线性代数

-31 课 线性变换

一. 知识概要

本节介绍线性变换，从线性变换概念谈起，然后从基和坐标的角度介绍了线性变换的矩阵形式。使得我们对线性变换问题具有更深入的了解。

二. 线性变换

2.1 线性变换

首先直接给出线性变换满足的条件：

【线性变换要满足的性质】

（1） $T(v+w)=T(v)+T(w)$

（2） $T(cv)=cT(v)$（由此有：$T(0)=0$）

或者可将上面两个式子联系起来，即任何一个线性组合的线性变换等同于 T(v)和 T(w)的同样的线性组合：

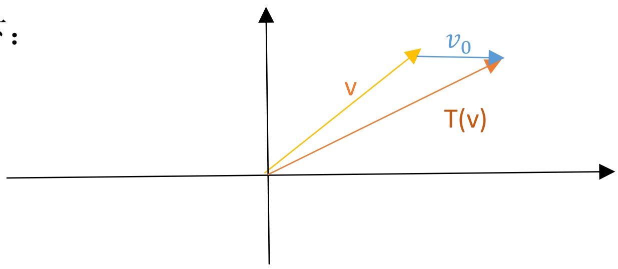
$$T(cv+dw) = c\ T(v) + dT(w)$$

其中的 T(..)可以理解为一种函数，有输入，有输出。将输入进行一番作用后得到输出。

例如投影，其输入就是一个向量，输出则是向量在直线上的投影。而投影本身就是一种线性变换。因为满足线性变换的两条性质。这一点画图即可判断，很简单。

接下来通过几个例子来了解这部分内容。

【例 1】平面平移：假如将整个平面都沿着某一方向平移 v_0。这是一个线性变换吗？示意如下：



分析：

显然不是的。假设我们将向量 v 长度增加一倍，变换之后的向量显然不是原向量变换结果 T（v）的两倍。甚至这个变换连最基本的 T(0)=0 都不满足。

【例 2】变换 $T(v)=\|v\|$（求长度）是一个线性变换吗？

分析：

对于零向量而言，$T(0)=0$

如果向量翻倍，它的长度也翻倍，没错。但是如果乘的是负数的话，这个变换就不满足数乘的性质了。例如：这里 $T(-v) \neq -T(v)$

故这不是线性变换。

【例 3】给定变换：对于平面内的任意向量，将其逆时针旋转 45 度得到一个新的向量。

分析：

这满足上面的两条性质。旋转后相加或数乘对线性变换运算无影响。所以这个操作 T 是线性变换。

【例 4】给定线性变换：$T(v)=Av$（A 是矩阵，假设 $T:R^3 \to R^2$）

分析：

直接代入性质验证：

$A(v+w)=Av+Aw$，满足性质一；

$A(cv)=cAv$，满足性质二。

所以这是一个线性变换。

根据上面这个例子我们发现，矩阵 A 必然是一个 2x3 型的矩阵。或者说，这个 2x3 矩阵的运算将原本的三维向量变换成了二维向量。这里我们发现一点：可以使用矩阵来描述线性变换。

2.2 线性变换的基向量与坐标

由于对于向量空间而言，所有向量都可以表示为基向量的线性组合。因此，理解一个线性变换，我们只需要了解基向量是怎么变换的：

$$T(v_1)、T(v_2)、T(v_3)\ldots T(v_n)$$

（其中，v_1 到 v_n 是输入空间的一组基。）

而知道这些就足以确定任何 v 的线性变换，为什么？因为向量空间内的向量 v 是基向量的线性组合。而由线性变换的两条性质易得：**任意的向量的线性变换都可以用基向量的线性变换的结果进行表示。**

数学式表达为：

设 $v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$，则有：

$$T(v) = c_1T(v_1) + c_2T(v_2) + \ldots + c_nT(v_n)$$

借着这个表达式，这里渗透一些坐标的概念：

我们知道在笛卡尔坐标系下。坐标就是 x，y——这实际上是取坐标轴上的一组单位向量作为基向量的结果。这里，我们推广来看：

将 v 表示成基向量的线性组合：$v = c_1v_1 + c_2v_2 + \ldots + c_nv_n$。这里对应的系数即为一组“坐标”。这说明：坐标来自一组基。因此，v 的坐标是一组数字，表示 v 由多少个基向量组成。

2.3 线性变换对应的矩阵表示

接下来，我们试着使用矩阵 A 来表示一个线性变换 T($R^n \to R^m$)
我们设 T 表示从 n 维空间到 m 维空间的变换。这里，我们需要两组基：

（1）输入空间的一组基来描述输入向量。
（2）输出空间的一组基，以确定输出向量的坐标。

令 $v_1, v_2, \ldots v_n$ 是来自于 R^n 空间的一组基，$w_1, w_2, \ldots w_m$ 是来自于 R^m 空间的一组基。对于一个向量，通过基将它表示出来，就能得到它的坐标。

