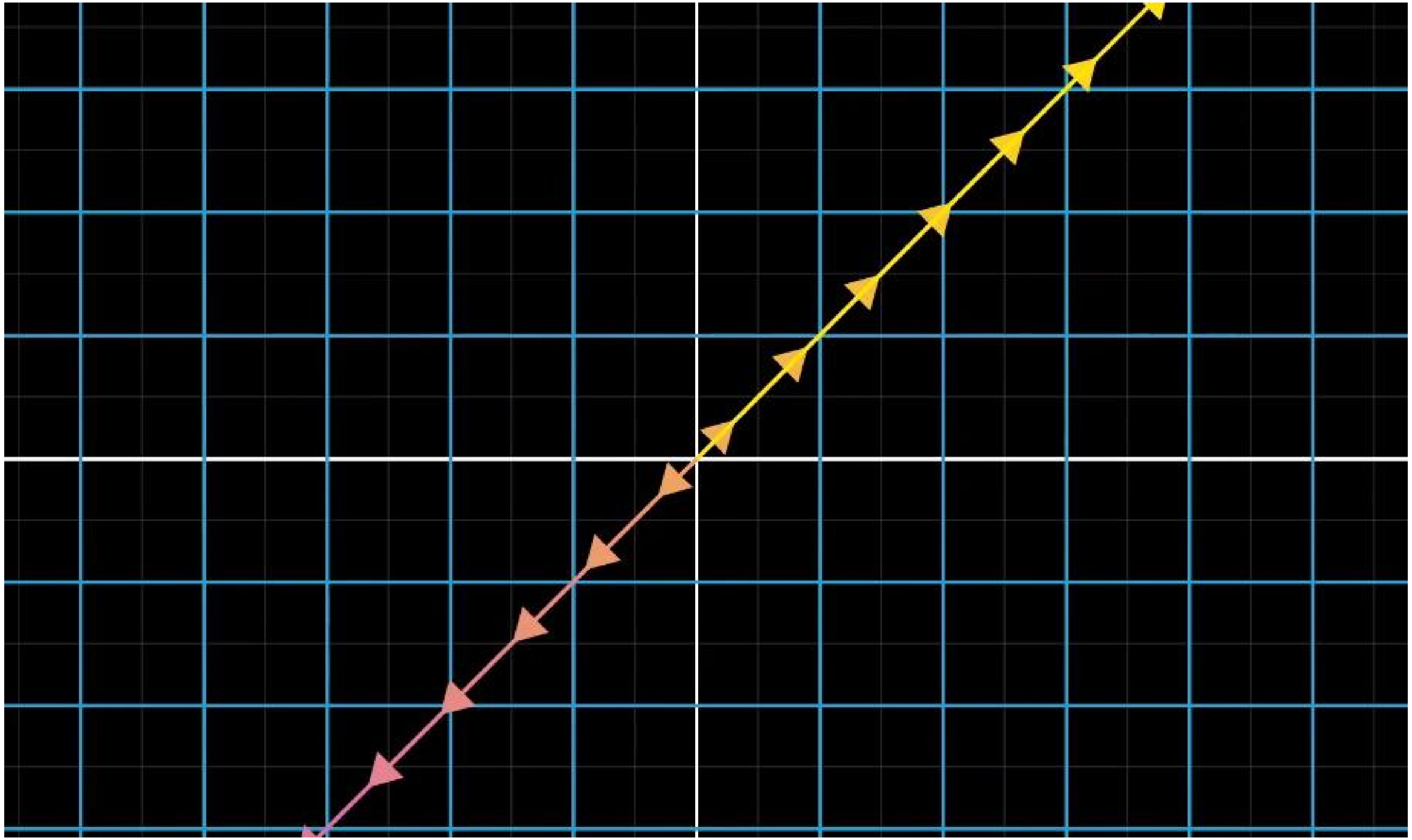
那么，在所有矩阵中主对角线元素为1，其他元素都为0的矩阵，就代表什么都不做的变换。
这种矩阵称为单位矩阵，记作I或者E.

