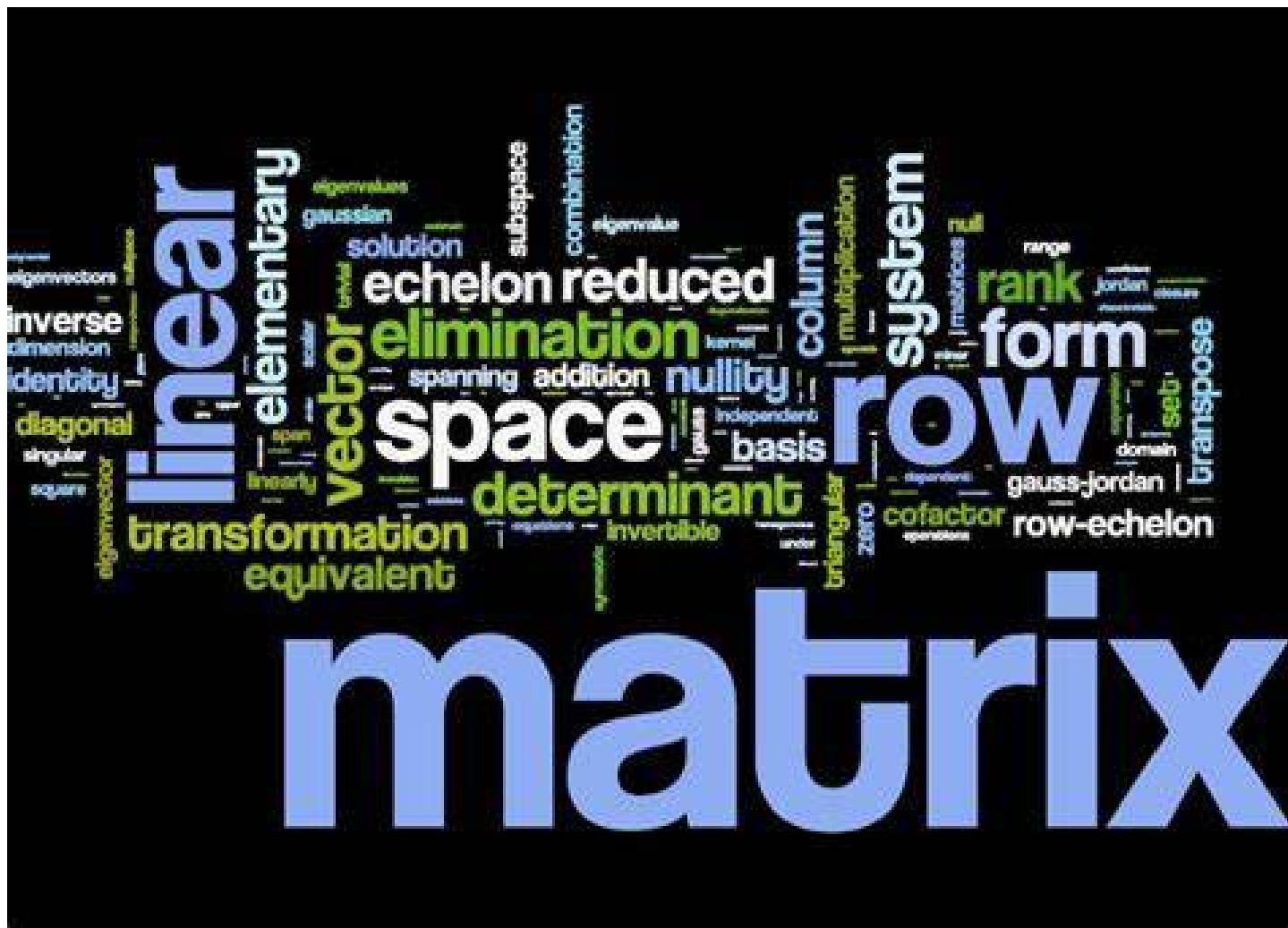


知乎

首发于
机器学习中的数学拾贝

线性代数的本质(10)-特征值与特征向量



初瞳

幸得识卿桃花面，从此阡陌多暖春。

关注她

20 人赞同了该文章

10.0 总结

♡线性变换的两种理解方式

- 将基向量变化后的位置视为矩阵的列，也就是新的基向量【依赖于坐标系】
- 利用特征向量和特征值理解线性变换，不依赖于坐标系的选择

♡特征值与特征向量

特征向量：一个向量经过线性变换，仍留在它所张成的空间中

特征值：描述特征向量经过线性变换

已赞同 20



2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝♡求解特征值就是寻找一个 λ 使得 $\det(A - \lambda I) = 0$

♡当特征向量的数量足够多时，这些特征向量就可以构成特征基。利用特征基可以简化矩阵的幂次计算。

为了方便理解，将总结前后copy了一份，与10.5章节的内容保持一致。

10.1 特征值与特征向量的意义

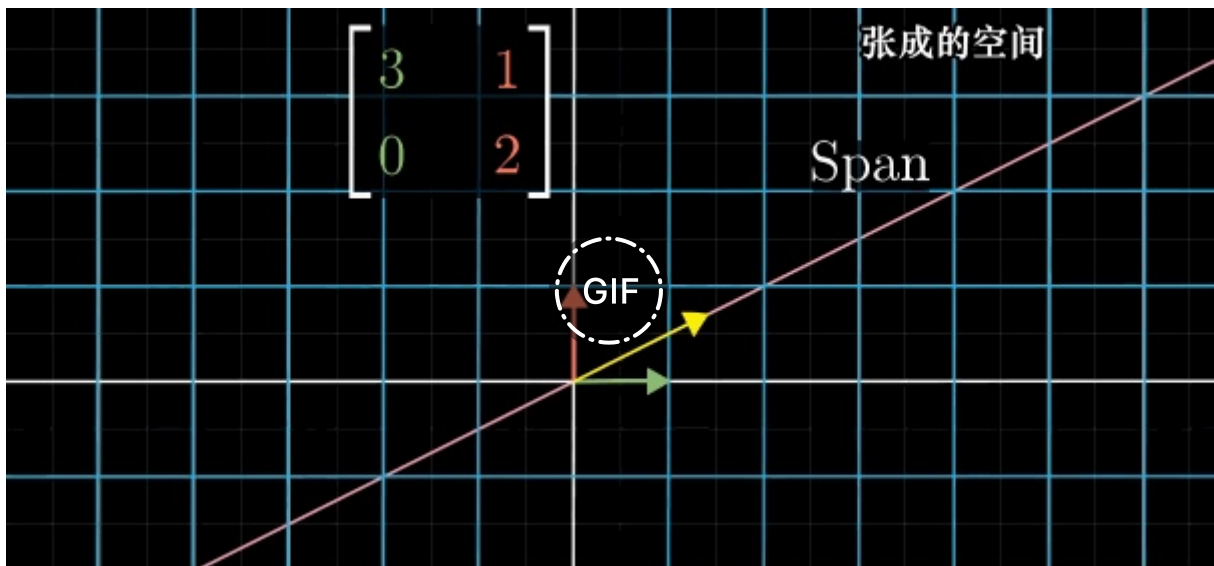
总结：

特征向量：一个向量经过线性变换，仍留在它所张成的空间中

特征值：描述特征向量经过线性变换后的缩放程度

正文：

大部分的向量经过线性变换都离开了它所张成的空间。如图中黄色的向量经过线性变换 $\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$ ，离开了它所张成的通过原点的直线张成的空间。



