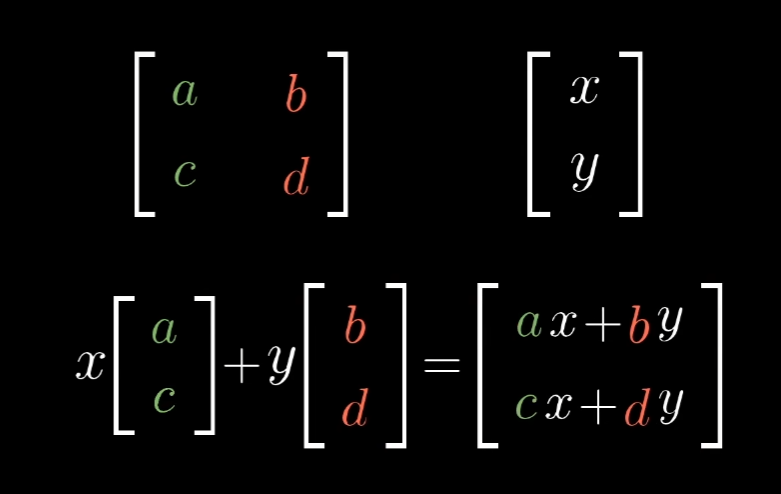
# 线性代数的本质

1. 向量究竟是什么

1．从物理角度看，向量是箭头，可移动，任意位置上的向量都和起始点在O点的向量一致；从数学角度看，向量就是列表，代表一组数。

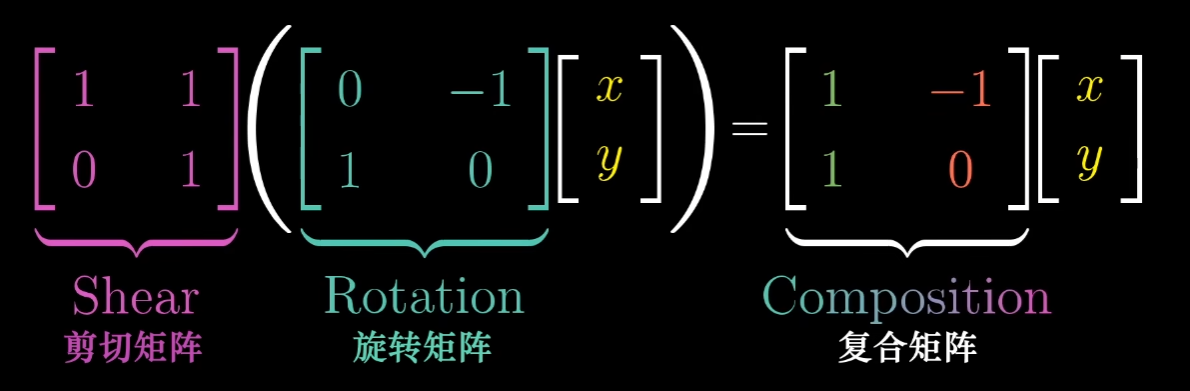
2．两种基本运算：相加和数乘

1. 线性组合、张成的空间与基
2. 在用数字描述向量时，要注意我们有基向量做基础，本质是基向量放 缩再相加。
3. 用向量的终点代表向量，这样就不用考虑箭头，把它们看成点。
4. 线性组合：一种角度是，去掉该向量不对维度产生影响，那么他们是线性相关的；另一种角度是，该向量可以被表示为其他向量的线性组合
5. 矩阵与线性变换
6. 线性变换：首先，变换和函数其实意义相近，但变换更强调运动的变化，适合描述向量。这里我们讨论其中的一种情况，线性变换。两个特征，一是原点不动，二是保持直线不曲。总的来说就是平行且等距。 一个观点是，线性变换是向量的函数。又一观点是，线性变换是对图形的变换。
7. 如何确定线性变换后的落脚点：首先，线性变换的一个很好的性质——平行且等距，让公式化成为可能：已知v是由i和j线性组合得到，那么变换后的v是由变换后的i和j得到，前面的放缩数字保持不变

