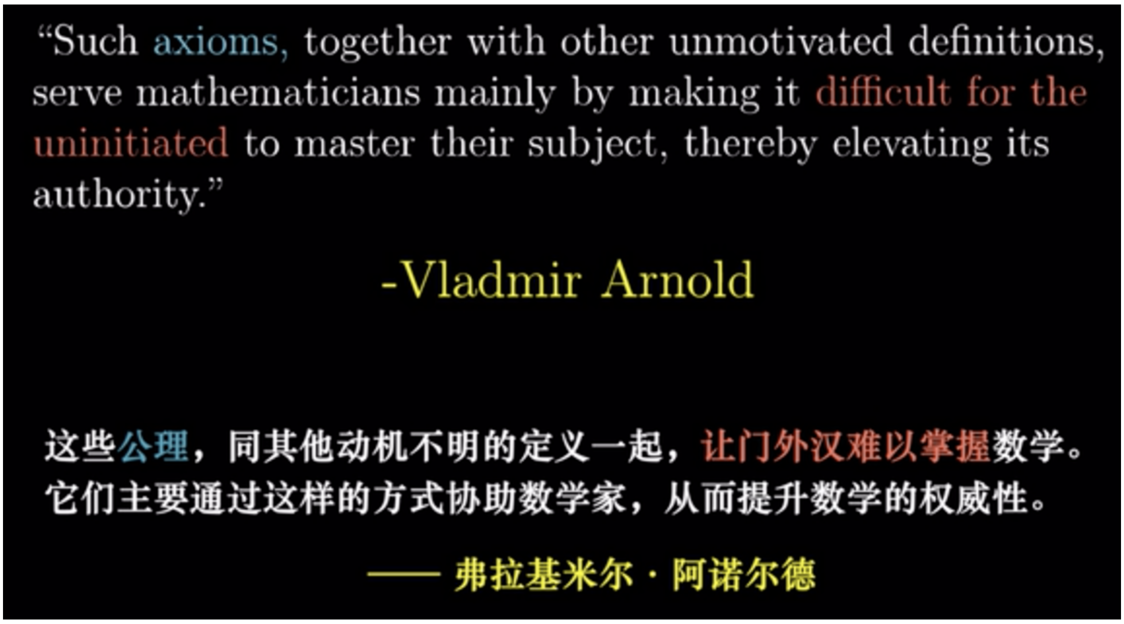
13.抽象的线性空间



什么是向量？ 比如说，它是一个二维向量。从根本上说，它是平面内的一个箭头，而我们可以用坐标来描述它。

或者说，它是一个实数对？而我们只是将它形象理解为平面内的一个箭头

又或者说，这两种观点只是更深层次的东西的体现？

一方面，将向量解释为一组数字给人感觉清晰明了。四维向量，乃至一百维向量看上去就像是可以操作的真实具体的概念；另一方面，四维空间之类的只是一个模糊的几何概念，难以解释。

但是另一方面，对于那些在实践中运用线性代数的人，尤其是熟悉基变换的人来说，他们通常感觉所处理的空间独立于坐标存在，而且坐标描述实际上有些随意，它依赖于你所选的基向量。

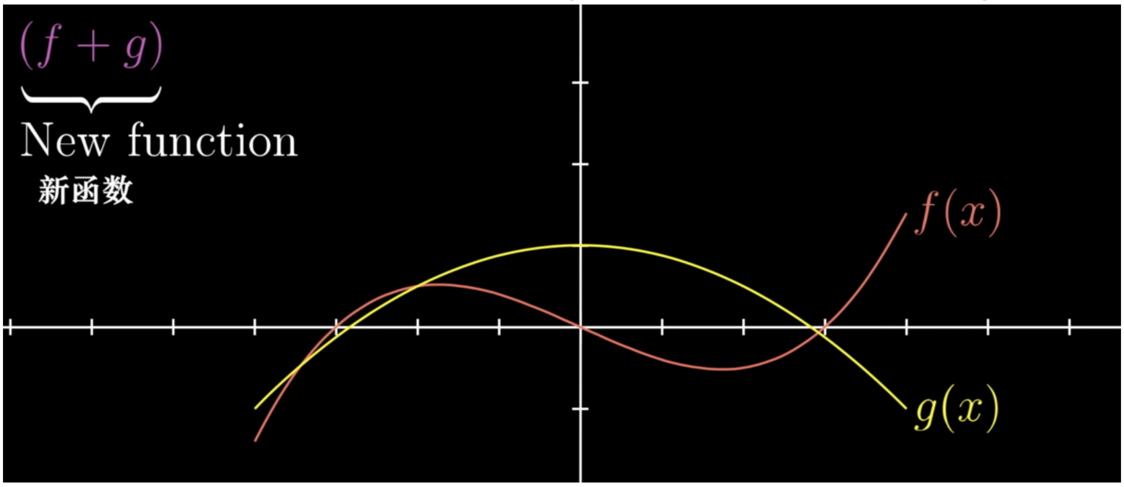
线性代数中的核心话题，如行列式和特征向量等，它们似乎不受所选坐标系的影响。行列式告诉你的是一个变换对面积的缩放比例；特征向量则是在变换中留在它所张成的空间中的向量。这两者都是暗含于空间中的性质，你可以自由选择坐标系，这并不会改变它们最根本的值。