但是也有一部分向量留在了它所张成的空间，这样线性变换起到的作用就是**压缩或者拉升**如：

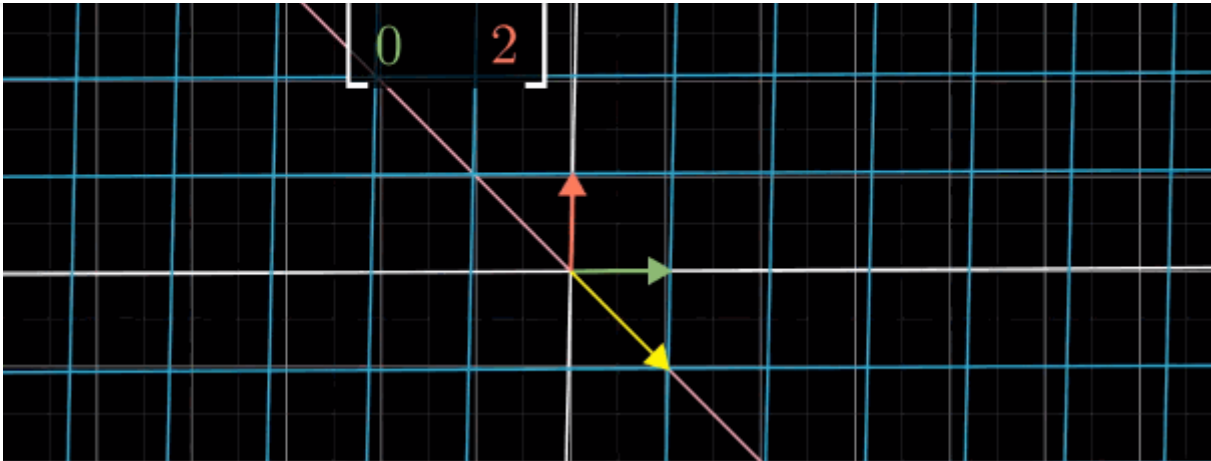
已赞同 20

2 条评论

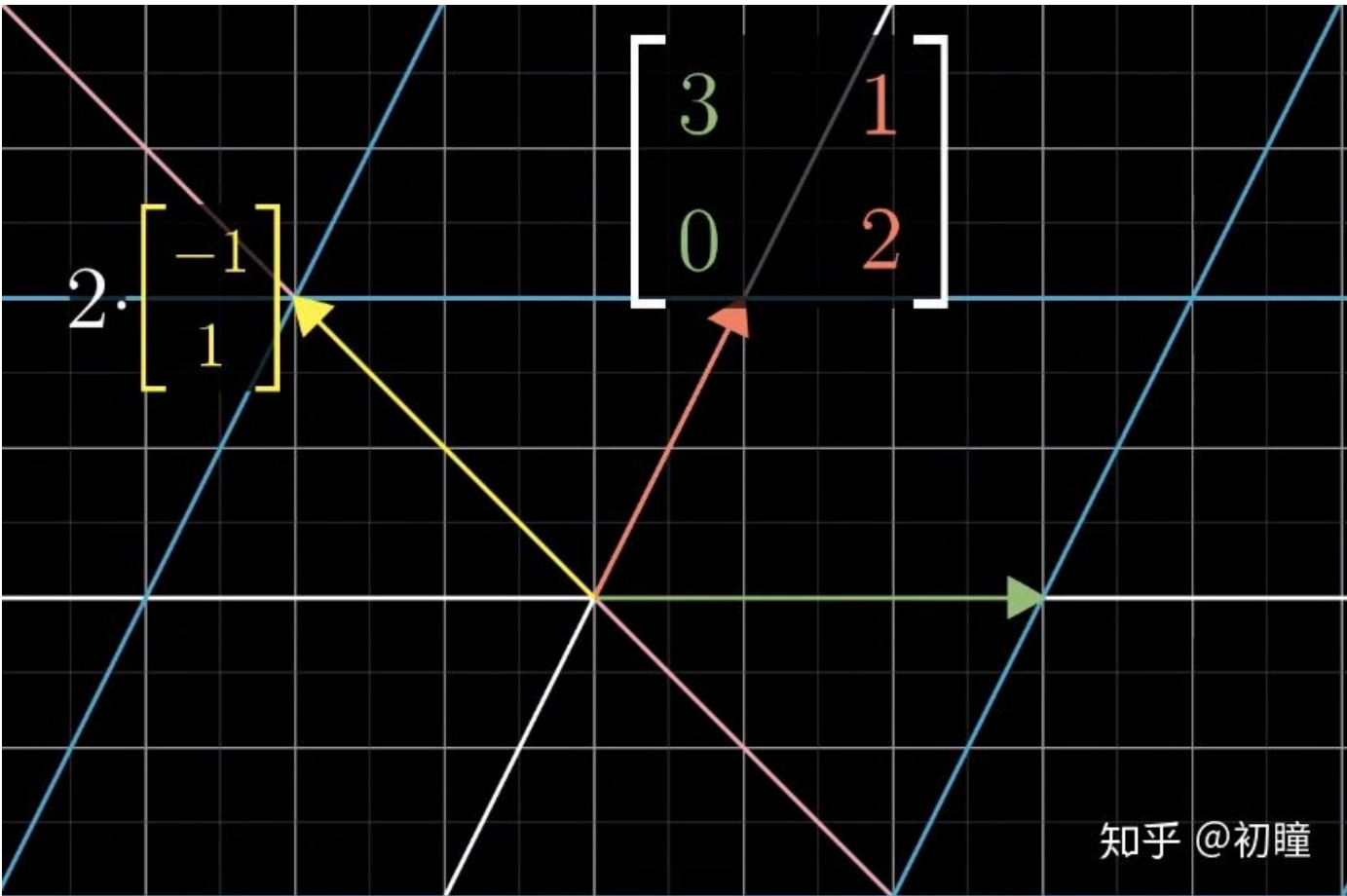
分享

★ 收藏

...

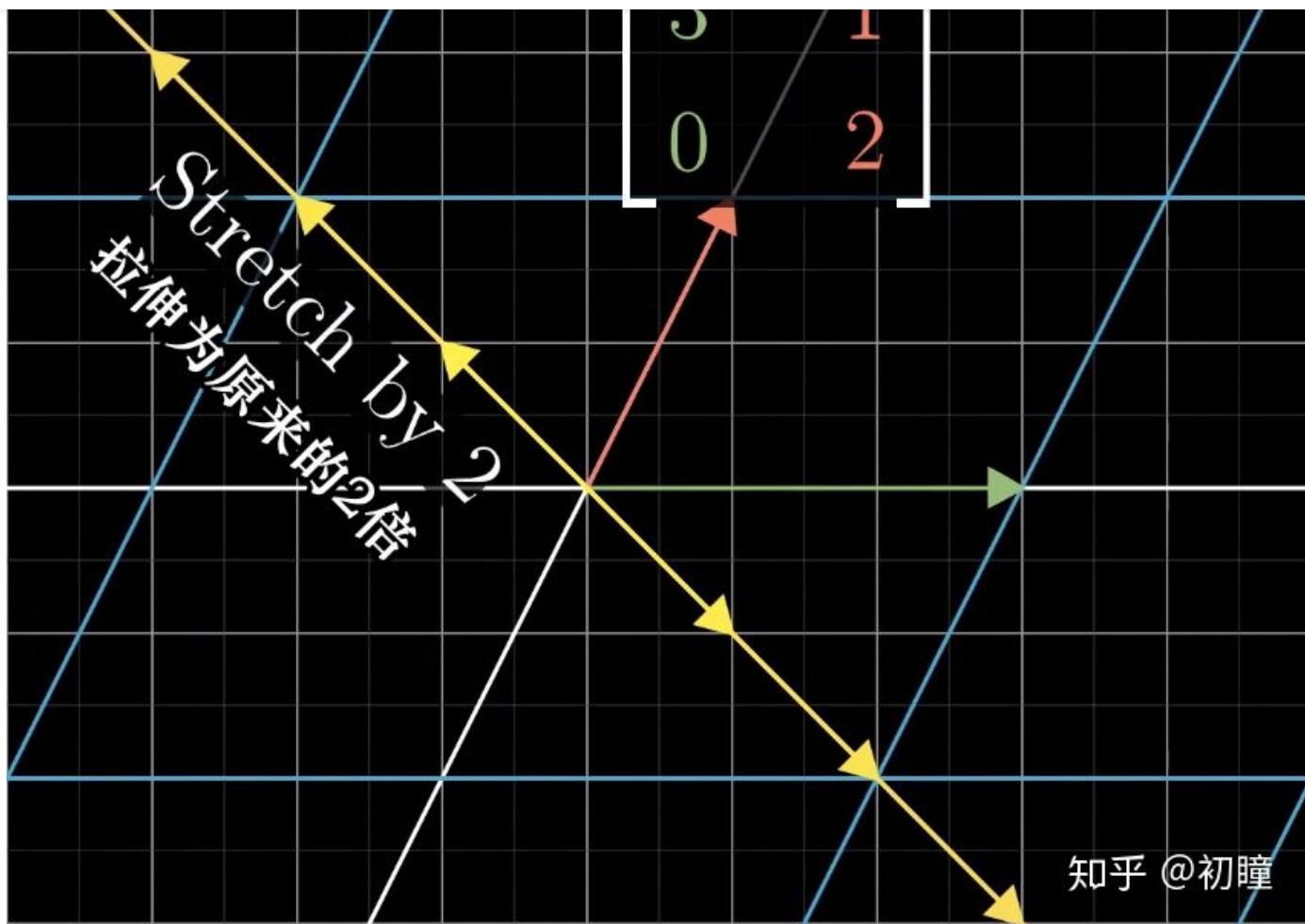


特别的, 对于 $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ 经过线性变换 $\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}$, 相当于将该向量拉升了两倍。