所以 $A^{-1}A = E$

但是，在线性代数里，当det(A)=0时，空间被降维后就没有对应的逆变换A^{-1}能变回去了。

因此当det(A)=0，逆不存在。

不可逆矩阵称为奇异矩阵。
可逆矩阵称为非奇异矩阵。



核（Kernel):

所有被线性变换变成零向量的向量组成的集合，称为该变换的“核”，即:

$$Ker\mathcal{A} = \{v \in V: \mathcal{A}v = 0\}$$

在几何上，核就是即将在变换后落在原点的向量的集合。

值域（Range）：

所有被线性变换映射出的向量的集合，称为该变换的“值域”，就是$\mathcal{A}v$。

秩

矩阵的秩 ： 线性变换后空间的维数。

满秩：秩与矩阵的列数相等

例如，对于2*2矩阵，秩最大为2.也就是基向量最多张成整个2维空间，此时det≠0.
对于一个满秩变换来说，唯一能在变换后落在原点的是零向量自身。

不满秩：变换后维度降低，此时det=0.
若不满秩，则有很多向量被压到原点。图例见上页。

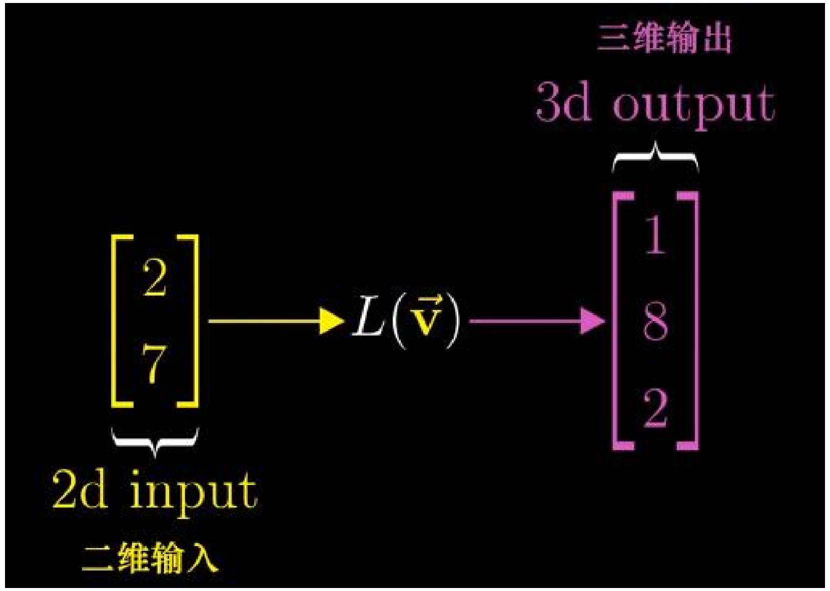
（注：秩的原始定义：向量组中线性无关的向量的个数。）

7. 非方阵

以上的大部分例子，我们都是用2*2的矩阵来举例做变换，表示二维到二维的映射。那么，非方阵呢?

非方阵表示不同维数之间的线性变换。

例:
3*2矩阵表示二维映射到三维
2*3矩阵表示三维映射到二维



请计算 $\begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}\begin{pmatrix} g \\ h \end{pmatrix} = ?$