但是，如果向量根本上并不是由一组实数构成，或者它们的本质其实与空间并不是那么相关。这不禁让人产生疑问：数学家所说的"space"或'spatial'是什么意思？

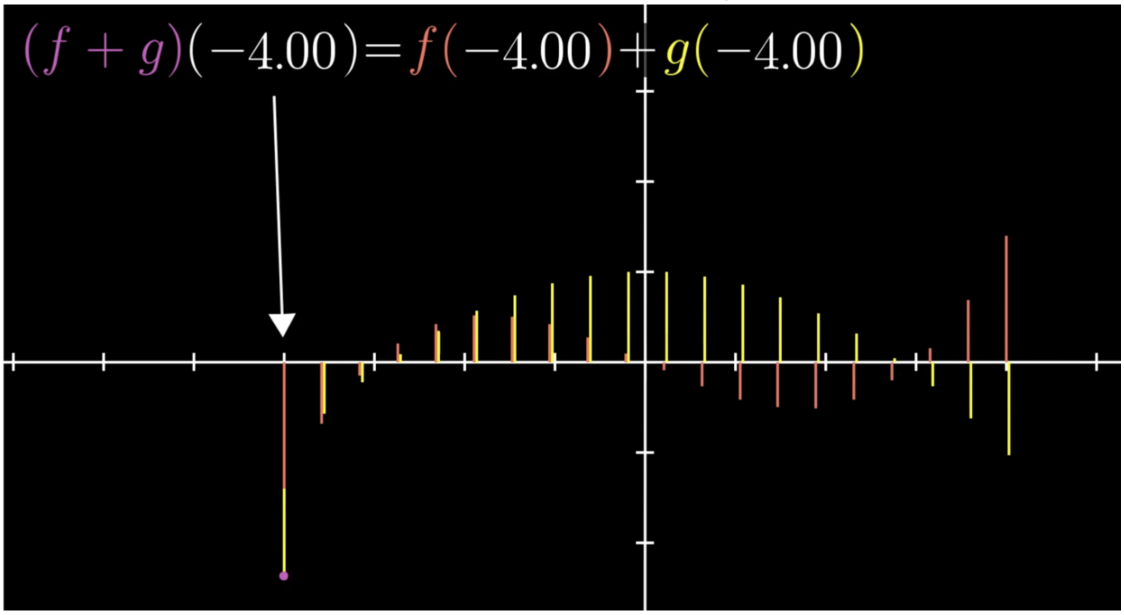
为了进一步说明这是怎么回事，我们来讨论一种即不是一个箭头，也不是一组数字，但是同样具有向量特性（vector-ish qualities）的东西——函数。