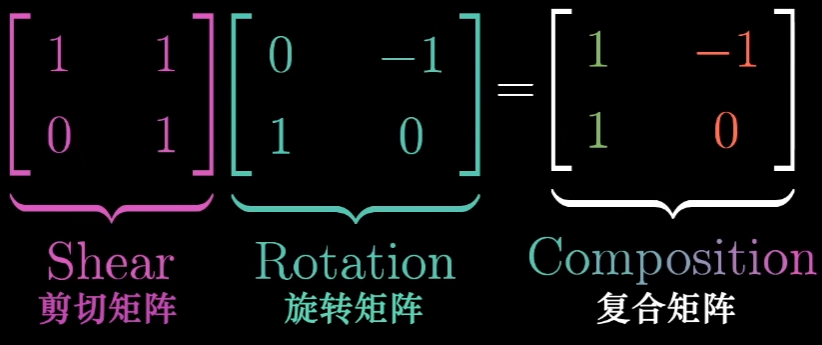


如果不作变换，即：基向量是原始的i和j，上述的变换过程同样适用。

1. 矩阵乘法与线性复合变换
2. 复合变换：字面意思，做多次变换。



进一步推广到矩阵乘法



用数值做解释就是，先追踪i的变换，再追踪j的变换（这里一定要牢记上节课对于矩阵变换的象形解释。

首先，线性变换针对的是一个向量，所以先把i给拿出来，做shear变换，再对j用同样操作，再线性变换的时候一定要注意视角的转变，即把什么矩阵看成向量。

\*矩阵乘法本质上就是复合变换

1. 矩阵变换的顺序：
   1. 交换律：顺序是会影响最后的效果的，用变换形象的思考即可明白这个道理。
   2. 结合律：不会影响最后结果，本质上都在用同一个顺序做线性变换，思考方式同上。
2. 行列式
3. 行列式是线性变换后面积是原来的x倍or（因为初始体积是1）平行六面体的体积x，x即为行列式。但就这来看，在高维上没办法直观理解。
4. 如果行列式为负，在二维空间表示平面翻转，但大小不变。在三维空间表示取向改变，标准是“右手定则指定方向是否改变”
5. 一个显然？的性质：det（M1M2）=det（M1）det（M2）
6. 行列式为0，表压缩为一条线或一个平面或一个点。。。
7. 逆矩阵、列空间与零空间
8. 逆矩阵：字面意思，线性变换的逆过程；

逆矩阵的存在前提:det(A)（行列式）不为0。—> Why？

det（A）= 0代表该线性变换降维，逆矩阵即为升维，此时有无穷多个变换，但是“NO function does this”