当矩阵是1*2时，是2维到1维的变换。

例：

$(a,b)\binom{c}{d}=ac+bd$

这就是下节所介绍的点乘（内积）。换句话说，点乘就是把空间压缩到一维。

8. 点乘，内积

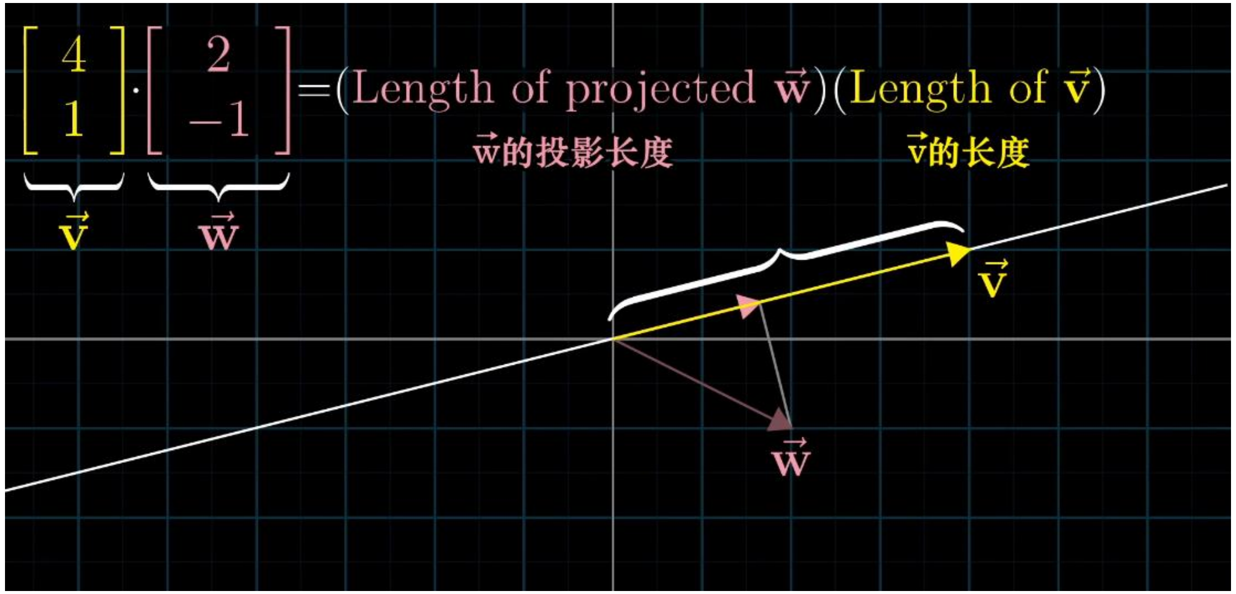
点乘也叫内积。

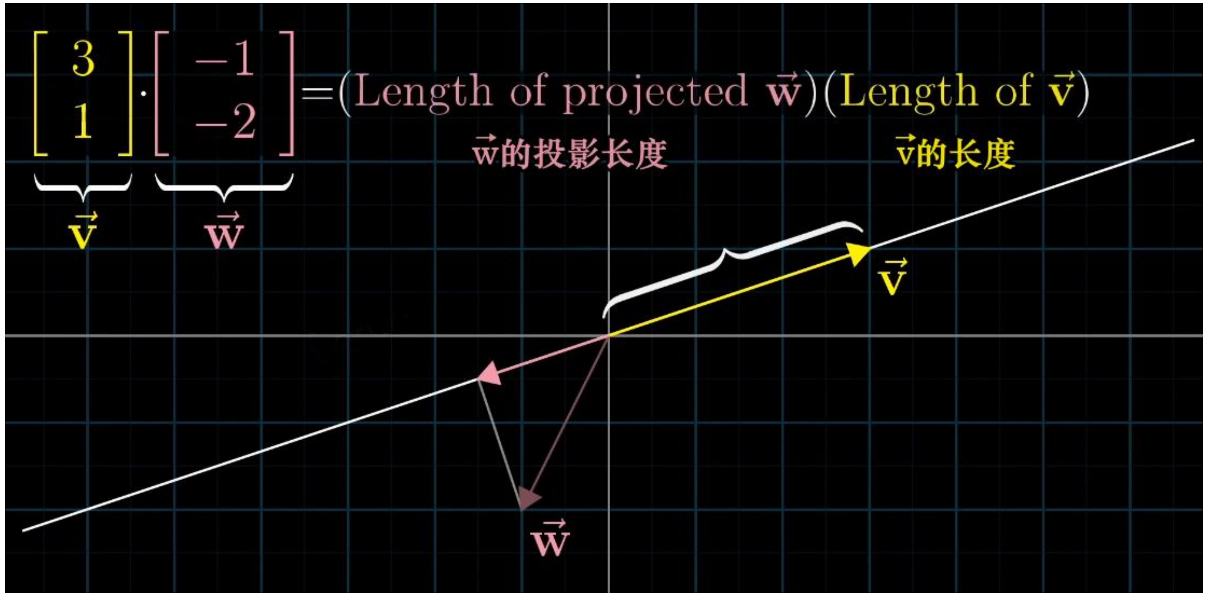
2个向量做内积，例如 $\vec{v} * \vec{w}$ 就是把向量$\vec{w}$向向量$\vec{v}$作投影，此时