从某种意义上说，函数实际上只是另一种向量。

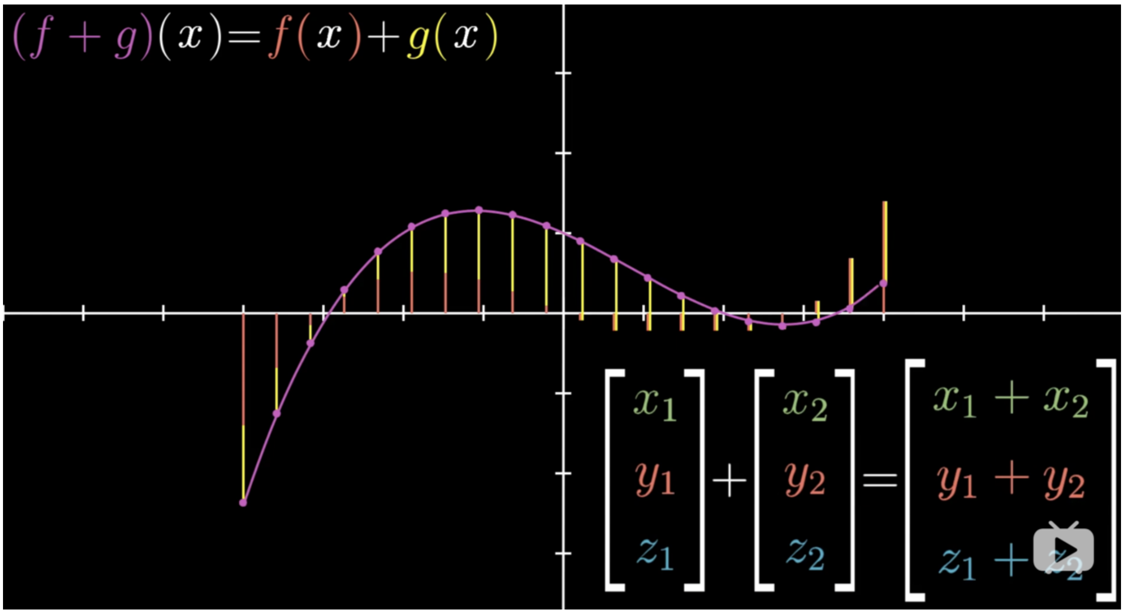
与向量加法类似，我们可以将两个函数f和g相加，从而获得一个新函数f+g。



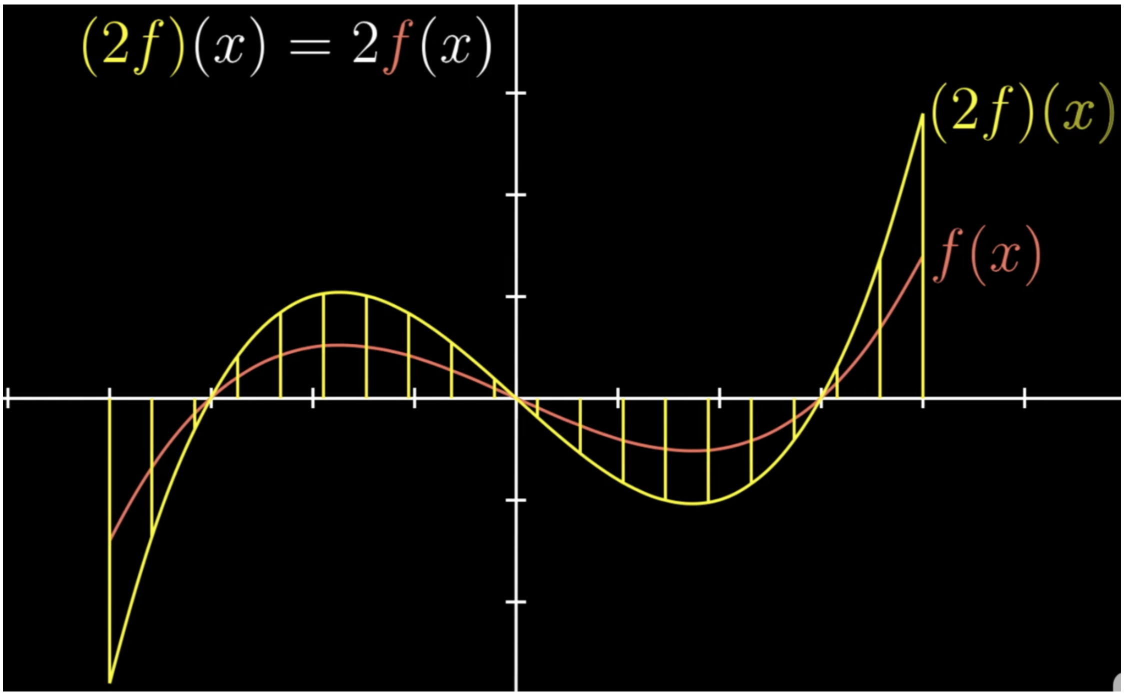
这个新函数在任意一点处的值（比如-4处），就是f和g在这一点的值的和



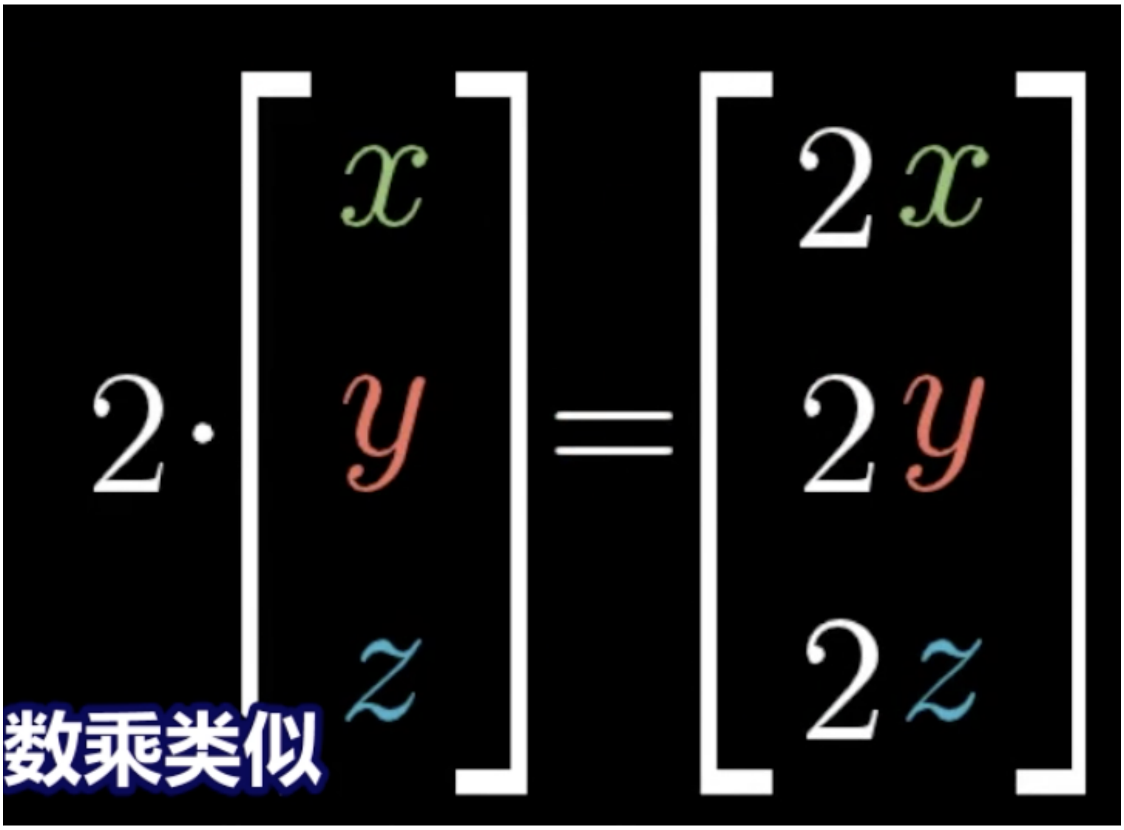
确切的说，这个函数在任意一点x处的值等于f函数和g函数在x处的值的和，这和向量对应坐标相加非常类似。