\*det（A）= 0时，Ax = v（x，v是向量）是否有解？

——>可能有，当v恰好在A的列空间上时有，否则没有。

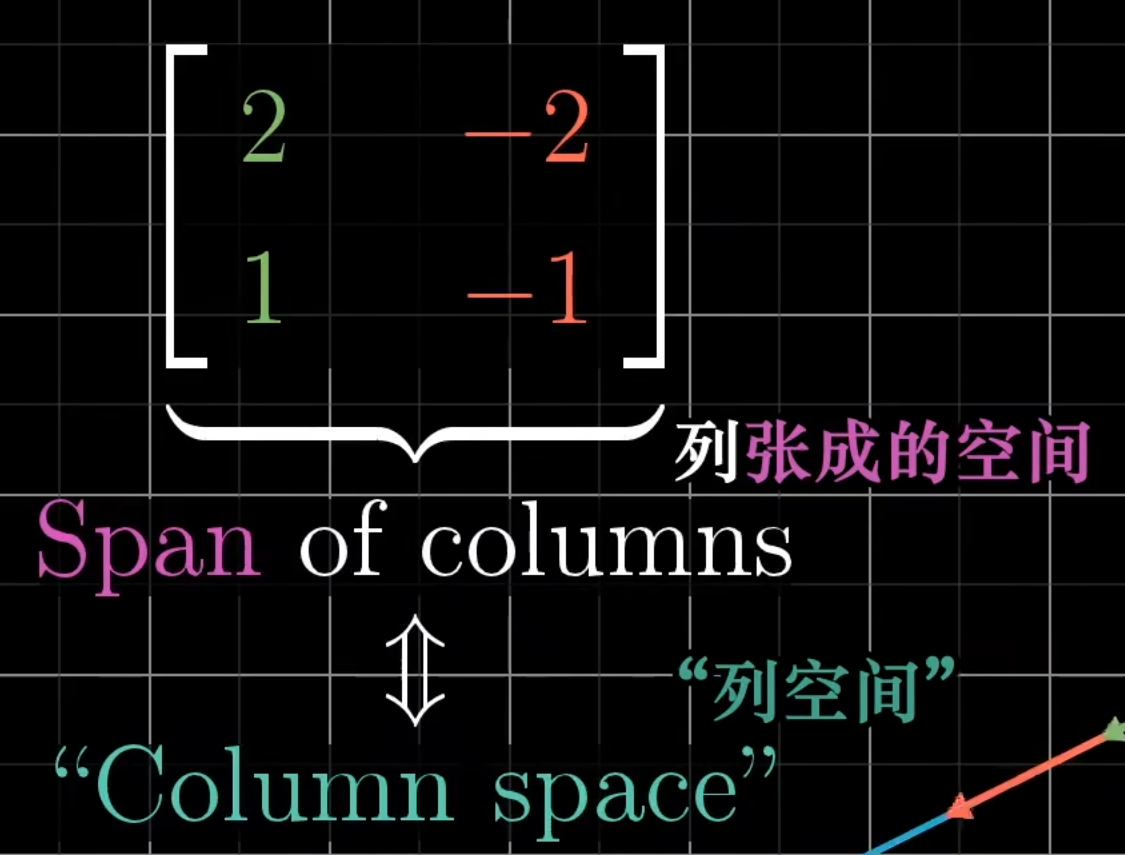
此外，把三维空间压缩成一维比压缩成二维更难有解

1. 列空间：在讲列空间之前，我们要说明一些其他概念
   1. 秩or rank：表线性变换后的维度；eg：一个2×2矩阵，

Rank = 1表压缩为一条直线，rank = 2表“满秩”，即变换后

空间维度不变。这样看来，秩用于描述矩阵

* 1. 列空间：换句话说，列空间就是矩阵的列所张成的空间。

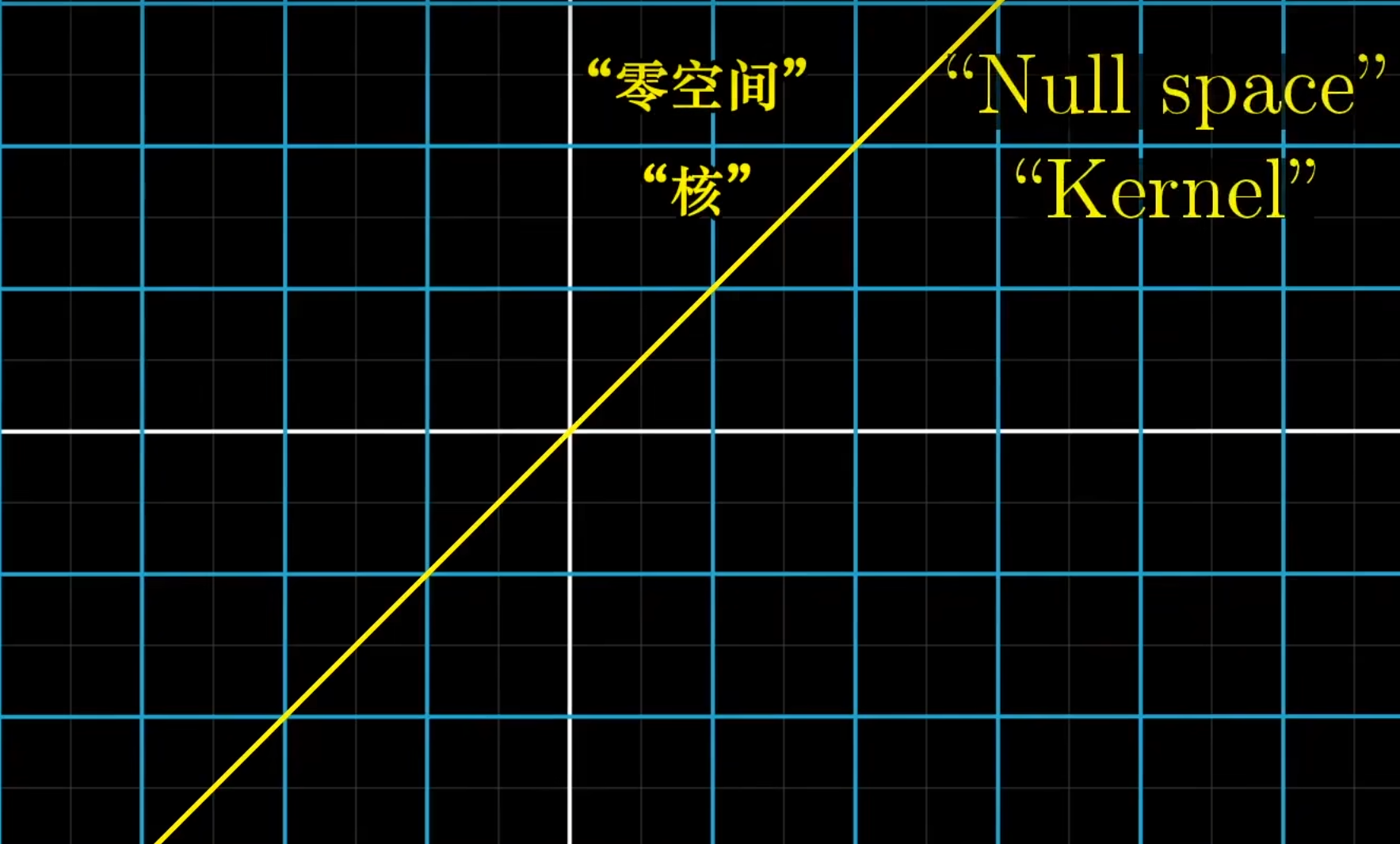


就这样看的话，列数实际上是秩的最大值。

零向量一定在列空间里，由于线性变换原点不动的性质

1. 零空间：压缩后变为零向量的向量集合

Eg：当平面压缩为直线，在平面内的某一条直线在压缩后会变成一个点.



1. 非方阵

矩阵行数代表变换后的空间维度，列数代表矩阵对象的维度（列空间）

这样看就显而易见了：

1. 行数大于列数——>升维

列数大于行数——>降维 \*注：非满秩的方阵实际上时伪降维？因为非满秩的方阵实际上是从高维度的角度看变换，z轴还存在，但非方阵是消除z轴.

1. 矩阵乘积的思考
2. 线性方程组的思考

（2. 3.详见知乎<https://www.zhihu.com/collection/713676588>）

这里简单带过： 先降维后升维再降维—>可逆不损失信息，方程组有解

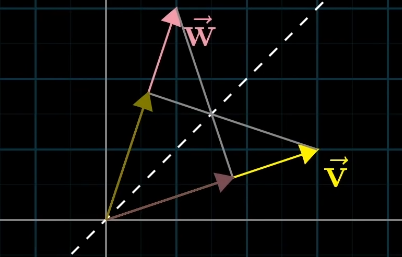
先升维后降维再升维—>不可逆损失信息，方程组无解，或者有无穷多的解.