内积=投影长度 * $\vec{v}$的长度

内积一般记作（α, β）。注意容易和行向量弄混，需结合上下文加以区分。

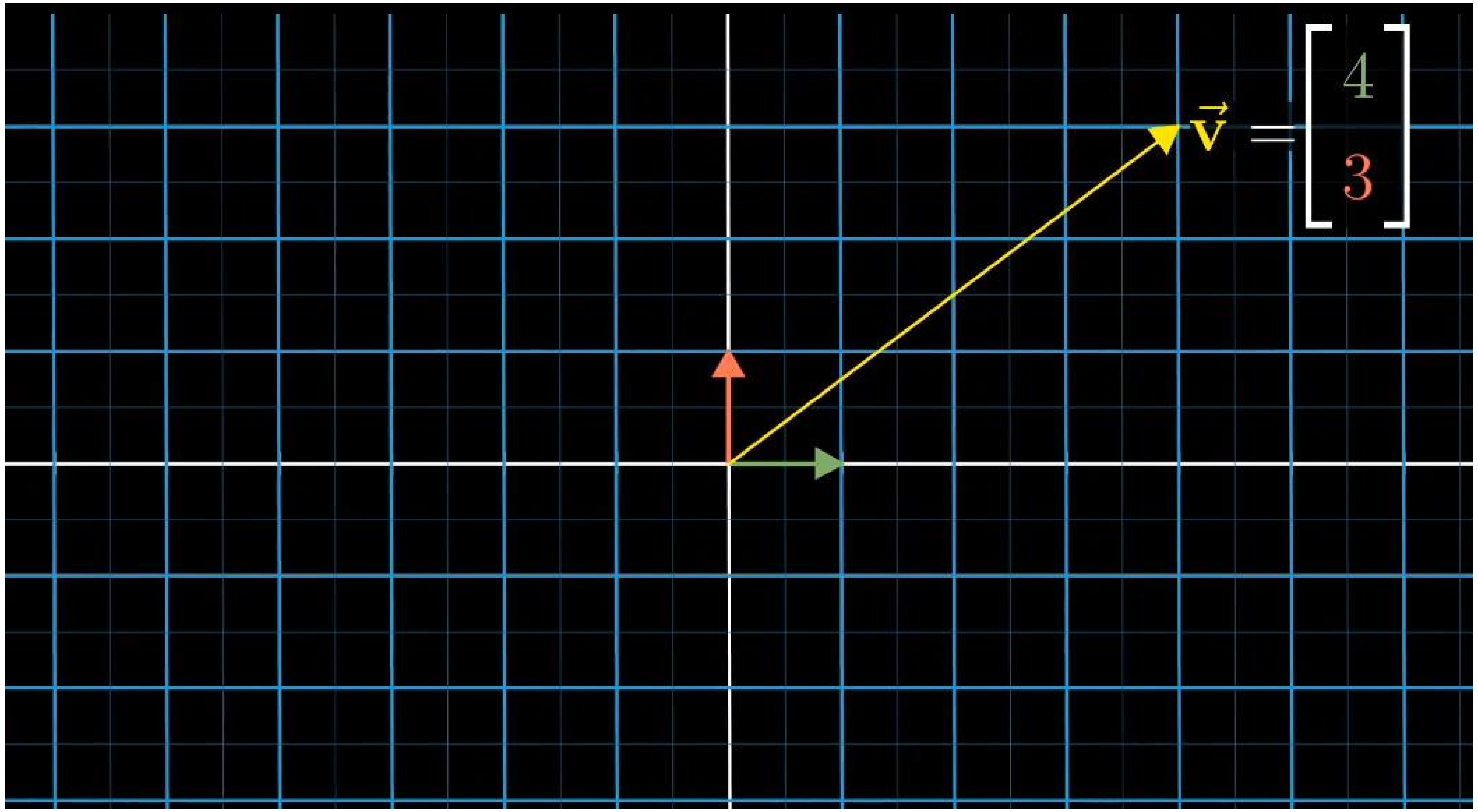
同侧投影内积为正，否则为负。





点乘也可以看作把空间压缩到一维。