类似的，函数与一个实数相乘也有合理的解释：

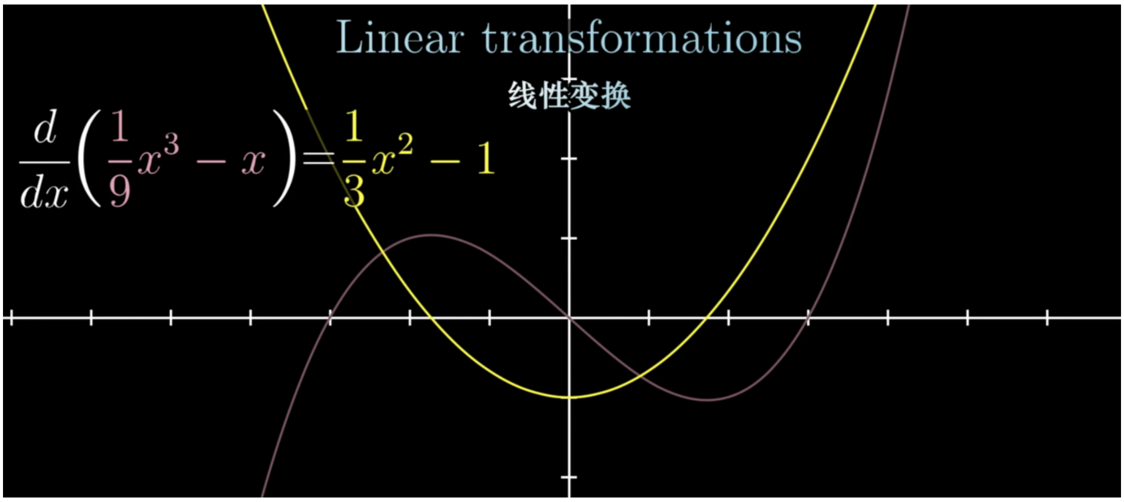


这再次和向量对应坐标数相乘类似



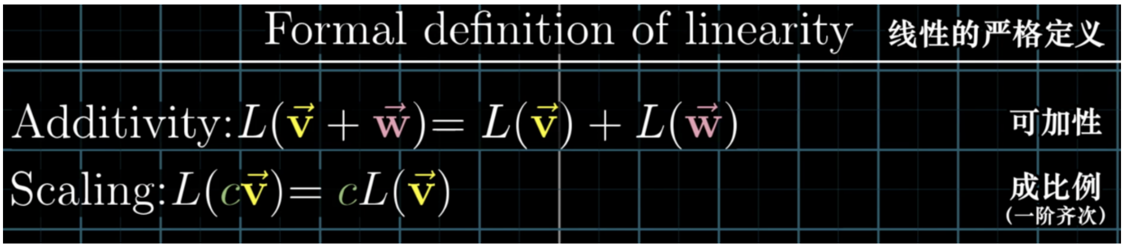
因为对向量所能做的操作不过相加和数乘两种，所以，最初以空间中的箭头为背景考虑的线性代数的合理概念和解决问题的手段能被我们取出来，然后应用于函数。

例如，函数的线性变换有一个合理的解释：这种函数接收一个函数，并把它变成另一个函数。从微积分中，我们可以找到一个常见的例子——导数。它将一个函数变为另一个函数：



那么，一个函数变换是线性的是什么意思呢？

我们需要看一下线性的严格定义（相对抽象而且符号繁重）。但是，抽象性带来的好处是我们能得到一般性（general ）的结论，它不仅适用于箭头，也适用于函数。满足一下两条性质的变换是线性的：