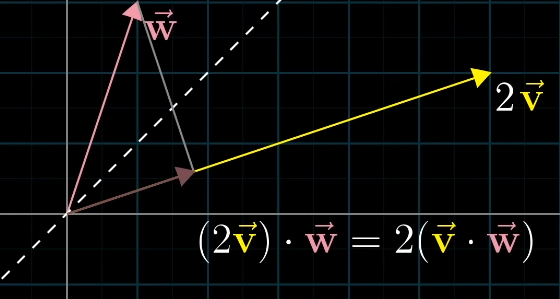
1. 点积与对偶性

点积实际上就是向量乘法：各坐标分别相乘再相加得到结果.

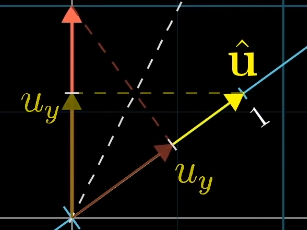
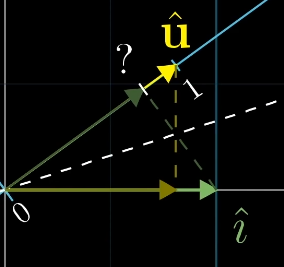
1. 向量乘法的投影顺序无关性：向量乘法的几何意义是向量投影到另一向量上并相加，方向相同为正，相反为负，垂直为0。

Why？先从两向量长度相同入手，再扩展到不同长度

长度相同时，显然顺序无关

长度不同时，对称性被破坏，但实际上这一过程可以看作是长度相同的两个向量投影后再放缩，这么看顺序就无关了。

1. 为什么点积这一运算过程，即坐标相乘后相加，会和向量投影有关系？

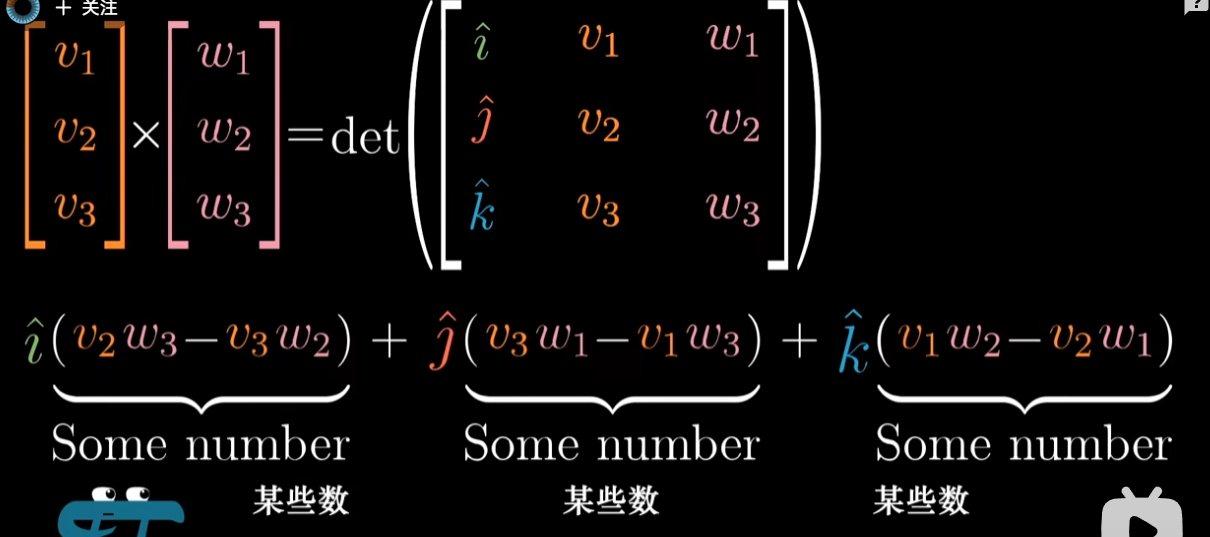
Why？ 向量投影这一过程其实是一个线性变换的过程是一个1×2矩阵线性变换。 首先是单位向量，从二维输入到一维输出的过程中，（矩阵里的数字是变换后的基向量坐标）所以先明确i和j变换后的坐标位置，这里通过对称性可以知道。

故现在就知道了矩阵是【ux，uy】

1. 叉积
2. A）标准介绍：先从向量围成的图形面积入手，叉积得到的结果 是（三维）“垂直于该平面并且长度数值为平面面积的向量，方向由右手法则确定”，二维则是面积，即叉积的结果是数，有正负。