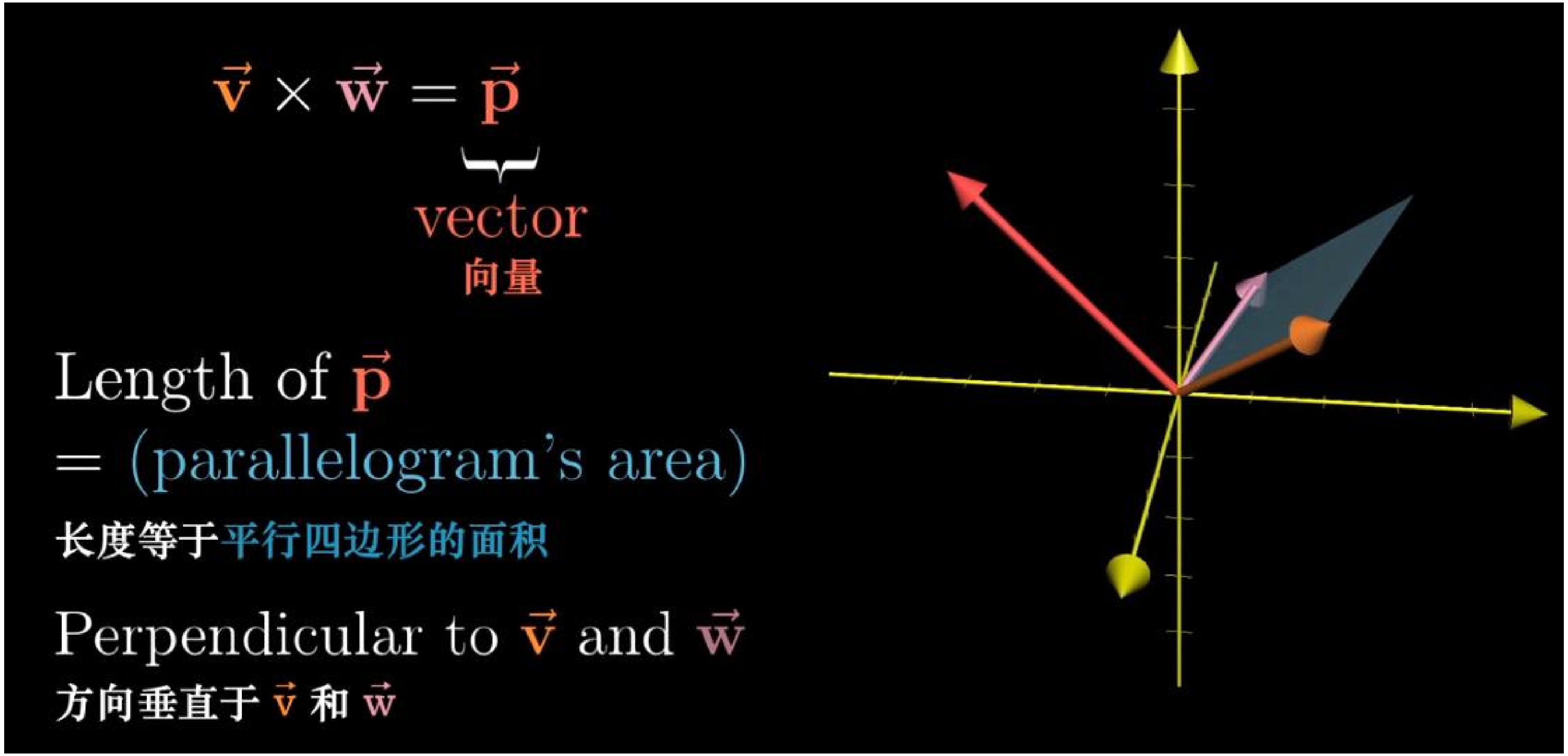
$$(1,-2)\begin{pmatrix}4\\3\end{pmatrix}=\begin{pmatrix}1\\-2\end{pmatrix}*\begin{pmatrix}4\\3\end{pmatrix}=-2$$