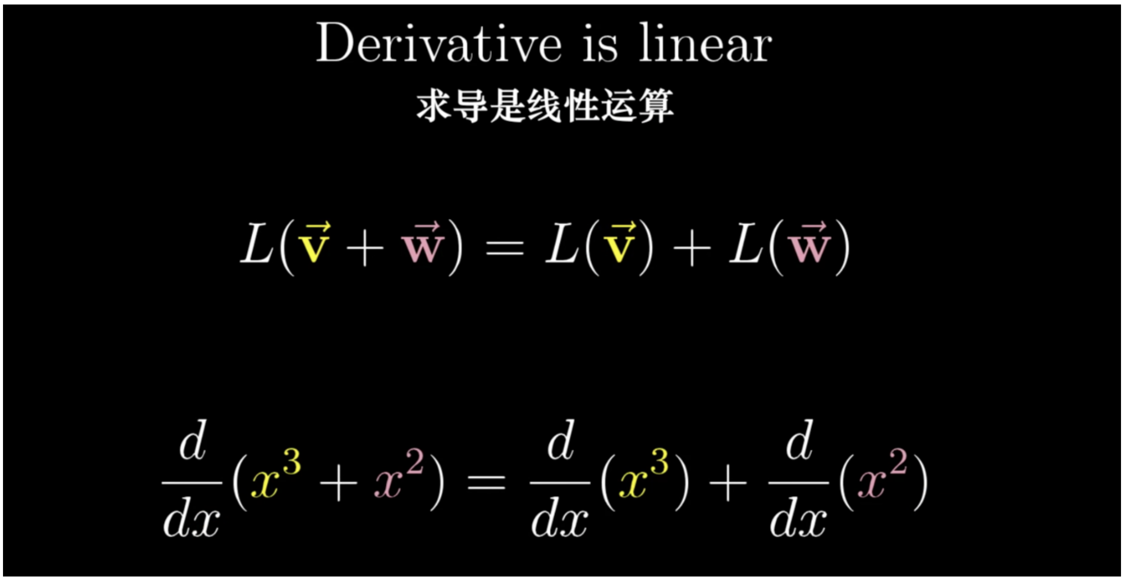


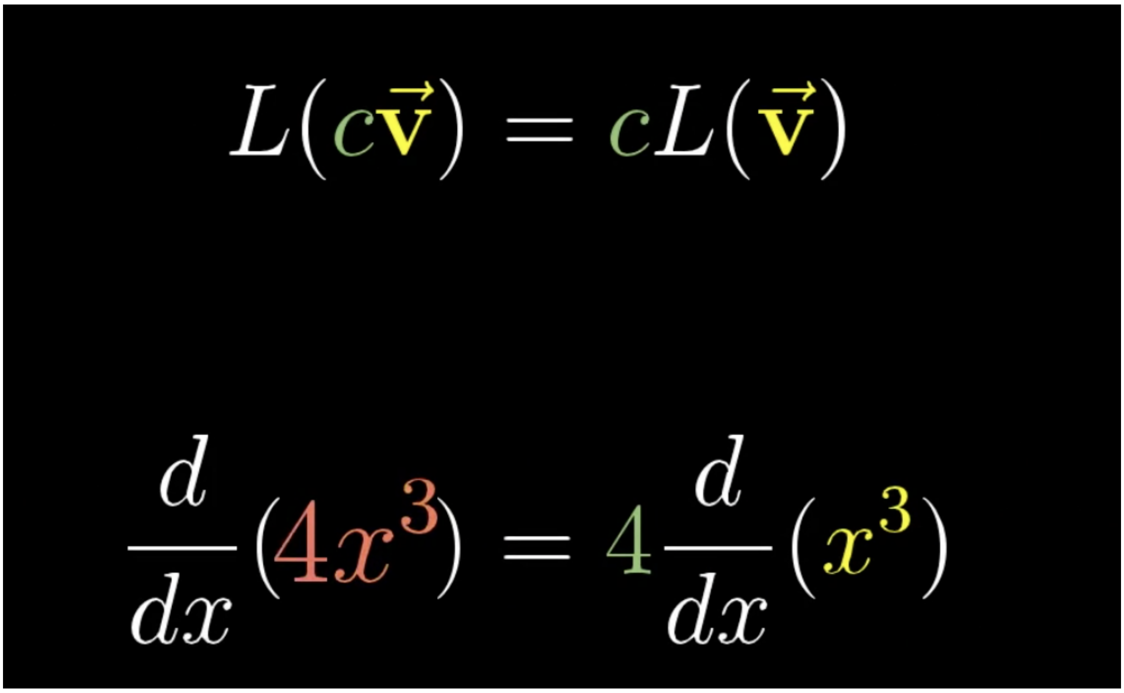
- 可加性意味着如果你把两个向量相加，然后对它们的和应用变换，得到的结果和将变换后的v与w相加一致。
- 成比例是说，一个向量与某个数相乘然后应用变换，得到的结果和变换后的v与这个数相乘一致。

你经常会听到的一种描述方法是线性变换保持向量加法运算和数乘运算。我们以前讨论过的“网格线保持平行且等距分布”的概念，只是这两条性质在二维空间这一特殊情况下的体现。

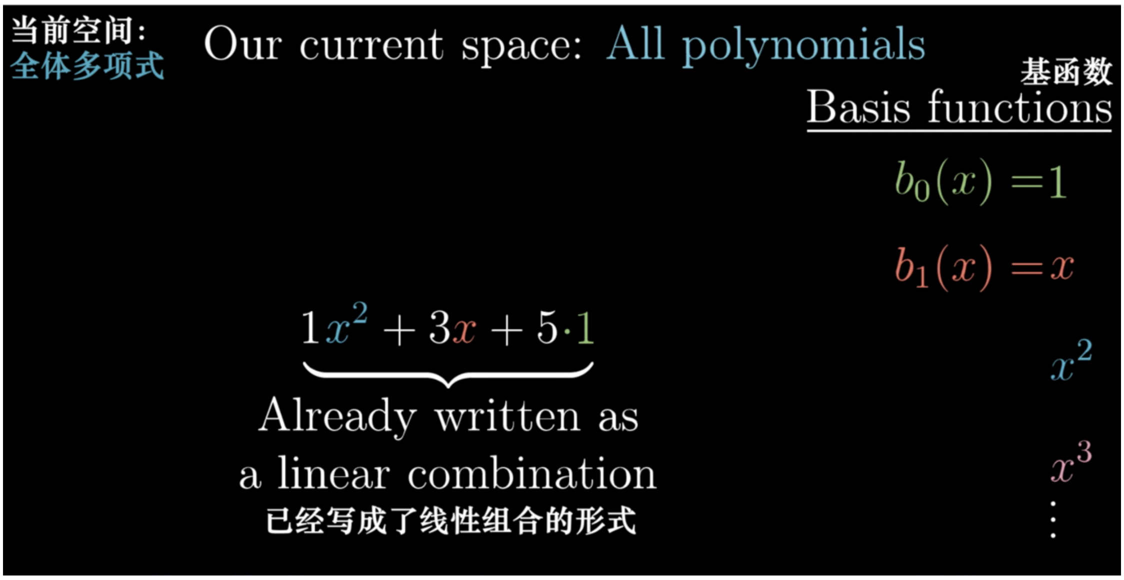
这两条性质的一个重要推论是：一个线性变换可以通过它对基向量的作用来完全描述，这使得矩阵向量乘法称为可能。因为任一向量都能表达为基向量以某种方式进行线性组合，所以求一个向量变换后的结果，实际上就是求出变换后的基向量以相同方式进行线性组合的结果。