\*在这个地方，自然的想法是，“既然二维空间里叉积的结果是面积，那么三维空间里叉积应该是体积？”，但很显然定义不是这么定义的。。。BTY，这已经比较接近了。

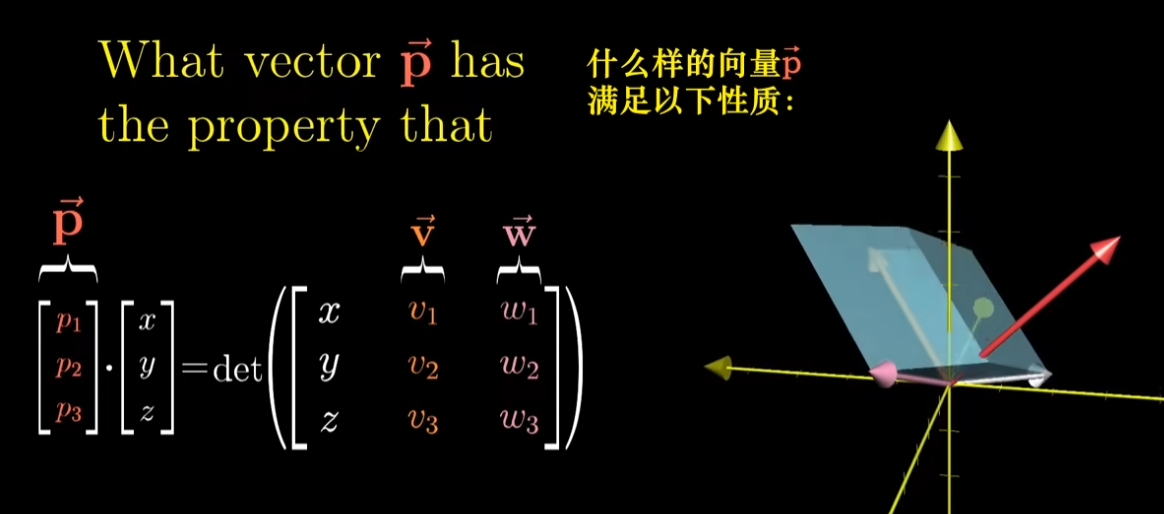
B）如何计算？ （三维）这里使用了一个3×3的矩阵，只要求出其det即可。



解释：等式右边几何意义是三个向量组成的平行六面体的体积（含正负），把v和w两个向量定死，这实际上是一个线性函数，输入一个三维向量的坐标xyz，输出其体积，这里用一个思想：“一个高维到数轴的线性变换或者说是线性函数必然有一个对偶向量”因此，计划是“1. 以v和w为基础做出一个线性变换 ; 2. 找出他的对偶向量 ；3. 证明对偶向量是垂直于平面且长度为面积的向量. ” 这样就把这个计算公式和其几何意义联系起来了，简单来说就是“把det（…）的几何意义和叉积的几何意义联系起来”det（）—> 线性函数—>对偶向量—>垂直于…向量

解释一下线性变换，该线性变换是体积，即数值上是底面积乘高，

找到一个可以实现这一结果的向量，即长度为底面积且垂直于底面，



这个向量和xyz点乘的结果恰好就等于该线性变换的结果。

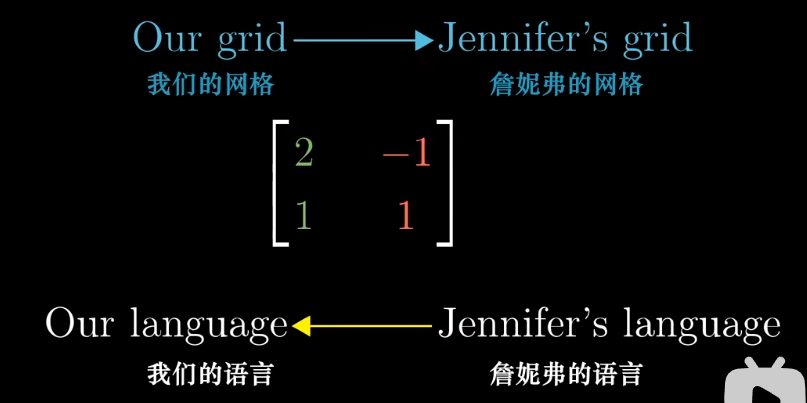
因此，我们找到了该线性变换的对偶向量，而这个向量就是叉积的几何意义所代表的向量，在此，我们就找到了该公式和叉积几何意义的联系。

十、 基变换

基变换：由变换后的基向量（由基视角描述）组成的矩阵。

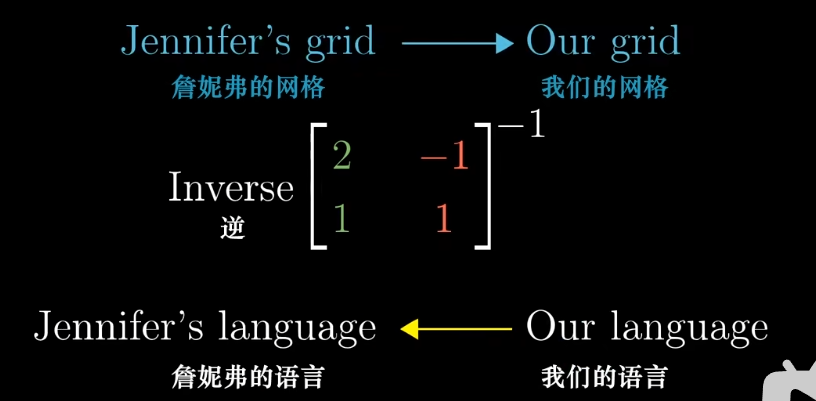
作用：将变换后视角描述的向量转换为基视角描述的向量。

1. 同一向量由不同语言描述：



更好的理解：同一个向量用不同的语言描述。

* + - 将我们的语言转变为Jennifer的语言，则用逆矩阵



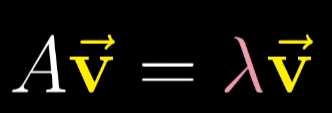
1. 同一变换用不同语言描述：

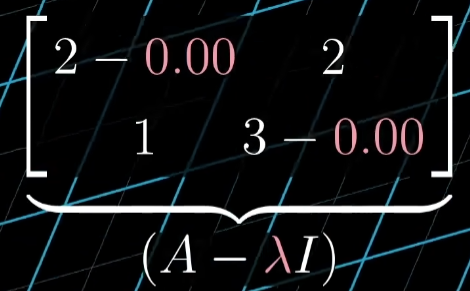
A-1·M·A：A是基变换，M是要描述的变换。当然，此时的对象是由Jennifer的语言描述的向量。