线性变换暗示一点, 对于处在它所在的对角线上的任何向量也仅被拉伸了2倍。

知乎

首发于
机器学习中的数学拾贝

他们的共同点就是线性变换后留在了原来向量所张成的线性空间里边，对于其他的向量则离开他原来张成的向量空间

已赞同 20



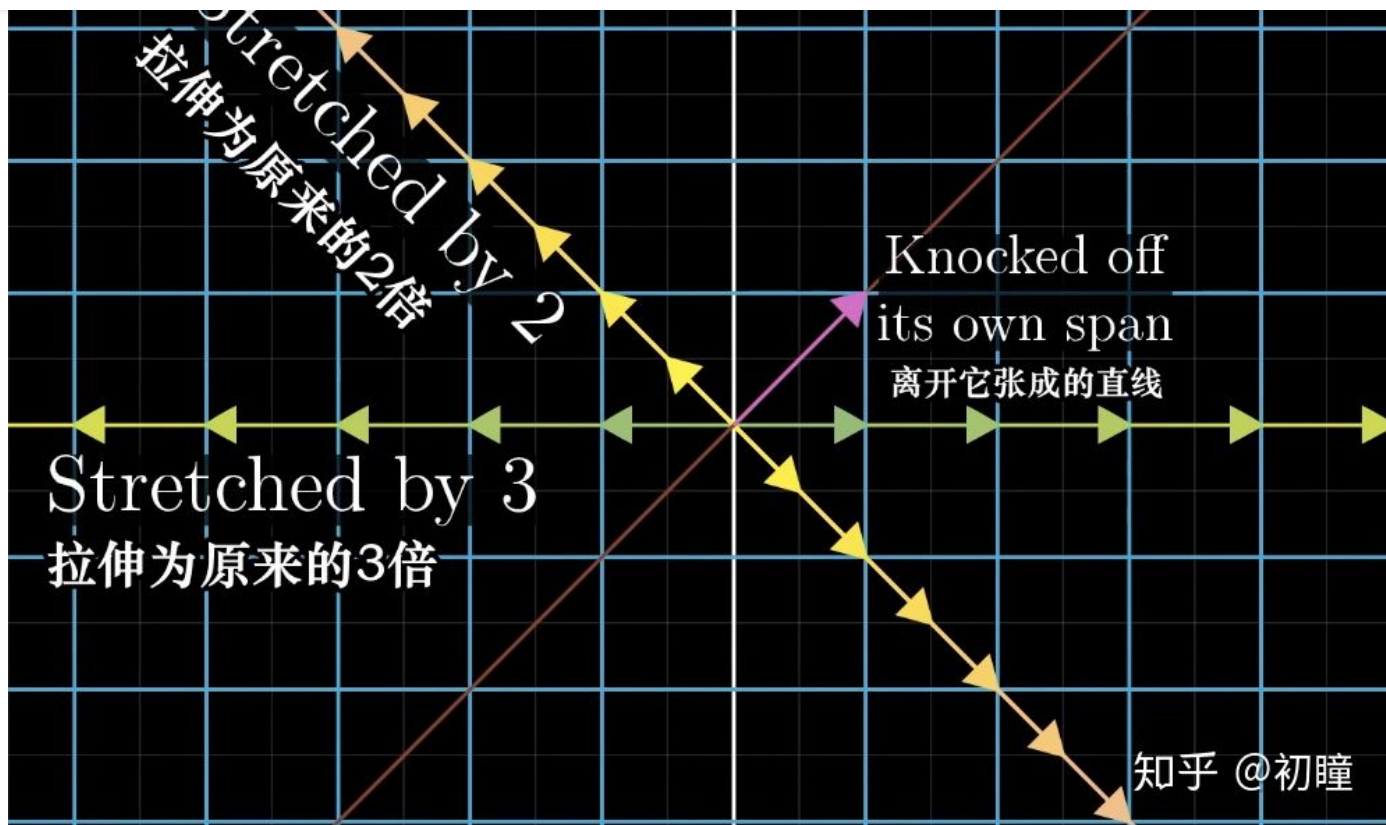
2 条评论

分享

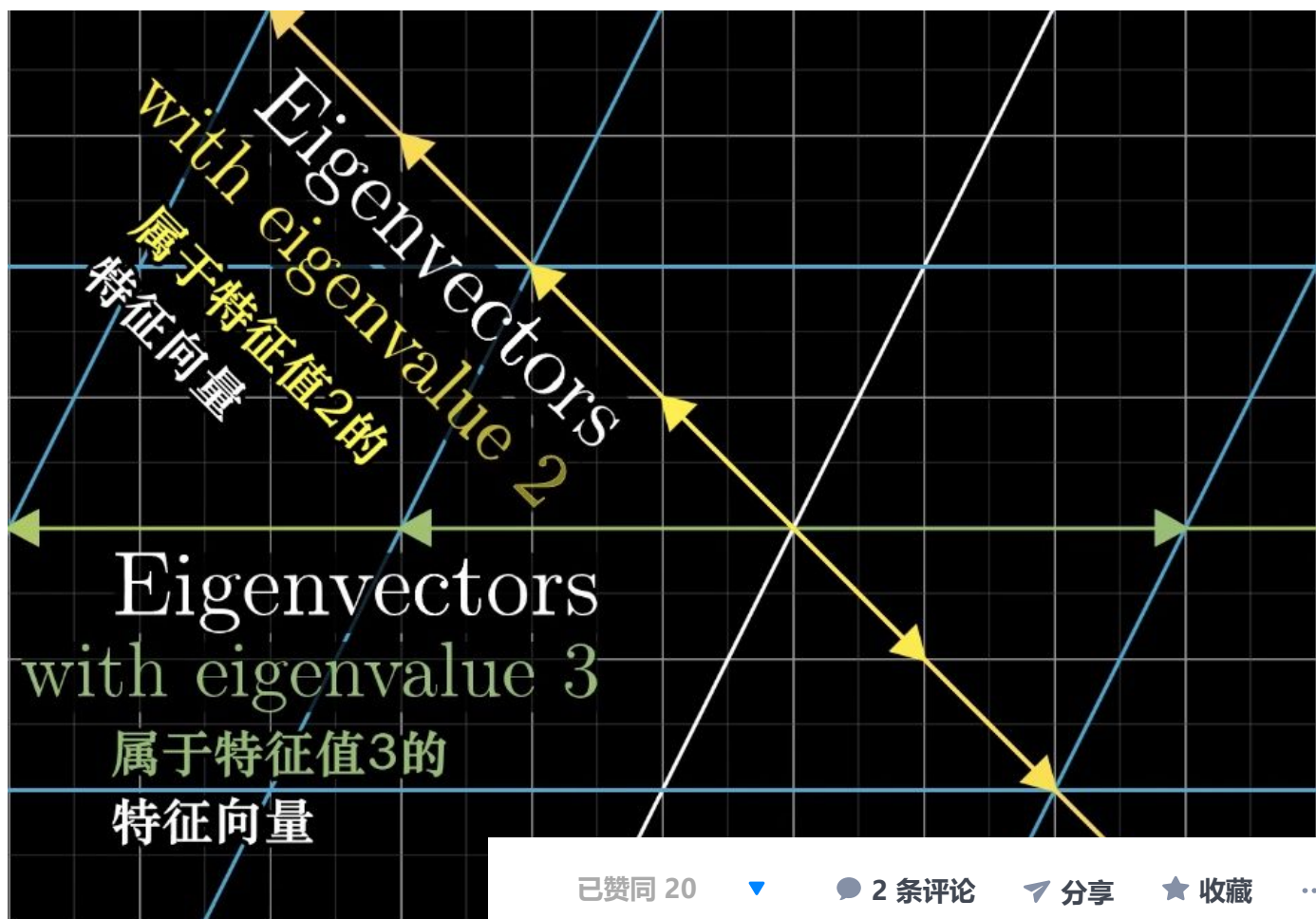
★ 收藏



知乎

首发于
机器学习中的数学拾贝

因此将这些特殊的向量称为“特征向量”，每个特征向量都有一个对应的值，称为“特征值”，目的是衡量特征向量经过变换后的压缩或者拉伸因子。