小技巧：线性代数中，和投影有关的问题，往往需要应用内积（点乘）。

9. 叉乘，外积

实际应用中，外积（叉乘）一般用来计算多维几何体的体积，例如重积分。



10. 特征值与特征向量

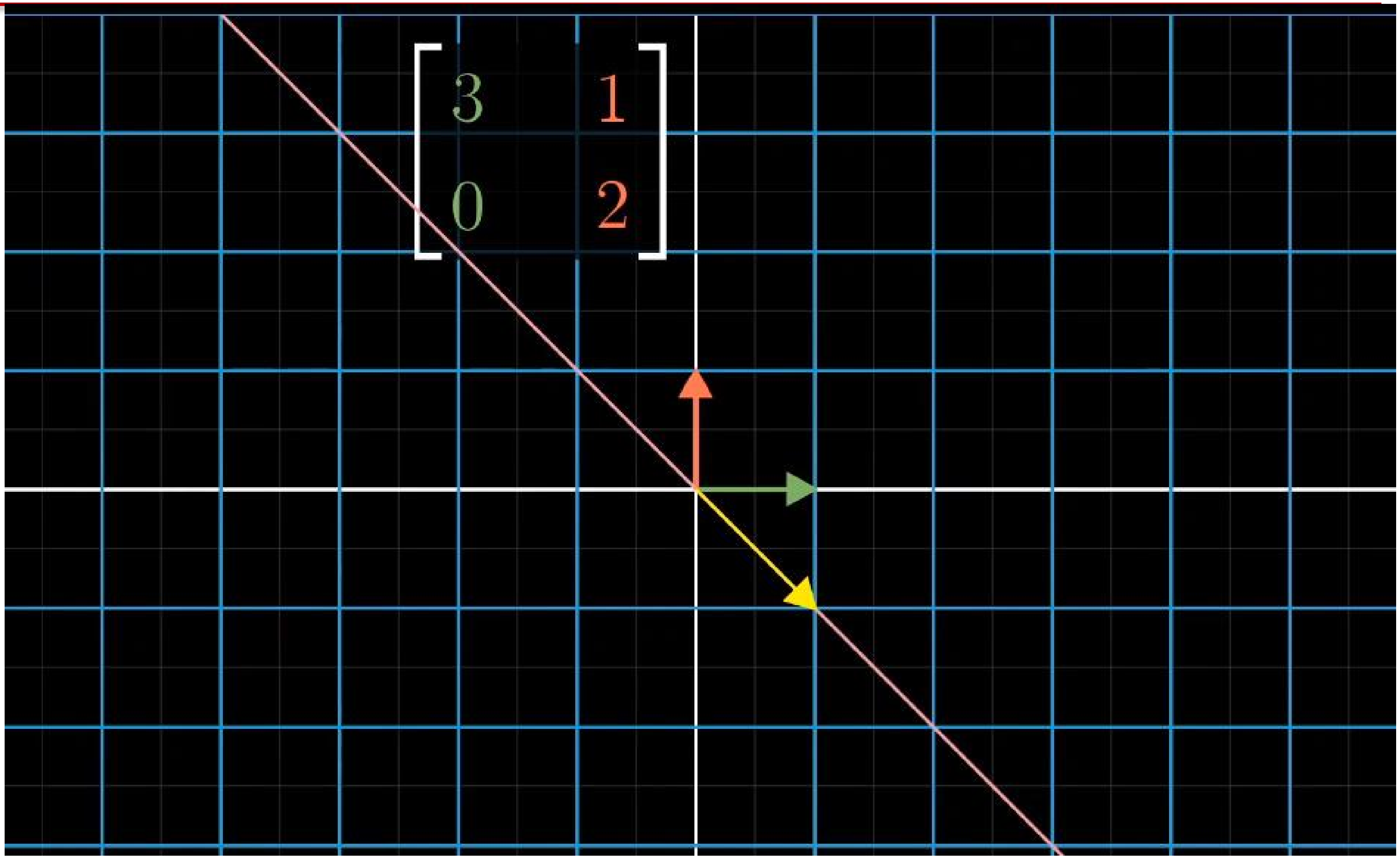
线性变换中，大部分向量在变换后会离开它的张成空间，但也有一些会留在它所张成的空间中。

以 $\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ 为例:

绝大部分向量在变换后都离开了该向量的张成空间，例如Y轴的单位向量。
但也有2个向量留在了它们分别的张成空间中，这两个向量是
45度线$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ 和 X轴$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