这一点对函数类说同样正确。比如说，学微积分的学生经常会用到一个事实即求导具有可加性和成比例性：





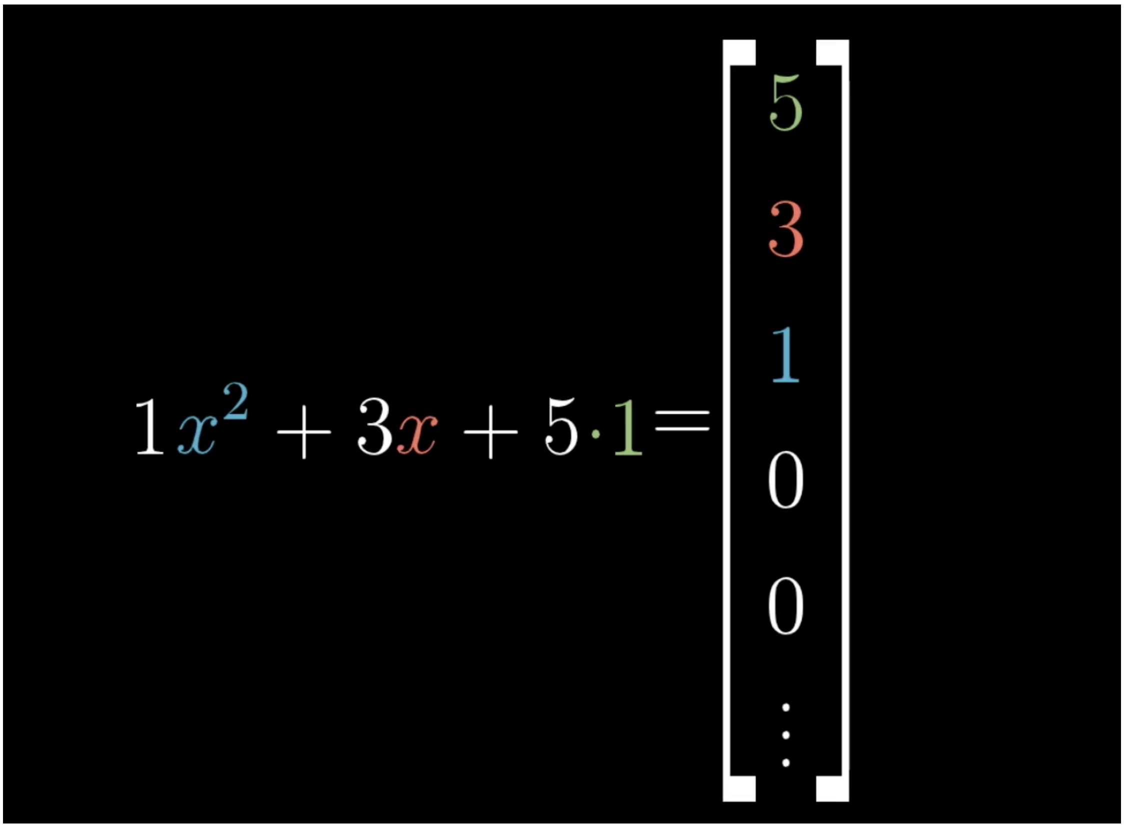
为了真正掌握这里的类比关系，我们来看看用矩阵描述求导是什么样的。我们把眼光限制在多项式空间上。首先要做的就是给这个空间赋予坐标的含义，这需要选取一个基。因为多项式已经写出了线性组合的形式，所以我们很自然地选取x的不同次幂作为基函数：