已赞同 20

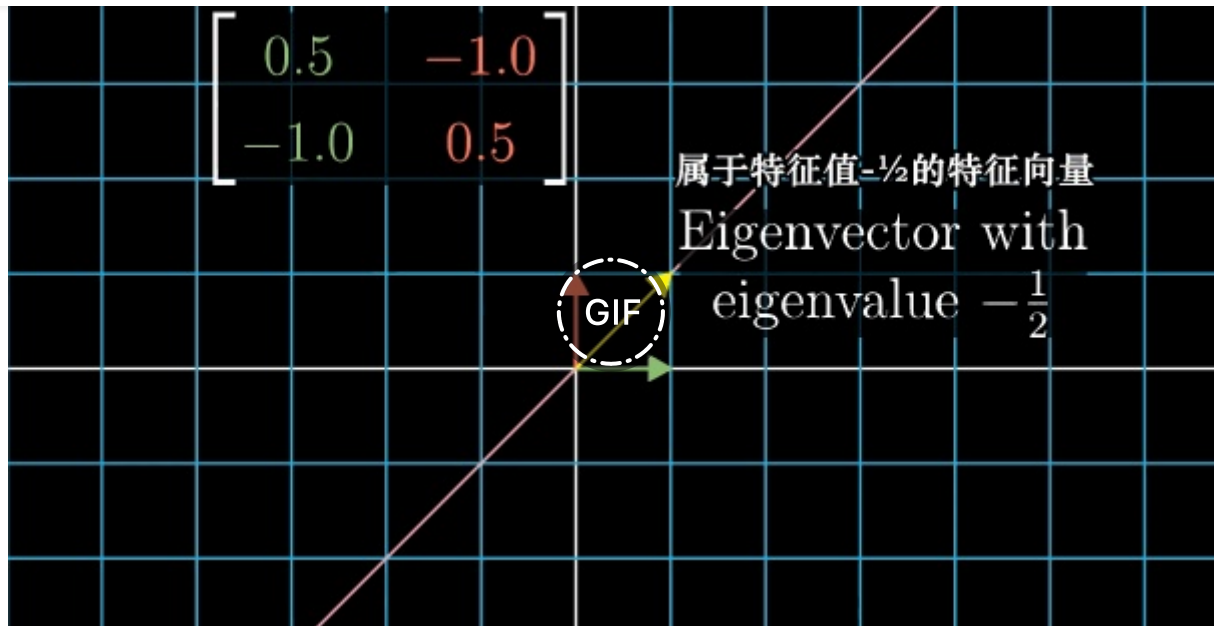
2 条评论

分享

★ 收藏

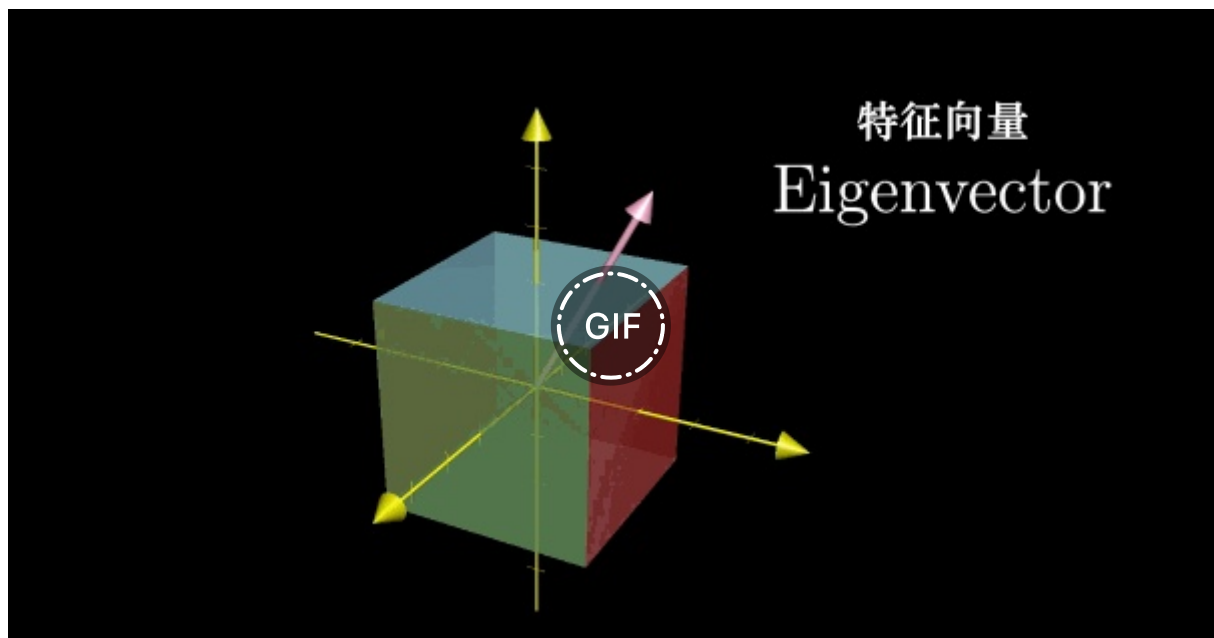
...

知乎

首发于
机器学习中的数学拾贝

10.2 三维空间中的特征值与特征向量

对于三维空间中的旋转，如果能够找到对应的特征向量，也即能够留在它所张成的空间中的向量，那么就意味着我们找到了旋转轴。特别地，**这就意味着将一个三维旋转看成绕这个特征向量旋转一定角度**，要比考虑相应的矩阵变换要直观。