图例见下页，请首先回忆一下张成空间的概念。

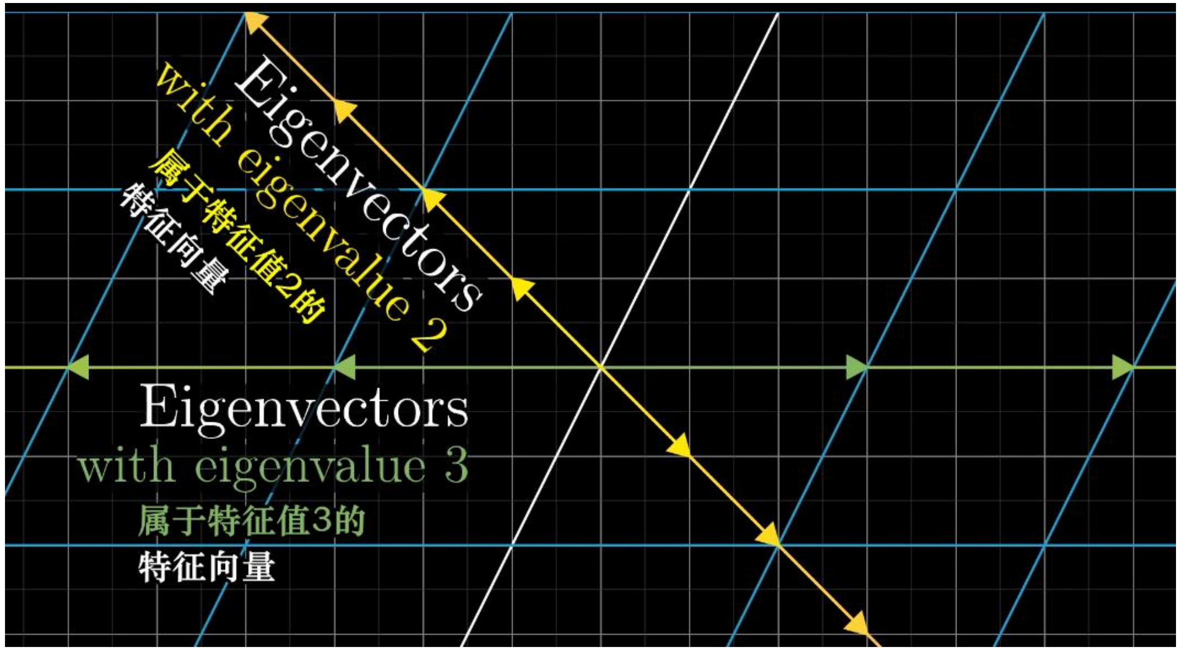
请注意绿色和黄色向量分别被拉伸了多少倍?



上例中，$\begin{pmatrix}1\\-1\end{pmatrix}$向量被拉伸了2倍，则$\begin{pmatrix}1\\-1\end{pmatrix}$叫做属于特征值2的特征向量。

$\begin{pmatrix}1\\0\end{pmatrix}$被拉伸了3倍，则$\begin{pmatrix}1\\0\end{pmatrix}$就是属于特征值3的特征向量。

公式 $A\vec{v} = \lambda\vec{v}$



注意:

1. 实数域中，并不是所有的矩阵都有特征值和特征向量。

2. 如果所有的基向量都是特征向量，此矩阵一定是对角阵，例如$\begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix}$

3. 更高维度的空间中，特征向量往往是变换的旋转轴。

需要掌握的计算

1. 向量运算

2. 矩阵的加法、数乘。矩阵乘法。

3. 矩阵的转置

4. 矩阵的初等变换（高斯消元）

5. 求矩阵的逆

6. 求矩阵的秩（行阶梯型与行最简型）

7. 求矩阵的特征值和特征向量

（选学）矩阵计算补充：打洞原理

古人云：龙生龙，凤生凤，华罗庚的学生会打洞。

打洞原理是专门用于分块矩阵的计算的。
我们知道，初等行（列）变换等于矩阵左（右）乘一初等矩阵。

例：

$$\begin{pmatrix} E_n & 0 \\ -CA^{-1} & E_m \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} E_n & -A^{-1}B \\ 0 & E_m \end{pmatrix}=\begin{pmatrix} A & 0 \\ 0 & D-CA^{-1}B \end{pmatrix}$$

用谁去干掉别人，就从谁的对角开始顺时针打洞。
本题用A干掉C（行变换，对应左乘）和B（列变换，对应右乘），所以都从对角的D开始打洞。
最后的结果必定是 $-\Diamond\triangle^{-1}\square$