下面我们来举一个具体的例子。

【例】

设定平面内的线性变换：将平面内的向量投影到一条定直线上。

我们选定该直线上的一个单位向量作为第一个基向量 v_1，在直线的垂直方向上我们可以取到另一个单位向量作为基向量 v_2。这两个基向量构成了输入空间的一组基。

注意：这里的变换是从平面到平面，所以输入向量与输出向量共用一组基。
那么现在的问题是：我们怎么确定矩阵 A 来描述变换？

step1：分析变换对于基向量的影响

我们先看该变换对于这两个基向量的影响。

对于投影而言，$T(v_1)=v_1$，$T(v_2)=0$

于是，对于一个输入向量 v（c_1，c_2）而言，得到的输出向量就是（$c_1,0$）

Step2：将变换写成矩阵形式

找到描述这一变换过程的矩阵 A，将该变换写成矩阵形式：

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}=\begin{bmatrix} c_1 \\ 0 \end{bmatrix}$$

回顾这个例子：选取的基一组与直线同向，一组与直线垂直垂直。之前介绍投影时说过，它们实际上都是投影的特征向量。所以得到的矩阵 A 是一个对角阵，对角线上都是特征值。

这就证明了结论：

如果以特征向量为基，可以得到线性变换的矩阵 A 是对角阵∧，对角线上都是特征值

此外，值得注意的是：选取特征向量作为基得到的矩阵是最好的，可以尝试使用其他基（例如坐标系下的标准基（1,0）与（0,1））对应得到的变换矩阵 A 将会相对复杂许多。

好的，通过上面的介绍，我们就知道了，线性变换过程可以用矩阵 A 来表示。那么现在如何确定矩阵 A 呢？

令 $v_1,v_2,\dots v_n$ 是来自于 R^n 空间的一组基，$w_1,w_2,\dots w_m$ 是来自于 R^m 空间的一组基。

求矩阵 A 的方法：

$$\begin{aligned} T(v_1) &= a_{11}w_1 + a_{21}w_2 + .. + a_{m1}w_m \\ T(v_2) &= a_{12}w_1 + a_{22}w_2 + .. + a_{m2}w_m \end{aligned}$$

其中 a 表示 A 矩阵中对应位置的元素。也就是说，计算 A 矩阵的方法是：**选取基向量$\boldsymbol{v_1}$，$\boldsymbol{v_2}$.. $\boldsymbol{v_n}$。通过线性变换得到输出向量 T($\boldsymbol{v_1}$)，T($\boldsymbol{v_2}$)..在输出空间里，T($\boldsymbol{v_i}$)就是输出基的线性组合（**$a_{11}w_1 + a_{21}w_2 + .. + a_{m1}w_m$之类**），这个线性组合的系数就是矩阵 A 的第 i 列。**

通过这种方式得到的矩阵 A，如果给定输入坐标，就有：

A(输入的坐标) = 输出坐标

（向量输入坐标乘上矩阵 A = 输出的空间中的坐标）

例如，输入的坐标为（1,0，..0）。这就意味着输入向量为基向量v_1=$\begin{bmatrix}1\\0\\0\\\dots\\0\end{bmatrix}$，使用 A 乘上输入坐标。得到的是 A 的第一列元素：a_{11}，a_{21}，.. a_{n1}这样的一组数字。而它们就是 T（v_1）的坐标。这种计算方法对所有基向量都成立。

【例】设定线性变换$T=\frac{d}{dx}$（求导数）

输入：$c_1+c_2x+c_3x^2$　　　基：1，x，x^2

输出：c_2+2c_3x　　　基：1，x

下面我们来求对应的矩阵 A

分析：

A 应当满足：

$$A\begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix}=\begin{bmatrix}c_2\\2c_3\end{bmatrix}$$

可以求得：

$$A=\begin{bmatrix}0&1&0\\0&0&2\end{bmatrix}$$

三. 学习感悟

本节针对线性变换进行了解释，尤其是最后的矩阵解释，确定输入输出空间的基后：

• 当已知线性变换矩阵 A 与输入的坐标，则可以使用 A 乘上输入的坐标，进而得到输出坐标。再使用输出坐标与输出空间的基进行组合得到输出的向量。

• 而如果已知输入坐标与输出坐标，就可以用另一个角度理解上面提到过的 A(输入的坐标) = 输出坐标。即在这个关系式中代入基向量的坐标如（1,0，..0），这样一来 A 的各列即为已知的输出坐标。或者说是已知的输出向量剥离输出空间的基之后得到的坐标。进而确定 A 的各列，确定 A 矩阵。