此时对应的特征值为1，因为**旋转并不改变任何一个向量**，所以向量的长度保持不变。

10.3 计算特征值与特征向量

已赞同 20



2 条评论

分享

★ 收藏

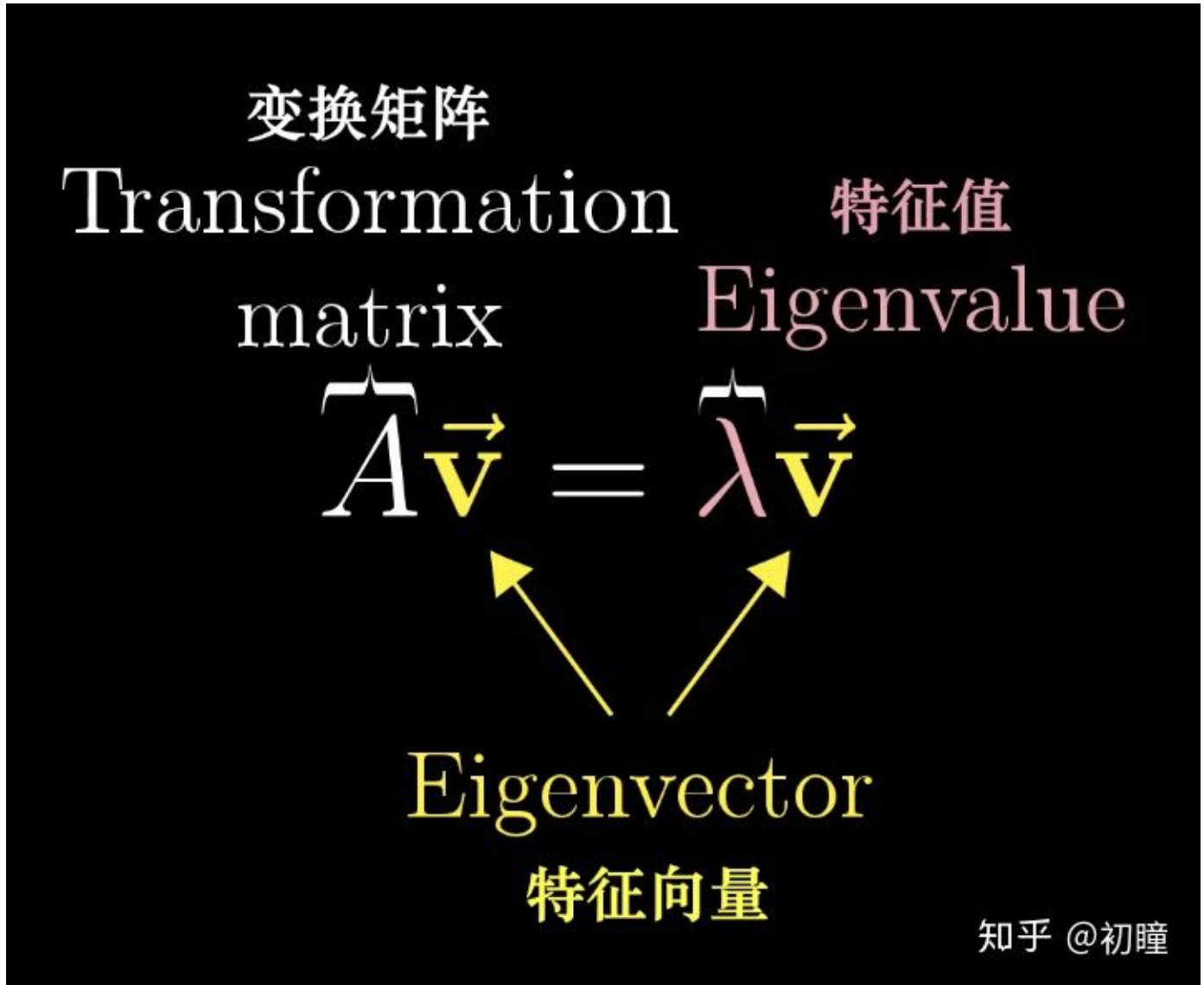


知乎

首发于
机器学习中的数学拾贝

理解线性变换作用，如果采用矩阵的列可以看做变换后的基向量，这就依赖于特定的坐标系。为了减少对坐标系的依赖，更好的办法就是求出对应的特征向量与特征值。

10.3.2 计算特征值与特征向量



线性变换

已赞同 20



2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝

Matrix-vector multiplication

$$\overbrace{A\vec{v}} = \underbrace{\lambda\vec{v}}$$

Scalar multiplication

向量数乘

知乎 @初瞳

矩阵乘法

可以看得出，线性变换的理解与矩阵乘法的理解是一致的。矩阵乘法的作用效果就是将任意向量乘以 λ 。这个矩阵的列代表着变换后的基向量。

下一步就是：将等式左右两侧都转化为矩阵向量乘积的形式

与数 λ 相乘Scaling by λ

与这个矩阵相乘



Matrix multiplication by

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

$$\lambda \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_I$$

知乎 @初瞳

已赞同 20



2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝

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

This matrix looks something like

这个矩阵是这种形式的

$$\begin{bmatrix} 3 - \lambda & 1 & 4 \\ 1 & 5 - \lambda & 9 \\ 2 & 6 & 5 - \lambda \end{bmatrix}$$

知乎 @初瞳

等式变换的最终形式

这样就是寻找一个向量 \vec{v} 使得这个新矩阵 $A - \lambda I$ 相乘的结果等于0

显然, $\vec{v} = \vec{0}$ 恒成立, 但是我们要寻找的是非零解。基于行列式的知识, 就是相当于求 $\det(A - \lambda I) = 0$, 当且仅当矩阵所代表的线性变换将空间压缩到更低的维度时, 才会存在一个非零向量, 使得矩阵与它的乘积为零向量。

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

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

空间压缩

知乎 @初瞳

已赞同 20



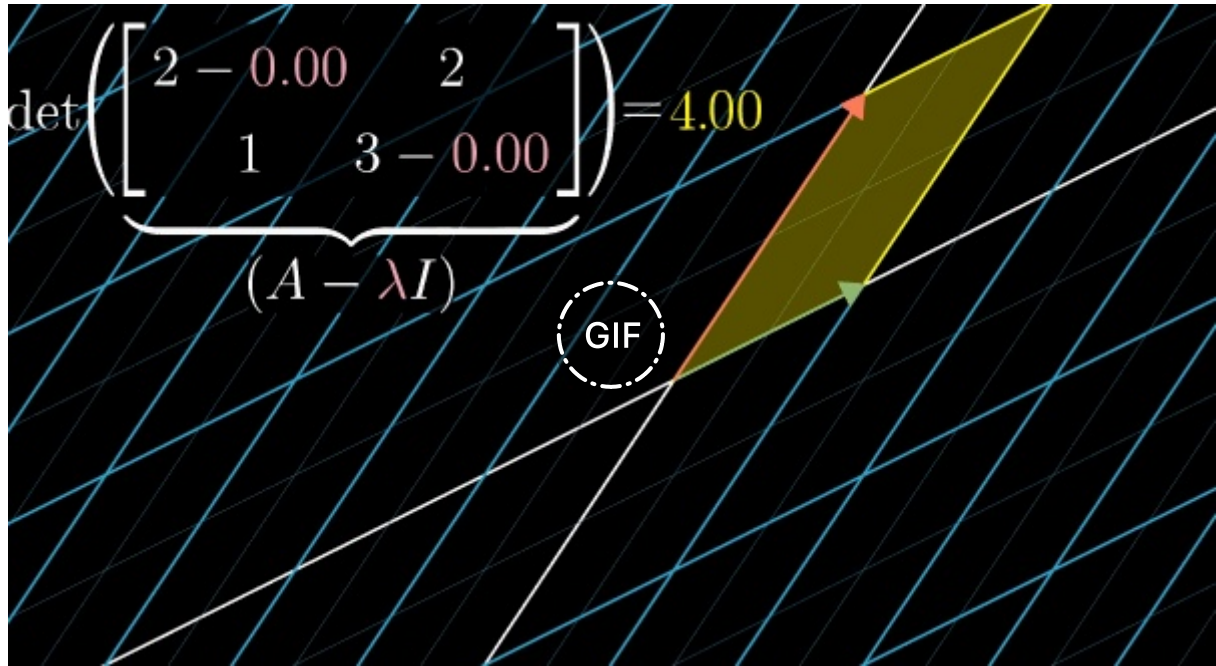
2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝

总结一下就是：

$$\begin{aligned}
 A\vec{v} &= \lambda\vec{v} \\
 A\vec{v} - \lambda I\vec{v} &= 0 \\
 (A - \lambda I)\vec{v} &= 0 \\
 \det(A - \lambda I) &= 0
 \end{aligned}$$

知乎 @初瞳

10.3.3并非所有的矩阵都存在特征向量

二维线性变换不一定有特征向量。（没有实数解表明他没有特征向量）如：逆时针旋转90度变换，向量离开了原始张成的特征空间

已赞同 20



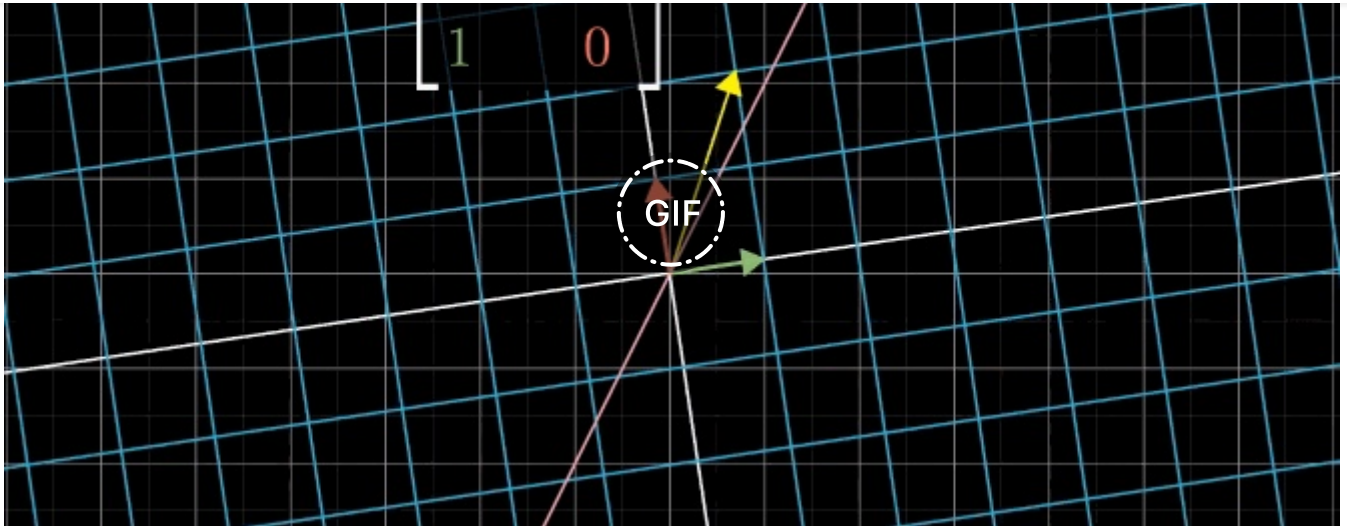
2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝

而我们尝试求解其特征值：

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

$$\lambda = i \text{ or } \lambda = -i$$

知乎 @初瞳

与 i 相乘在复平面中表示为90度旋转，这和 i 是这个二维实向量旋转变换的特征值有所关联。值得注意的一点就是，特征值出现负数的情况一般对应于变换中的某一种旋转。

10.4特征基与对角化

属于单个特征值的特征向量有可能不在直线上，如果特征向量的数量足够多，就能够张成空间。

(1)特征向量不能张成空间

已赞同 20

2 条评论

分享

★ 收藏

...