十一、 特征向量与特征值

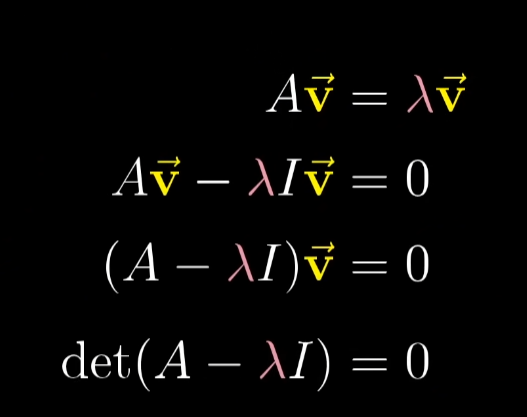
特征向量：在线性变换中不发生旋转的向量，只伸缩。

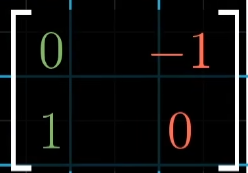
特征值：倍数，eg：二分之一，一…

符号语言：

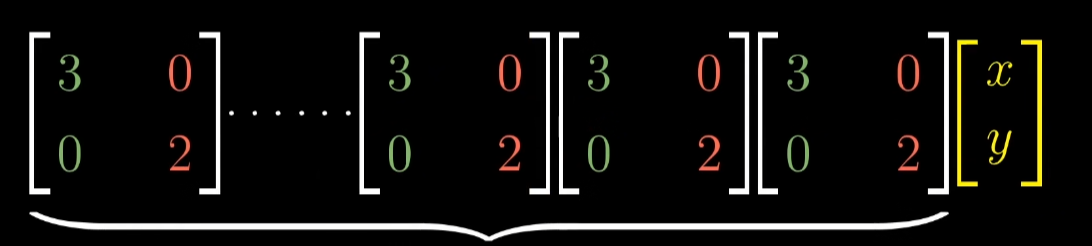
1.如何得到这个特征值和特征向量，这里省略推导过程，描述一个直观解法：当矩阵对角发生改变时，网格会逐渐压缩or拉伸，在这个过程中，特征向量所在的线会逐渐压缩or拉伸，但不会旋转，那么这就是我们要的特征向量，我们把矩阵的对角数进行变化like，逐渐变化对角（压缩空间）直到det等于零，这时的值就是特征量，在这个压缩过程中变为零向量的向量就是特征向量。

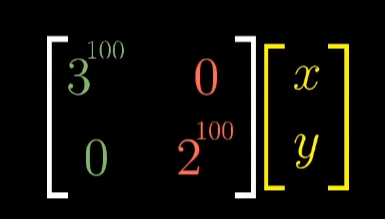
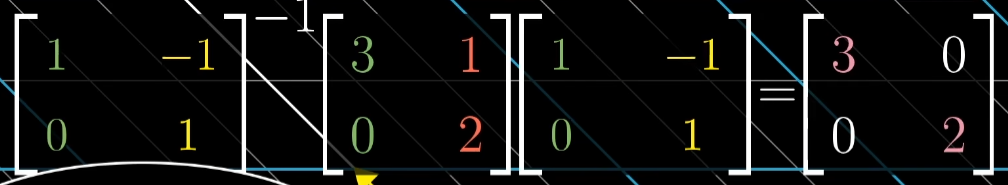
\*这里有一点就是，特征向量从基础网格线性变换后不旋转，从线性变换的结果压缩为直线也不旋转。

原因在这，特征向量定义的等价替换。

此外，纯旋转是不会有特征向量的，Like

1. 特殊的对角矩阵
2. 这种变换的特点是基向量是特征向量，表现为网格伸缩，但不旋转。也叫“特征基”
3. 对角矩阵有个很好的性质是，他的多次变换十分好算。



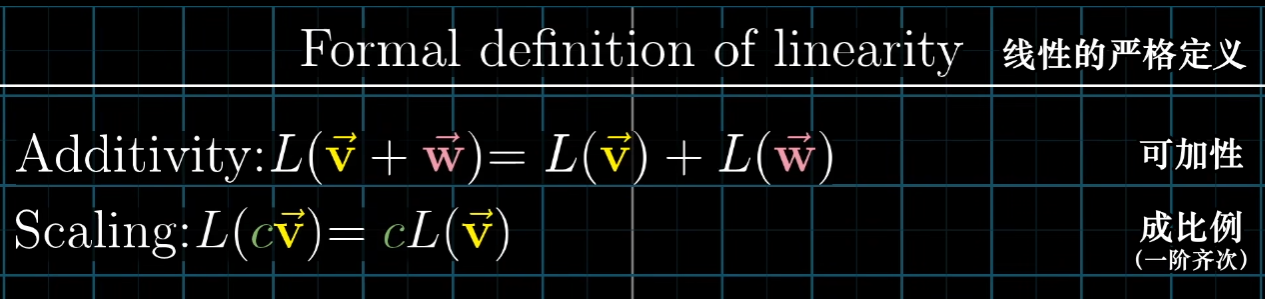
相比之下，非对角矩阵的100次幂就很难算。所以有个计算技巧：“把非对角矩阵转换为对角矩阵进行100次幂的计算，再还原回原来的矩阵。”like

