Preface

这一系列的几篇文章 (预计 20 篇之内吧) 根据 Gilbert Strang 的 Introduction to Linear Algebra 5th Edition 对 Linear Algebra(线性代数) 和 Sheldon Axler 的 Linear Algebra Done Right(3rd Edition) 的一些知识，以及自身的学习经历进行梳理和复习。书里讲到的知识点，我会根据自己的知识进行补充，尽量让它们更加直白容易理解。这些知识对于控制理论的学习是基础性的。我们不要求对每个细节都牢牢记住，但要求熟悉每一个基本概念。在后续其他学科的学习中，我们会时常用到 linear algebra 的许多知识，特别是在线性系统的学习之中，到时候我们就会发现特意回顾它们的价值。文中会时不时出现中英文的词汇的交替使用，一来是为了让读者熟悉这些核心概念的英语，二来是笔者写作的一点习惯。

这里选用了 ElegantBook 作为 Latex 模板写一个短幅的快速复习教程。在此对模板开发团队表示感谢。笔者也不是数学专业出身，对这门学科的浅薄理解如果有误，请君不吝指教。转载请联系笔者，并且注明出处，谢谢。

邮箱：frankmsj94@gmail.com

三脚猫 Frank

2019 年 12 月 12 日

于 California, USA

第 2 章 线性映射和矩阵初步
Linear Maps and Matrix

2.1 线性映射

我们在上一章回顾了向量和向量空间（线性空间）的一些定义和性质。这一章我们来讨论两个向量空间上的映射，我们在 $\mathbb{R}$ 上的两个向量空间 V 与 W 上定义一种**线性映射 (Linear Map)**。我们给出下面比较正式的定义：

定义 2.1. 线性映射

Linear map 是一个从 V 到 W 的映射 $T : V \to W$，并且满足以下两个性质

- 叠加性：$T(u+v) = T(u) + T(v)$，任取 $u, v \in V$
- 齐次性[a]:$T(\lambda) = \lambda T(u)$，任取 $u \in V, \lambda \in \mathbb{R}$

[a]homogeneity: the quality or state of being all the same or all of the same kind.

当向量空间中的向量是与数字有关的时候，我们更多地把映射称为函数 (Function)，那么线性映射就会指线性函数 (Linear Function)。线性函数的例子我们再熟悉不过了，最简单的一次函数 $y(x) = kx$。但是注意 $y(x) = kx + b$ 就不是线性函数，请自己采用定义体会一下。我们把后者称为 affine function(仿射函数)。线性的概念在这里要重新理解一下，因为国内初高中教材对于线性的定义是模糊的。Linear Map 有时候也被称为**线性变换 (Linear Transformation)**。从 V 到 W 的所有 linear map 的集合记为 $\mathcal{L}(V, W)$。

下面要给出一个关于线性映射存在且唯一的定理，在此之前我们明确一个结论，从向量空间的定义中我们知道**一个向量空间 W 中，任何向量的线性组合任然属于 W**[1]。

定理 2.1

设有向量空间 V 中一组基 $v_1, v_2, \ldots, v_n$ 和 W 中的 n 个向量 $w_1, w_2, \ldots, w_n$。一定且唯一存在一个线性映射 $T : V \to W$ 使得

$$Tv_j = w_j \tag{2.1}$$

j = 1,2,3,...,n.

[1]向量空间的基本性质

证明 先证明存在性。给定两个向量空间 V 和 W，我们可以定义一个如下映射 $T:V\to W$

$$T(c_1v_1+c_2v_2...+c_nv_n)=c_1w_1+c_2w_2...+c_nw_n$$

$c_j\in F$. 由于 $v_1,v_2,...v_n$ 是一组基，所以任何一个 V 中的元素都可以在 T 的作用下在 W 中找到对应的像，并且这个像我们定义为了这 n 个 W 中向量的线性组合，系数正好和逆像的系数一样。根据 linear map 的定义我们马上就可以验证这是一个线性映射。Existence 得到证明。我们每次令 $c_j=1$，其余系数都为 0，则得到 $T(v_j)=w_j$.

