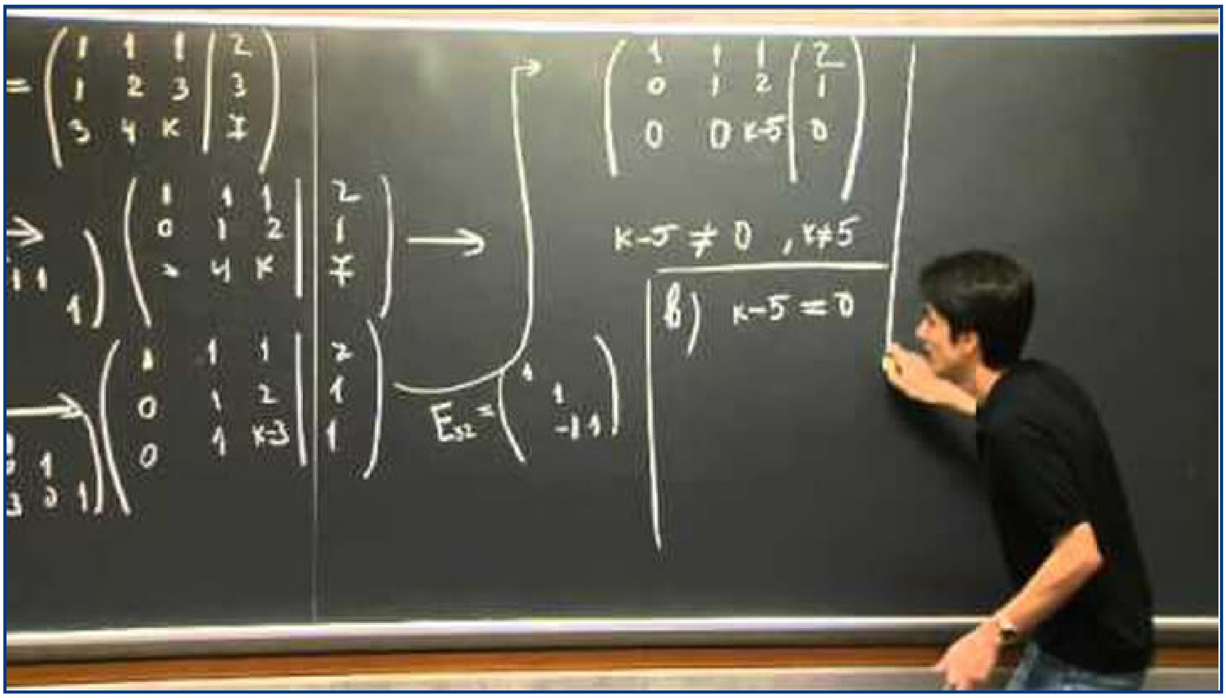
理解矩阵乘法

作者： 阮一峰

日期： 2015年9月 1日

大多数人在高中，或者大学低年级，都上过一门课《线性代数》。这门课其实是教矩阵。



刚学的时候，还蛮简单的，矩阵加法就是相同位置的数字加一下。

$$\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ 5 & 3 \end{pmatrix}$$

矩阵减法也类似。

矩阵乘以一个常数，就是所有位置都乘以这个数。

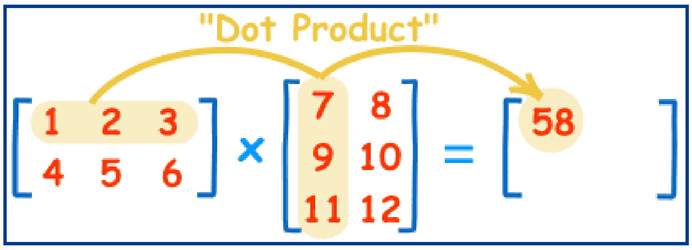
$$2 \times \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 8 & 6 \end{pmatrix}$$

但是，等到矩阵乘以矩阵的时候，一切就不一样了。

$$\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \times \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 7 & 8 \end{pmatrix}$$

这个结果是怎么算出来的？

教科书告诉你，计算规则是，第一个矩阵第一行的每个数字（2和1），各自乘以第二个矩阵第一列对应位置的数字（1和1），然后将乘积相加（ 2 x 1 + 1 x 1），得到结果矩阵左上角的那个值3。



也就是说，结果矩阵第m行与第n列交叉位置的那个值，等于第一个矩阵第m行与第二个矩阵第n列，对应位置的每个值的乘积之和。

怎么会有这么奇怪的规则？

我一直没理解这个规则的含义，导致《线性代数》这门课就没学懂。研究生时发现，线性代数是向量计算的基础，很多重要的数学模型都要用到向量计算，所以我做不了复杂模型。这一直让我有点伤心。

前些日子，受到一篇文章的启发，我终于想通了，矩阵乘法到底是什么东西。关键就是一句话，**矩阵的本质就是线性方程式，两者是一一对应关系。**如果从线性方程式的角度，理解矩阵乘法就毫无难度。

下面是一组线性方程式。

$$\begin{cases} 2x + y = 3 \\ 4x + 3y = 7 \end{cases}$$

矩阵的最初目的，只是为线性方程组提供一个简写形式。

$$\begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

老实说，从上面这种写法，已经能看出矩阵乘法的规则了：系数矩阵第一行的2和1，各自与 x 和 y 的乘积之和，等于3。不过，这不算严格的证明，只是线性方程式转为矩阵的书写规则。

下面才是严格的证明。有三组未知数 x、y 和 t，其中 x 和 y 的关系如下。

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = y_1 \\ a_{21}x_1 + a_{22}x_2 = y_2 \end{cases}$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

x 和 t 的关系如下。

$$\begin{cases} b_{11}t_1 + b_{12}t_2 = x_1 \\ b_{21}t_1 + b_{22}t_2 = x_2 \end{cases}$$

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

有了这两组方程式，就可以求 y 和 t 的关系。从矩阵来看，很显然，只要把第二个矩阵代入第一个矩阵即可。

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

从方程式来看，也可以把第二个方程组代入第一个方程组。

$$\begin{cases} a_{11}(b_{11}t_1 + b_{12}t_2) + a_{12}(b_{21}t_1 + b_{22}t_2) = y_1 \\ a_{21}(b_{11}t_1 + b_{12}t_2) + a_{22}(b_{21}t_1 + b_{22}t_2) = y_2 \end{cases}$$

上面的方程组可以整理成下面的形式。

$$\begin{cases} (a_{11}b_{11}+a_{12}b_{21})t_1+(a_{11}b_{12}+a_{12}b_{22})t_2=y_1 \\ (a_{21}b_{11}+a_{22}b_{21})t_1+(a_{21}b_{12}+a_{22}b_{22})t_2=y_2 \end{cases}$$

$$\begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

最后那个矩阵等式，与前面的矩阵等式一对照，就会得到下面的关系。

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{pmatrix}$$

矩阵乘法的计算规则，从而得到证明。

（完）

文档信息

发表日期： 2015年9月 1日

更多内容： 档案 » 算法与数学