2.27：新感悟，感觉自己根本没看懂；这里再解释一遍加深记忆。对于中间的这个线性变换E，他的特征向量是a（1,0）和b（-1,1），对于ab来说，这个E对他们只有拉伸的效果，因此以ab作为基向量的空间中，该变换只有拉伸效果。

十二、 抽象向量空间

1.向量可以是任何东西,比如函数。

函数满足向量的可加性和成比例性，函数的线性变换的一个典型例子是导函数，相当于放了一个函数到另一个函数里面，那么，什么是线性的呢？

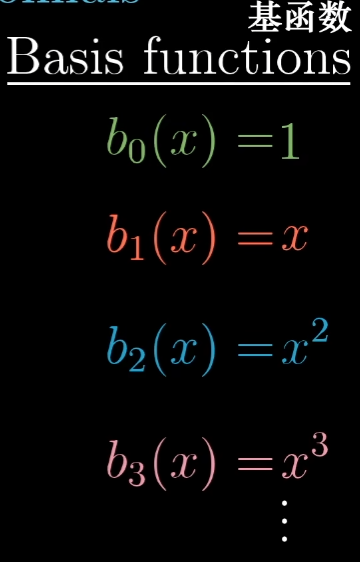
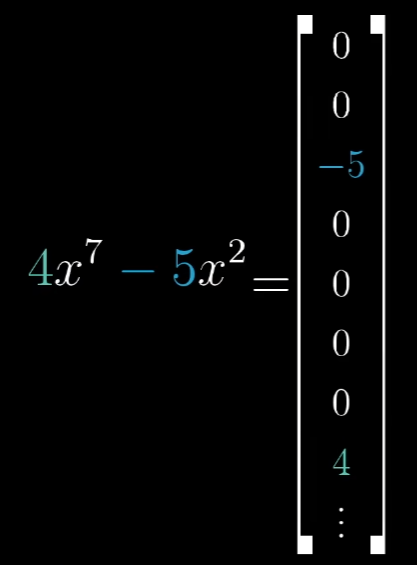


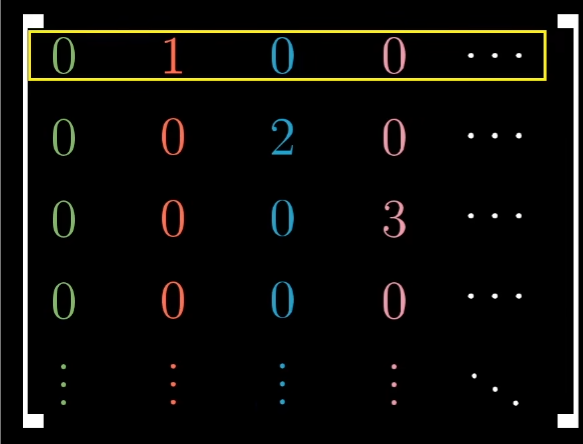
在前面的讲解中，我们提到网格线保持平行且等距是线性的，但这是具象的，尽管这只是线性这一性质在二维空间的体现，但你不能说他没有意义，因为具象往往更好理解，但不具有普适性，但抽象就具有普适性，这也是为什么教材的讲解偏向于抽象的原因。