再证明这个映射 T 对于这两组向量是否唯一存在。假设有另一个线性映射 $P:V\to W$ 也满足上面的 $P(v_j)=w_j$. 那么根据线性映射的性质有:

$$P(c_1v_1+c_2v_2...+c_nv_n)=c_1w_1+c_2w_2...+c_nw_n=T(c_1v_1+c_2v_2...+c_nv_n)$$

于是得到 $T=P$. Uniqueness 得到证明. □

为什么我们可以这么定义出一个映射呢？因为首先根据基的定义，我们可以用 $c_1v_1+c_2v_2...+c_nv_n$ 表示任何属于 v 的向量。然后对于每个 v，我们又可以把它与 $w_1,w_2,...,w_n$ 的某一个线性组合对应到一起。为了使得它是一个线性映射，我们这里使得系数一样，这是完全做得到的。由此我们证明了这时候 T 是一个线性映射，并且它是存在的，并且是唯一的。这里的唯一性是指给定一组基 $v_1,v_2,...,v_n$ 和其对应的 W 中的像只能唯一确定一个 T，不能确定两个。但是注意，再找另外一组基和另外一组像，我们可能还能得到同样的 T。

下面我们聊聊线性映射之间的运算。我们可以对线性映射集 $\mathcal{L}(V,W)$ 定义 addition 和 scalar multiplication。现在有两个 linear map：S 与 T，定义:

$$(S+T)(x)=S(x)+T(x),S(\lambda x)=\lambda S(x),\lambda\in F$$

我们验证 $(S+T)(x)$ 是否是一个 linear map 呢?

$$(S+T)(x+y)=S(x+y)+T(x+y)=S(x)+T(x)+S(y)+T(y)=S(x+y)+T(x+y)$$

同理可以验证数乘性质。我们发现这样定义运算得到 map 仍然是 linear 的，并且仍然是属于 $V\to W$ 的 map，于是乎我们有如下结论。

命题 2.1

所有 $T:V\to W$ 的集合 $\mathcal{L}(V,W)$ 是一个向量空间

这是个从向量空间定义出发的，通过我们上面分析就能得到的结论。两个 linear map 之和或者 linear map 与一个数相乘，仍然是 linear map。我们接下来还可以定义两个 linear map 的乘积：$(ST)(x)=S(Tx)$，我们同样可以验证 ST 这个映射仍然是一个 linear map，即**我们这样定义线性映射之间的乘积仍然是线性映射**。线性映射的乘积是有一些运算性质的，比如 $(T_1T_2)T_3=T_1(T_{2T}3)$，表明线性映射之间的乘积满足结合律；再如线性映射满足 $T(0)=0$。但是很重要的一点，它们

不满足交换律。一旦线性映射的乘积顺序发生了改变，那么它们便和原来不再相等。

说了这么多，你可能已经发现了，这和矩阵的运算和性质是一样的。**矩阵和一个线性映射是可以对应的**，或者说，**矩阵是线性映射的一种表示方法**。

2.2 矩阵

关于矩阵的运算方法我们都再熟悉不过了[2]，我们就不再这里赘述了。从实用的角度来讲，我们更关心的是数值算法如何计算一些矩阵相关的属性，比如特征值，求逆等等，特别对于大型矩阵来说。这个话题以后有机会再提。

在 Linear Algebra Done Right 上把矩阵与 linear map 放在一起讲，而 Introduction to Linear Algebra 中则是直接以列向量的线性组合来引出矩阵。我们已经很熟悉矩阵的一些基本运算了，但是对矩阵的不同理解会在一些需要分析推导的场合产生影响。我们首先从线性映射的角度讲讲矩阵。

根据上一节的定理，如果选定了 V 中的一组长度为 n 的基和 W 中的 n 个向量[3]，那么我们就可以唯一确定一个线性映射 T。但反过来，我们问：**给定了一个线性映射 T，我们是否同时能确定 V 中的基和其在 W 中的像呢？**答案显然是不能的。根据式2.1，我们发现即便是同一个线性映射，如果基 v_j 和像 w_j 是可以同时变化的。也就是说**虽然给定一组基和对应的像，它们之间的线性映射 T 是唯一的，但是同一个 T 可以由多组不同的基和像来确定**。现在如果给定一个线性映射 T，并且自行选定一组 V 中的基，可以知道必然唯一确定了 W 中的 n 个像。由于这 n 个像都是 W 中的元素，我们就自然可以用 W(假设为 m 维空间) 中的基 $(w_1, w_2, ..., w_m)$ 来表示这 n 个像。于是，我们自然得到了一个结论:

命题 2.2

给定 n 维向量空间 V 中的一组基 $(v_1, v_2, ..., v_n)$ 和 m 维向量空间的一组基 $(w_1, w_2, ..., w_m)$，如果有一个确定的线性映射 $T : V \to W$，那么就有:

$$Tv_k = A_{1,k}w_1 + ... + A_{m,k}w_m \tag{2.2}$$

成立。

根据式子2.2，我们可以定义一个矩阵来表示 T，前提是我们选定了 V 和 W 中的基。因为基不同，2.2中的系数显然就不同，于是定义的矩阵也就会发生变

[2]如果不熟悉可以参考两本书的 matrix 部分进行回顾。

[3]注意，W 空间中我们并不要求这 n 个向量是基。W 空间可以是不同于维度 n 的。

$$\mathcal{M}(T) = \begin{array}{c} \\ w_1 \\ \vdots \\ w_m \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & \dots & v_k & \dots & v_n \end{array} \\ \left(\begin{array}{ccccc} & & A_{1,k} & & \\ & & \vdots & & \\ & & A_{m,k} & & \end{array} \right) \end{array}.$$

图 2.1: Definition of matrix of a linear map T

化。我们定义一个矩阵，使得下面两个式子表达同样的映射关系：

$$Tv = T(c_1v_1 + \ldots + c_nv_n) = a_1w_1 + \ldots + a_mw_m \tag{2.3}$$

$$\mathcal{M}v = \mathcal{M}[c_1, c_2, \ldots, c_n]^{\mathrm{T}} = [a_1, a_2, \ldots, a_m]^{\mathrm{T}} \tag{2.4}$$

当我们把两个空间中的基给确定下来了，那么原来元素的坐标和其像的坐标在 V 和 W 中都确定了下来。矩阵 $\mathcal{M}$ 这里借用 LADR 书中的图2.1来显示这种定义。它的**每一列的系数就是原来 V 中的基向量 v_k 在 T 作用下，映在 W 中的像在选定的基 $\{w_m\}$ 下的坐标**。。

我们可以在求一个线性映射的矩阵表示时，可以选取两个空间的标准基[4]比如对于 $T(x,y) = (x+3y, 2x+5y, 7x+9y)$，选取 $T(1,0) = (1,2,7)$ 和 $T(0,1) = (3,5,9)$，我们就得到了矩阵的两列。于是我们有表达该线性映射的矩阵：

$$\mathcal{M}(T) = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix} \tag{2.5}$$

那么任何一个此二维空间下元素 (x,y) 在标准基下的坐标都可以由这个矩阵映射到一个三维空间中产生标准基下的坐标 (a,b,c)。所以说**矩阵作为线性映射或者线性变换的一种表示，前提是要明确两个空间中的选取的基**。

从另一个角度来看，一个矩阵 $\mathcal{M}$ 与向量 v 的乘积，可以看作是 $\mathcal{M}$ 的每个列向量以 v 中的系数进行线性组合的结果。用一个简单的例子说：

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} v_1 + \begin{bmatrix} 2 \\ 4 \end{bmatrix} v_2$$

知道这两点暂时就足够了。其他有关矩阵的更多属性在之后我们一步步再分开来讲。线性代数的核心问题是在**求解线性方程组**，而线性方程组是可以自然地由 $\mathbf{A}x = b$ 来表示的，其中 A 是一个 m-by-n 的矩阵。由此如果 b 是 $\mathbf{A}$ 中列向量的一种线性组合，那么我们自然可以解得其系数向量 x。接下来我们会先讲解求解线性方程组的有关问题，然后引出线代中非常重要的，关于矩阵子空间的讨论。

[4]标准正交基，默认大家都熟悉。

重点

- 线性映射满足叠加性和齐次性
- 线性映射的存在与唯一定理
- 线性映射的集合是一个向量空间
- 线性映射的加法、数乘、乘法仍然是线性映射
- 矩阵是一种线性映射的表示方法
- 矩阵与向量的乘积可以看成是列向量的线性组合。

Reference

- 2.1 节参考了 Sheldon Axler 的 Linear Algebra Done Right(3rd Edition) 第三章 3A
- 2.2 节参考了同书的第三章 3C 以及 Gilbert Strang 的 Introduction to Linear Algebra(5th Edition) 第一章第三节