所以，第一个基函数:$b_0(x) = 1$,第二个基函数:$b_1(x) = x$，第三个基函数:$b_2(x) = x^2$。以此类推。

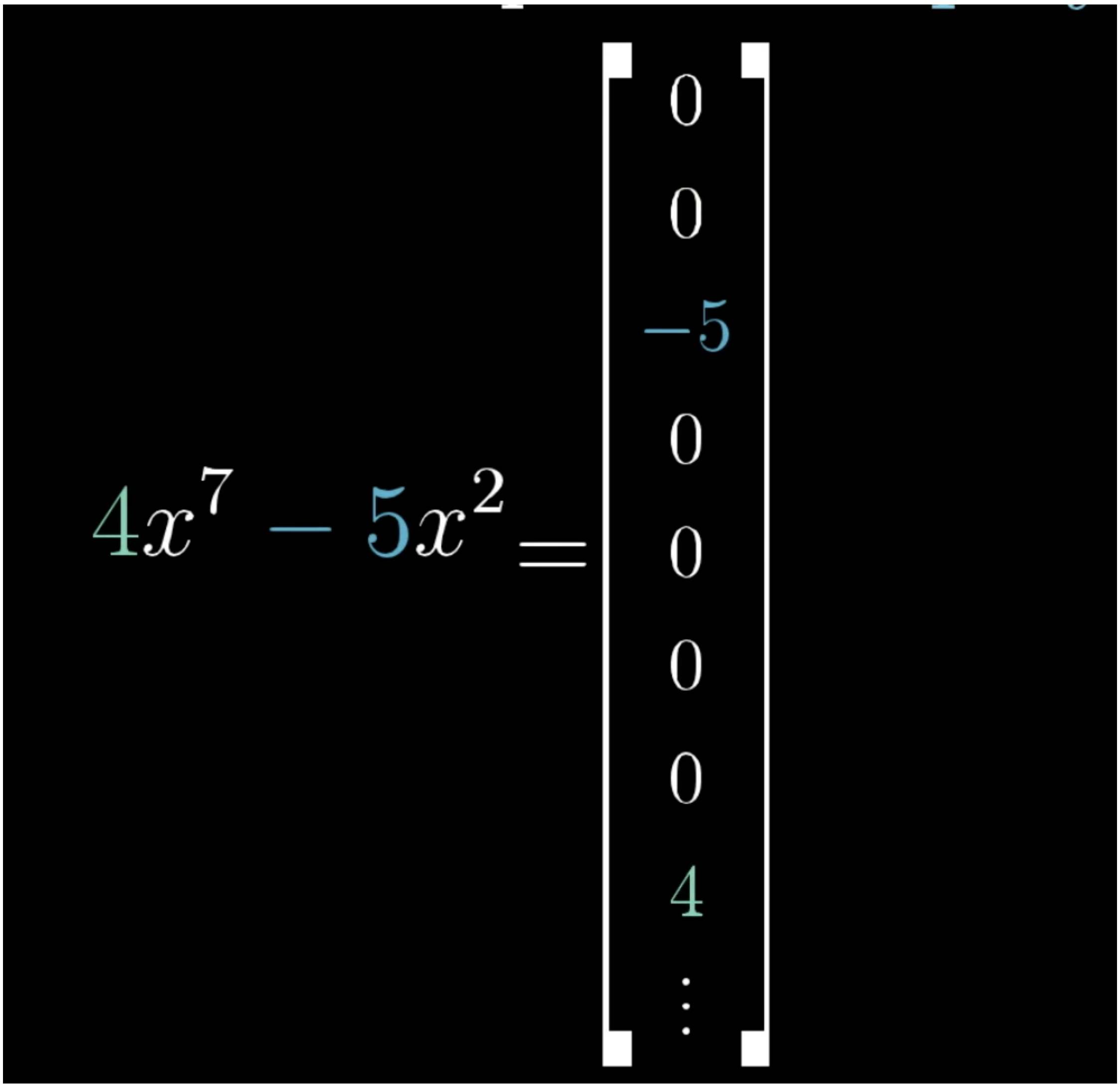
基函数在这里起到的作用和i帽，j帽在向量的世界中起到的作用类似。因为多项式的次数可以是任意高，所以这个基函数基也是无穷大的。不过没有关系，这只是说明我们把多项式当向量来处理时，它们会有无穷多个坐标。

比如描述$1x^2+3x+5\cdot 1$ 用坐标来描述的话就是5，3，1，然后更少无穷多个0：



可以这样理解它：5乘以第一个基函数，加上3乘以第二个基函数，加上1乘以第三个基函数，在此之后，其他基函数不再出现。

再看下一个例子：



我们接下来看以下用矩阵向量乘法来求导：

$$\frac{d}{dx}(1x^3+5x^2+4x+5)=\underbrace{3x^2+10x+4}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}\begin{bmatrix} 5 \\ 4 \\ 5 \\ 1 \\ \vdots \end{bmatrix}=\begin{bmatrix} 1\cdot 4 \\ 2\cdot 5 \\ 3\cdot 1 \\ 0 \\ \vdots \end{bmatrix}$$

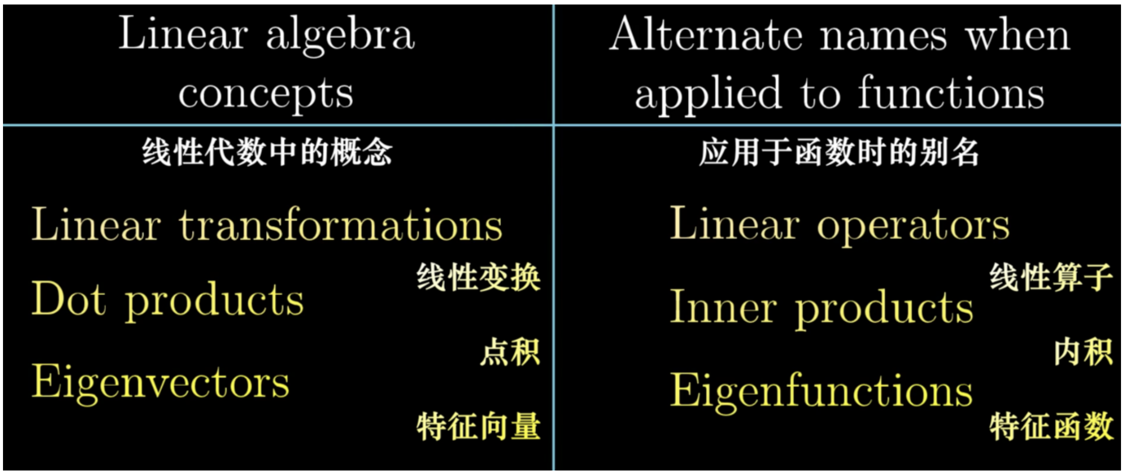
上面求导过程中的矩阵是如何得来的呢？是通过求每一个基函数的导数，然后把结果放在对应列：