1. 导函数的矩阵形式：在这里，我们定义一个包含全体多项式的空间，这里的基函数是，常数项，二次项，三次项….like，

b（x）=1，b（x）2 = x ， b（x）3 = x的平方…

这样的话，一个多项式就可以写成1×n的形式，导函数也可以写成矩阵的形式。

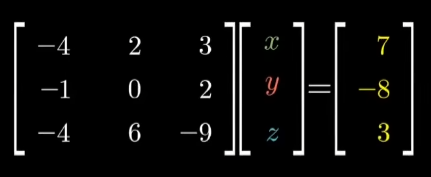
 

左图是导函数矩阵。（这里不做解释，这不是什么神奇的东西，只是这个矩阵运算恰好满足求导的规则，可以这么理解）

十三、克莱姆法则

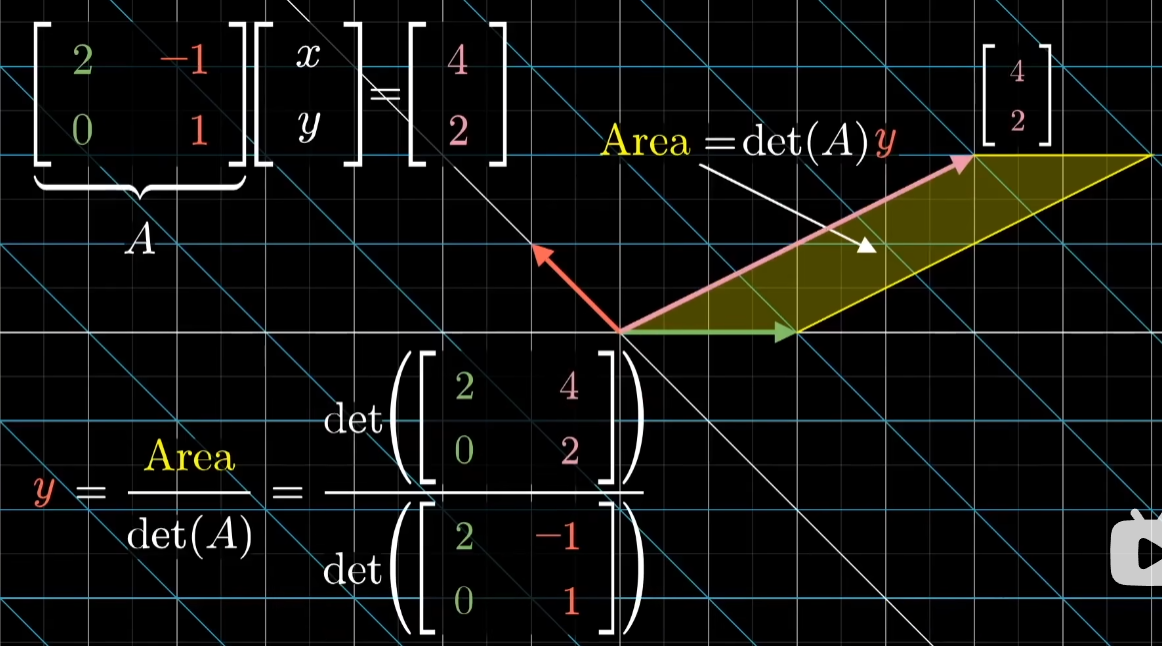
克莱姆法则介绍了求解未知数和方程数量相同的方程组，即方阵。

首先，我们要知道，求解未知数的几何意义，即“对一个未知向量进行线性变换后得到一个新向量”。

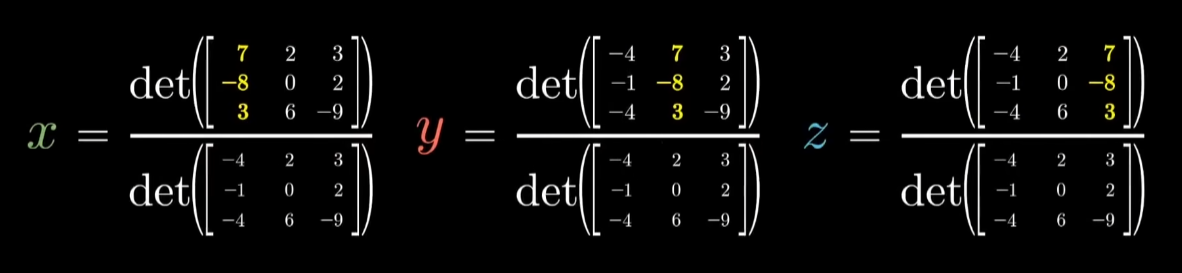


接下来，我跳过视频里引导的部分，直接说明克莱姆法则的几何意义

* + - 1. 求解xyz三个值or很多值，你要知道，这些量其实毫不相关，他们是相互独立的，因此，你只要会求解其中的一个量，那其他量也是同一种方法。
      2. 那么现在我们来试着解x，从二维角度看，我们找一组向量，一组是，另一组是（0，1），那么这两个向量组成的矩阵的面积就是x；线性变换后的面积很显然是x·det（A），也可以是由线性变换后的作为底边的向量和变换后的向量组成的矩阵的det。（这里有点绕，如图）



三维同理：



# MIT线代

## 第一节课

* + - 1. 列图像和行图像：前者是基向量线性组合描述，后者是求每个方程代表的图形所形成的交点。
      2. 描述Ax的方法：线性组合（btw用线性变换的视角看也不是不行，但是在这里我猜测做题肯定还是前者思考角度多一些）

## 第二节课

矩阵消元：让矩阵变成类似于阶梯型的矩阵。消元看行向量

矩阵乘法：承接上一节课的第二点，线性组合思考矩阵的计算效率会快一些，“取x个向量a，再取y个向量b”用列向量或行向量都可以。

左乘和右乘：再《线性代数的本质》系列里，作者一直在传递从右向左看矩阵的思路，并且把矩阵看作列向量的组成，但今天我们知道了第二种处理方法，行向量是从左向右看，显然的是，左乘和右乘是有相对性的，以哪个矩阵作为第一种变换显得格外重要。这里比较有趣的地方是，同一种变换like AB，你用两种视角看他的结果是相同的，这一点我其实还不能理解，这也肯定不是巧合。

## 第三节课

1. 如何得到一个矩阵里有具体坐标的值。首先，记住一些技巧，左行右列，依次乘下去后相加得到该点的数值。这也是为什么两个矩阵相乘的时候需要满足行和列的数相同的原因，第二个技巧是，得到的矩阵的行数是左矩阵的，列数是右矩阵的，“左行右列”。
2. 如何得到矩阵的一列？只要用到矩阵的几何意义，即向量线性组合。

行向量同理，但要注意是右乘。