知乎

首发于
机器学习中的数学拾贝

$$\det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = \underbrace{(1-\lambda)(1-\lambda)}_{\lambda=1} = 0$$

知乎 @初瞳

可以看到几何上得出的特征值与数值计算得出的向量值是一致的。对于单个特征值，却存在无数的特征向量。但是这些向量张不成二维空间。

(2)特征向量可以张成空间

矩阵 $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ ，相当于将空间中所有的向量拉伸2倍，它仅存在唯一的特征值2，但平面内任意一个向量都属于这个特征值的特征向量。

已赞同 20



2 条评论



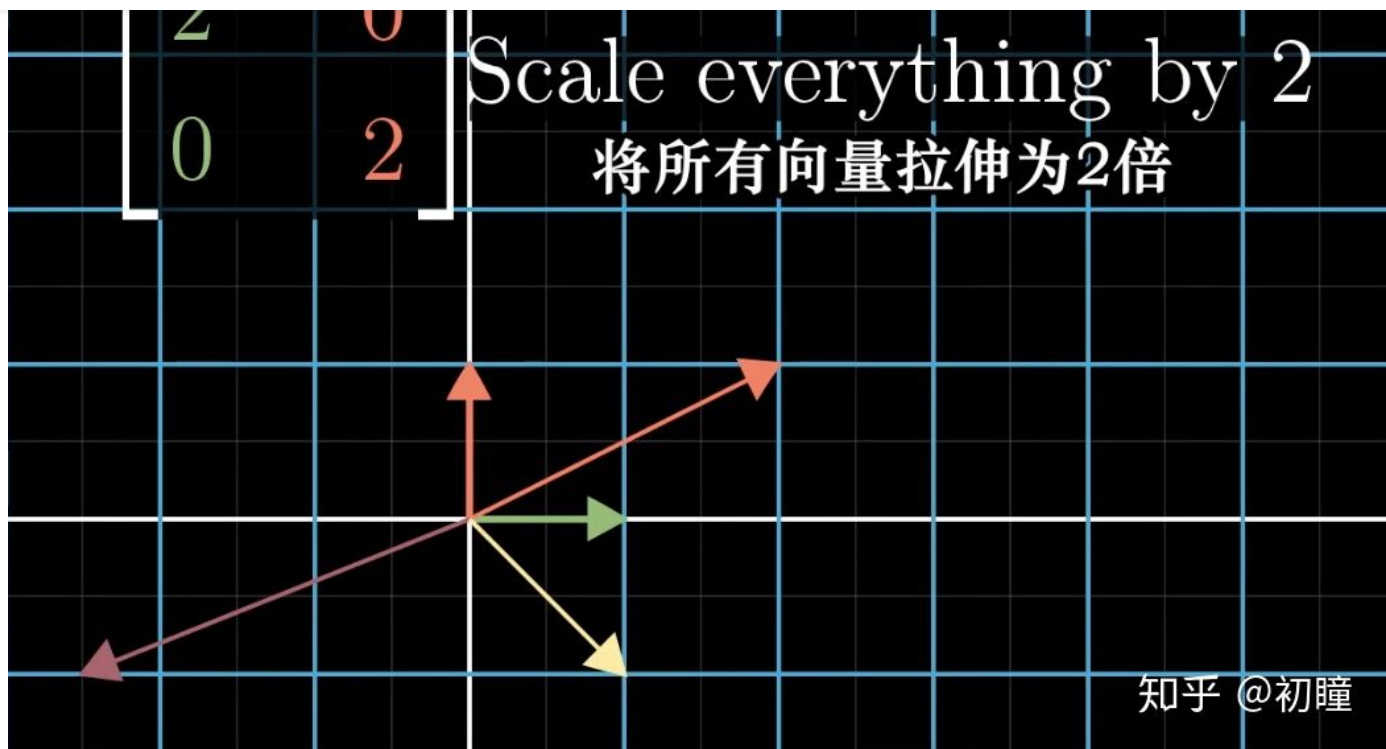
分享



收藏



知乎

首发于
机器学习中的数学拾贝

(3)对角阵

对于对角矩阵，除对角元以外其他元素均为0的矩阵被称为对角矩阵

对于对角矩阵，所有基向量都是特征向量，对角元素就是所属的特征值。

已赞同 20

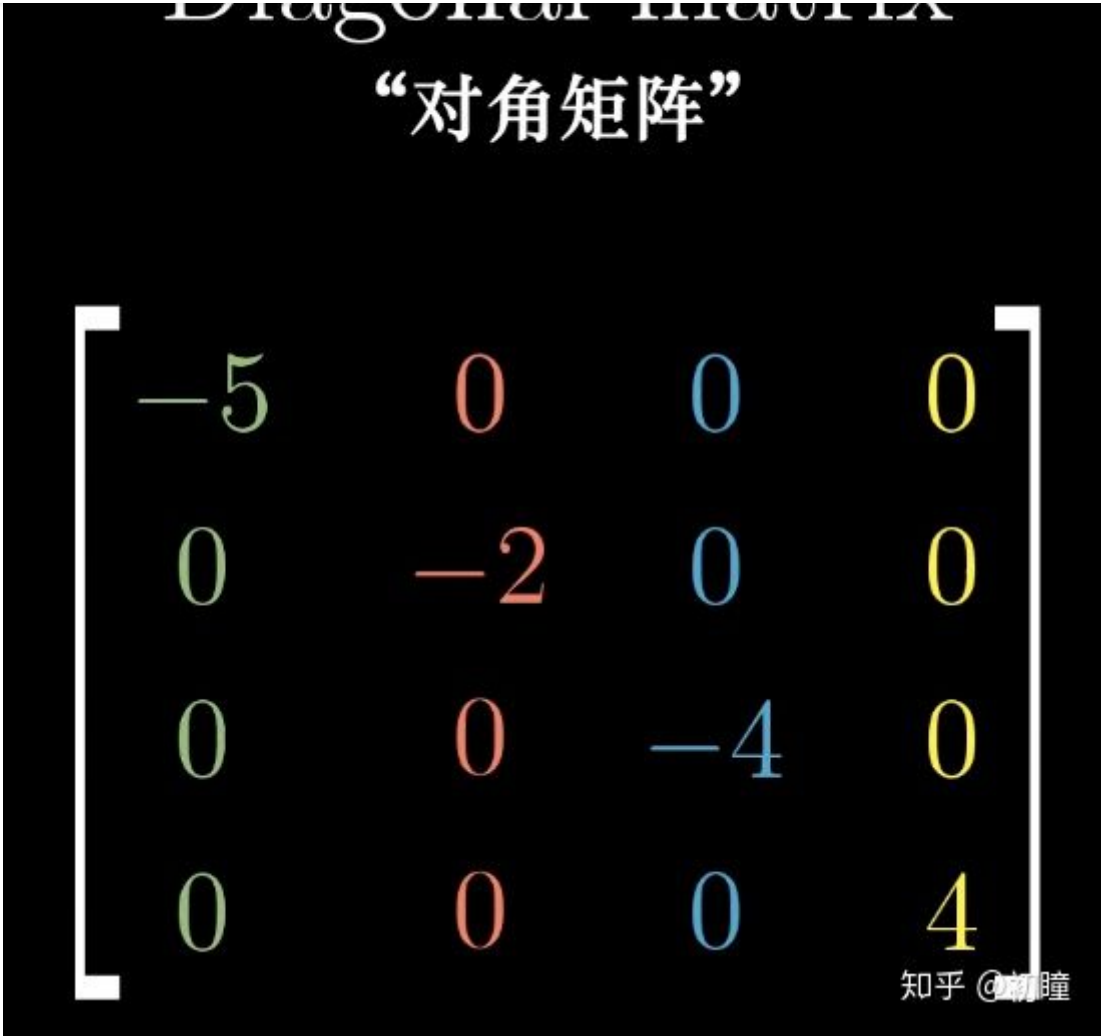


2 条评论

分享

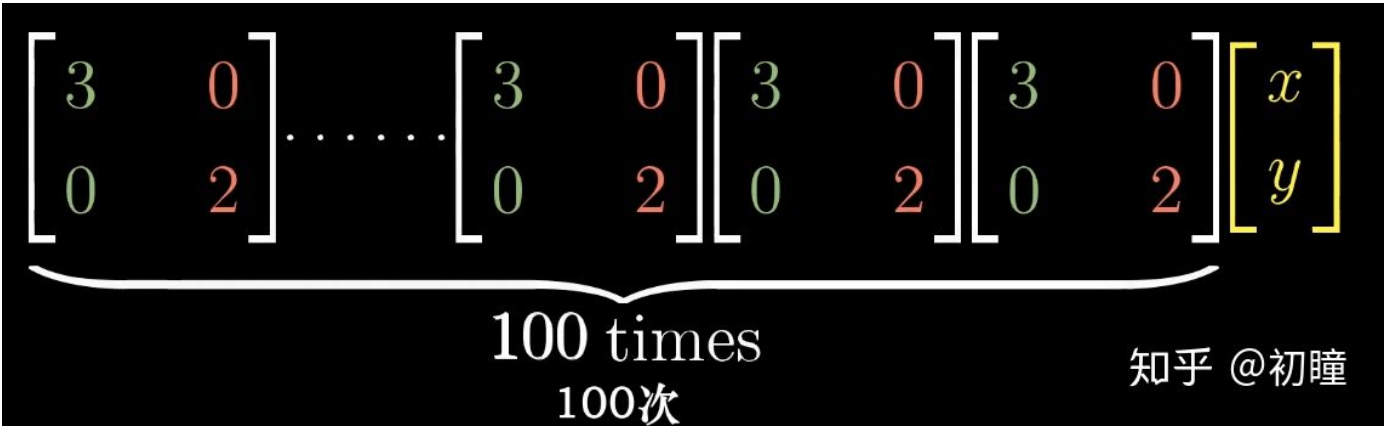
★ 收藏



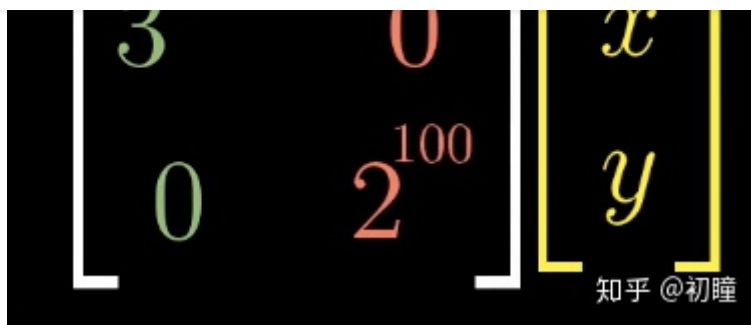


对角阵

对角阵的幂次：



知乎

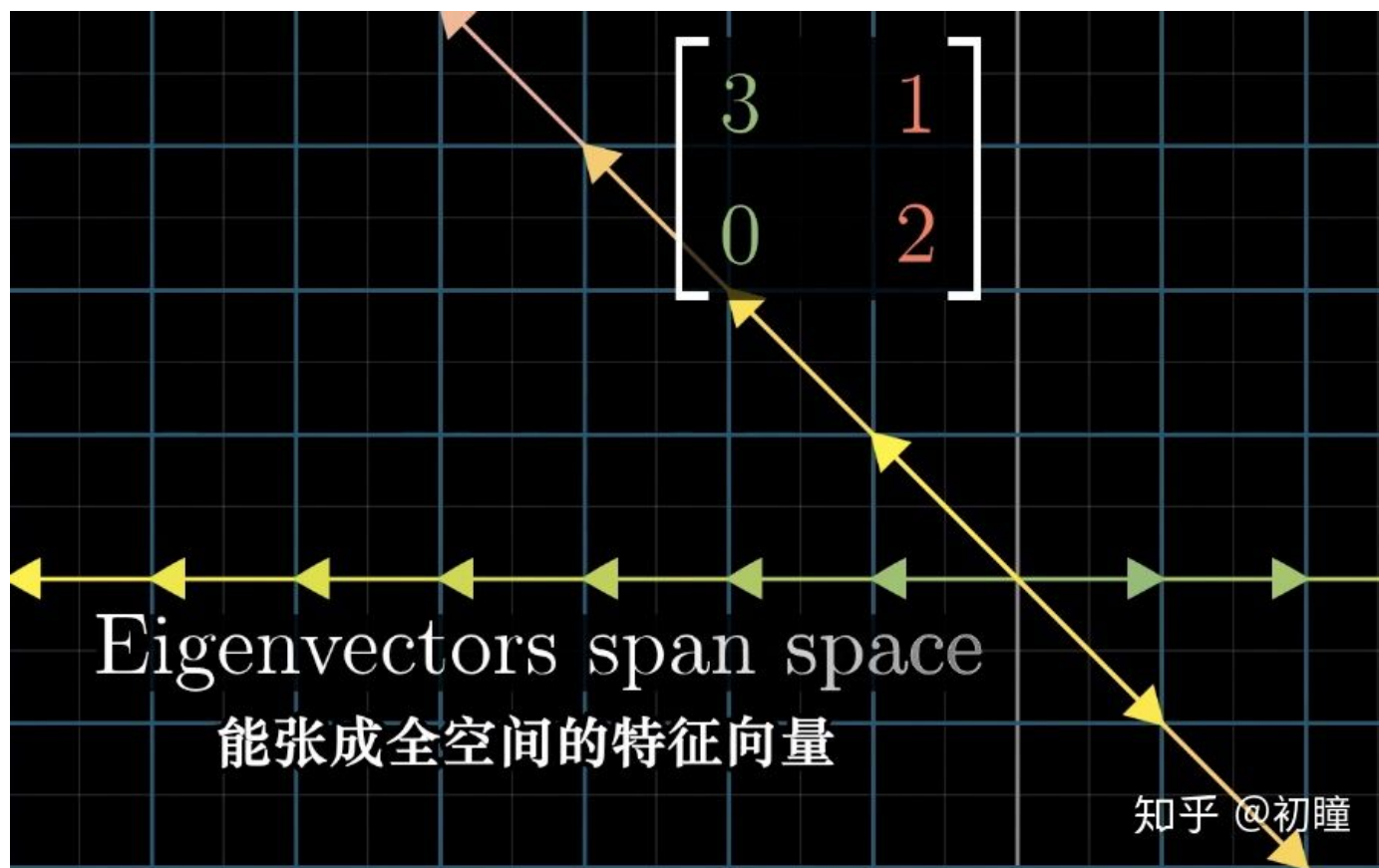
首发于
机器学习中的数学拾贝

非对角阵，计算100次，噩梦一场。

不是所有矩阵都能对角化，如剪切变换，它的特征向量不够多，不能够张成全空间。事实上，如果你能找到对应的特征基，矩阵运算就会变得特别轻松。

(4) 特征基