$$\frac{d}{dx}b_0(x)=\frac{d}{dx}(1)=0=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{d}{dx}b_1(x)=\frac{d}{dx}(x)=1=\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{d}{dx}b_2(x)=\frac{d}{dx}(x^2)=2x=\begin{bmatrix} 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

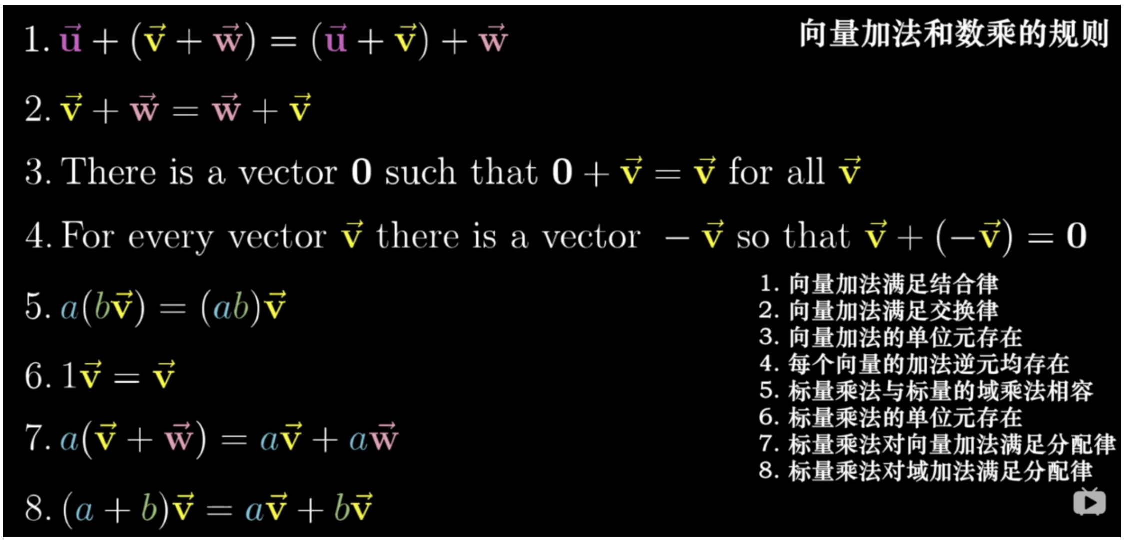
乍一看，矩阵向量乘法和求导是毫不相干的，但它们其实是一家人



最后，我们回到向量上来。数学中有很多类似向量的事物，只要你处理的对象具有合理的数乘和相加概念，不管是空间中的箭头，一组数、函数的集合，还是你定义的其它的东西。线性代数中所有关于向量、线性变换和其他概念都应用适用于它。

现在想象一下，你是一名发展线性代数理论的数学家。你想让你所做的所有定义和发现，不只对一个特殊情况适用，对其他类似向量的事物都有普适性（generality）。这些类似向量的事物，比如箭头、一组数、函数等，它们构成的集合称为“向量空间（vector spaces）”

为了使向量空间的概念统一，你需要做的是建立一系列向量加法和数乘必须遵循的规则：



这些规则被称为公理（axioms）。在线性代数的现代理论中，如果让所有已建立好的理论和概念适用于一个向量空间，那么它必须满足以上的8条公理。它本质上是一个清单，确保向量加法和数乘的概念确实是你所希望的那样。

这些公理并非是基础的自然法则，而是一个接口（interface）。一边连着你（发现

这些结论的数学家），另一边连着其他人（也就是想把这些理论应用于新的向量空间的人）。

作为一个数学家，不需要考虑人们定义的所有可能的奇怪向量空间了。你只需要根据这些公理证明你的结论。只要这些人的定义满足这些公理，他们就能顺利地应用你的结论，即便你根本没想过他们的这种情况。

因此，你往往会把你的所有结论抽象地表述出来，也就是仅仅根据这些公理描述，而不是集中于某一种特定的向量上。简而言之，这就是为什么你阅读的每一本教科书都会根据可加性和成比例性来定义线性变换，而不是用网格线保持平行且等距分布来定义。

所以，对于向量是什么这个问题，数学家会直接忽略。在现代理论中，向量的形式不重要，箭头、一组数、函数等待都无所谓，它可以是任何东西，只要向量相加和数乘的概念遵守上面的8条规则即可。

这就像问“3”究竟是什么一样。遇到具体情况时，它就代表三个东西的集合。但是在数学里，它被看做所有三个东西的集合的抽象概念，从而让你能用一个概念就能推出所有三个东西的集合。

向量也是如此，它有多种体现。但是数学把它抽象成“向量空间”这样一个无形的概念

> 普适的代价是抽象（abstractness is the price of generality）