基向量恰巧是特征向量，将会发生什么？利用特征向量可以将矩阵进行对角化。对于有的变换有许多特征向量，多到能选出一个张成全空间的集合。



这样就能变换你的坐标系，其中这些特征向量就是基向量。

已赞同 20



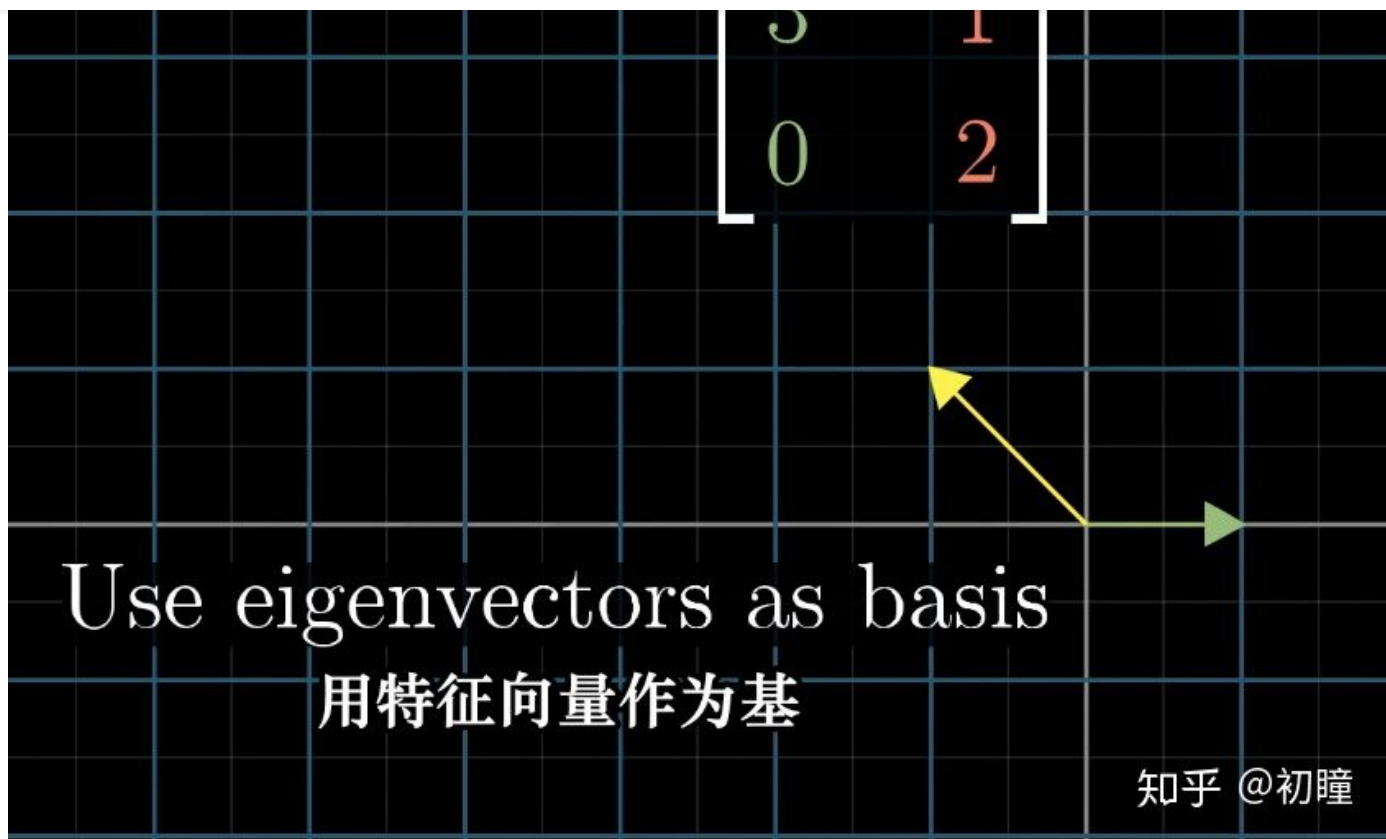
2 条评论

分享

★ 收藏

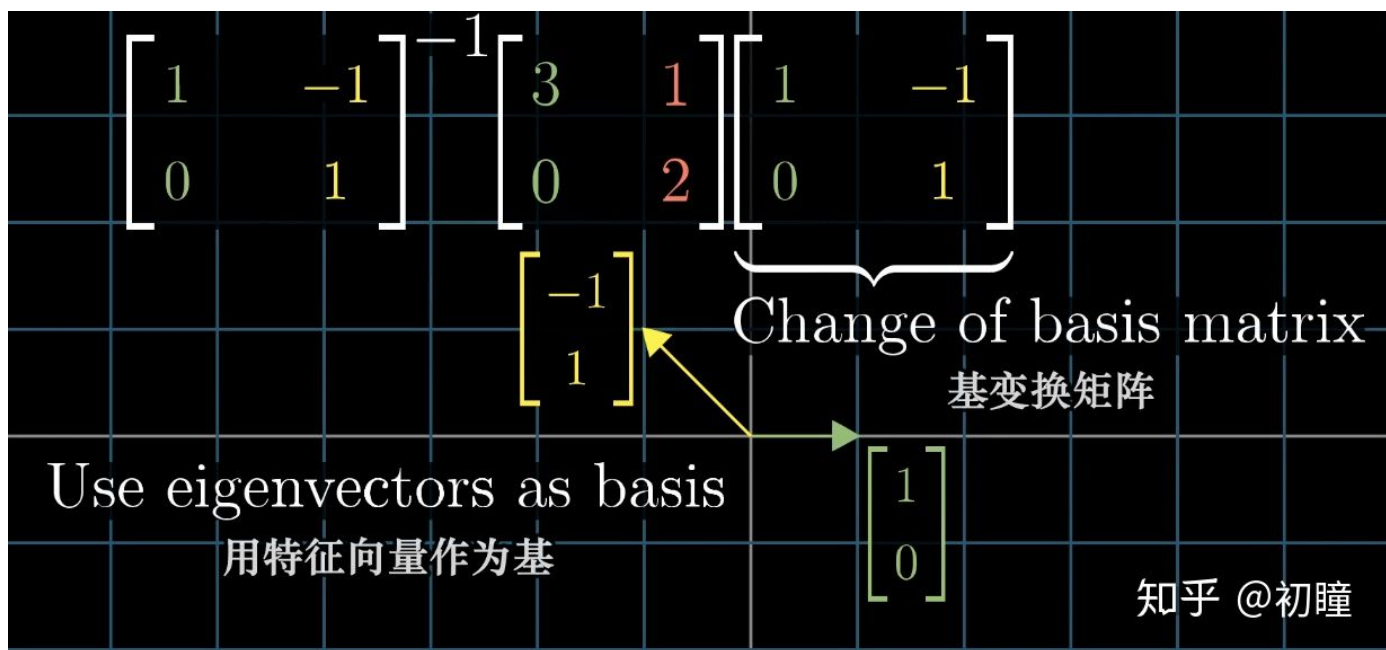


知乎

首发于
机器学习中的数学拾贝

选择特征向量构成基向量

取出你想用做新基的向量（特征向量 $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ）的坐标，然后将基向量按一定顺序构成矩阵的列，形成的这个矩阵就是“**基变换矩阵**”，



基变换

这样所得的矩阵代表同一个变换，只

已赞同 20



2 条评论

分享

★ 收藏



知乎

首发于
机器学习中的数学拾贝

对角化

用线性无关的特征向量来完成这件事情的意义在于：最终变换的矩阵必然是对角阵，且对角元就是对应的特征值。这是因为它处坐标系的基向量在变换中仅仅进行了缩放

(5) 利用特征基计算矩阵的幂次

特征值对应的所有的特征向量构成的集合被称为一个“特征基”

计算 $\begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^{100}$ ，一个更加容易的做法就是先变换到特征基，在那个坐标系计算100次，然后转换回标准坐标系。具体做法如下：

$$\begin{aligned} & \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \\ \Leftrightarrow & \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} \\ \rightarrow & \begin{bmatrix} 3 & 1 \\ 0 & 2 \end{bmatrix}^{100} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}^{100} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}^{-1} \end{aligned}$$

10.5 总结

♡ 线性变换的两种理解方式

- 将基向量变化后的位置视为矩阵的列，也就是新的基向量【依赖于坐标系】
- 利用特征向量和特征值理解线性变换，不依赖于坐标系的选择

♡ 特征值与特征向量

特征向量：一个向量经过线性变换，仍留在它所张成的空间中

特征值：描述特征向量经过线性变换后的缩放程度

♡ 求解特征向量就是寻找非零向量

已赞同 20

2 条评论

分享

★ 收藏

...

♡当特征向量的数量足够多时，这些特征向量就可以构成特征基。利用特征基可以简化矩阵的幂次计算。

为了方便理解，将总结前后copy了一份，与10.0章节的内容保持一致。
题图来源于：<https://www.ics.uci.edu/~xhx/img/courses/linearalgebra.png>
内容及图片来源：

<https://www.bilibili.com/video/av6540378/?...>
www.bilibili.com



编辑于 03-19

「真诚赞赏，手留余香」

赞赏

还没有人赞赏，快来当第一个赞赏的人吧！

线性代数 矩阵 矩阵论

文章被以下专栏收录



机器学习中的数学拾贝

关注专栏

推荐阅读

首发于
机器学习中的数学拾贝



性空间加
 $x_2 \in V$,
 $k_1 x_1 + k_2$
一笑生花

⇒ 切换为时间排序



 赞



 赞