

# 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3章 核空间、像空间与秩

### Kernel, Image and Rank

#### 3.1 $Ax=b$ 的两种理解

线性代数的核心问题是在于解线性方程组  $Ax=b$ , 然而高斯消元法以及一些改良方法, 我觉得应该放到数值计算那里再去细讲。默认读者是第二遍学习, 我这里就跳过基础的解线性方程组的问题了。对于矩阵  $A$  与向量  $x$  的相乘, 可以有两种理解方式:

- 等式左边看作是矩阵  $A$  中列的线性组合。如果  $b$  在它们所有可能的线性组合之中, 我们就可以找到对应的系数, 那么  $x$  有解。记  $A = [A_1|A_2|\dots|A_n]$ , 其中  $A_j$  代表  $A$  的第  $j$  列, 我们可以写成:

$$x_1 A_1 + x_2 A_2 + \dots + x_n A_n = b \quad (3.1)$$

我们会把  $A_j$  张成的空间叫列空间 (column space) 或者像空间 (image)。

- 矩阵乘法看作是行与向量之间的 dot product。记  $A = [t_1; t_2; \dots; t_m]$ , 其中  $t_k$  是  $A$  的第  $k$  行, 那么我们可以把  $Ax=b$  写成

$$Ax = \begin{bmatrix} t_1 x \\ t_2 x \\ \vdots \\ t_n x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (3.2)$$

#### 3.2 核空间 Kernel

首先回顾一个矩阵的**核空间 (Kernel)**, 也叫做**零空间 (null space)**。之所以标题叫 Kernel 而不是 null space, 其实是因为我个人在 paper 里见 kernel 的次数比较多, 写起来比较简洁。这完全是个人习惯, 为了让大家不至于忘记它还叫 null space, 我在下文就穿插使用了。

对于上面这个线性方程组, 其完整的解是应该是  $Ax=0$  的通解与  $Ax=b$  的特解之和。那么这里使得矩阵  $A$  满足  $Ax=0$  的所有  $x$  向量的集合就是矩阵  $A$  的 kernel 或者 null space, 我们记为  $\text{Ker}(A)$  或者  $\text{null}(A)$ 。

我们假设  $x \in \mathbb{R}^n$ , 那么我们可以证明  $\text{Ker}(A)$  是  $\mathbb{R}^n$  的一个子空间。证明其为 subspace 的过程我们简单叙述: 首先  $\text{Ker}(A)$  显然是  $\mathbb{R}^n$  的子集, 其次任何两个  $\text{Ker}(A)$  上的元素  $x, y$  之和,  $A(x+y)=Ax+Ay=0$ , 任然属于  $\text{Ker}(A)$ 。任意一个元素的

scalar multiplication 也同样在  $\text{null}(A)$  中封闭。于是矩阵的 null space 或者 kernel 就是  $\mathbb{R}^n$  的一个 subspace。特别的，零向量自身构成了一个零空间，因为任何矩阵与它的乘积都是零向量，并且加法与数乘都是封闭的。

事实上 kernel 和 null space 不光光可以针对矩阵。由于矩阵和线性映射的对应关系，Kernel 和 null space 也可以针对一个 linear map 而言。我们在这里把矩阵替换成  $T$  之后就可以同样定义一遍。回顾在映射的概念中，什么叫做单射 (injective)。一个映射是 injective 的，即满足  $f(x) = f(y) \rightarrow x = y$ 。考虑一个  $n$  维实向量空间  $V$  和一个  $m$  维实向量空间  $W$ ，有一个 linear map 记作  $T: V \rightarrow W$ ，如果  $T$  是单射的，那么由于  $T(0)=0$ ，零向量就是满足  $Tx=0$  唯一的向量。同理  $Ax=0$ ，如果  $x=0$  是唯一的解，那么我们必然知道  $A$  的各列是线性独立或者线性无关的。反过来，如果  $A$  的各列是 linear independent 的，那么  $A$  的 Kernel 或者 null space 就是 trivial 的，即由单一的零元素组成的。

Kernel 这个名字其实也很形象，就是  $V$  中的元素在  $T$  或者  $A$  的作用下，压缩到了原点，就像一个核一样。如果不巧，一个 basis 中的向量被压缩到了原点，那么其像空间的维度就会相应减少，我们下面会详细展开这个讨论。直观上我们能感觉到 Kernel 的维度如果变高了，说明更多信息都被压缩了。这就是说在  $T$  的作用下原空间降维了。

### 3.3 像空间 Image

一个  $m$ -by- $n$  的矩阵  $A$  的像空间 (Image) 也被称为列空间 (column space)，记为  $\text{Im}(A)$ 。同理我们也可以对映射进行类似定义，对映射定义时像空间也叫 range，记为  $\text{range}(T)$ 。 $\text{Im}(A)$  是  $A$  的所有列张成的空间，即所有列的线性组合构成的空间。它同样是一个子空间，但是却是  $\mathbb{R}^m$  的子空间。这一点我们同样要通过子空间的定义去证明。另外如果  $T$  是满射的 (surjective)，那么  $T$  的像空间就是  $W$ 。换做是矩阵，意味这其列向量张成了  $\mathbb{R}^m$  空间。

### 3.4 秩

秩 (rank) 的概念我们一定不陌生。一个矩阵的行秩是行张成空间的维度，列秩是列张成的维度。关于秩我们回顾两个重要的定理：

#### 定理 3.1

行秩与列秩相等



于是我们可以把行秩和列秩统一称为 rank，一般可以计算列秩。

**定理 3.2**

A 是一个  $m$ -by- $n$  的矩阵, 那么  $\text{rank}(A) = \text{Dim}(\text{Im}(A))$



上面的定理告诉我们一个矩阵的列秩是列空间或者像空间的维度。这很好理解, 本来我们定义像空间就是由矩阵的所有列张成的。那么列秩代表了列的最大的线性无关组中包含的列数, 也就是像空间最小 spanning list 的长度。

## 3.5 线性映射基本定理

线性映射基本定理, 也叫秩-零化度定理 (rank-nullity theorem)。我这里采用线性映射的说法, 再用矩阵去解释。我们不完整证明它, 只说说它的直观理解。

**定理 3.3**

$V$  是一个有限维空间,  $T$  是一个从  $V$  到另一个向量空间  $W$  的线性映射。那么  $T$  的像空间是一个有限维空间, 并且

$$\dim V = \dim \text{null}(T) + \dim \text{range}(T) \quad (3.3)$$



这里定理说  $V$  的维度就是  $T$  零空间的维度与像空间维度之和。我们从最简单的情况分析一下是不是对的。

### 3.5.1 null space 只有零向量

首先一个线性映射  $T$  的 null space 可以分为两种情况, 只有 0 或者还存在非 0 向量。我们之前说过  $T$  单射时, null space 只有 0 向量, 此时由于只有 0 向量, 它自身线性相关, 于是没有 basis, 维度为 0。此时我们考虑  $V$  上的一组基, 在  $T$  的作用下, 这组基的像  $Tv_j$  是落到了像空间  $\text{range}(T)$  中的, 并且不在 null space 中 (0 向量是唯一解), 那么我们证明一下这组基的像仍然是线性独立的, 并且上面基本定理成立。

**证明** 已知  $T$  的 null space 中只有 0 向量, 则  $\dim \text{null}(T) = 0$ 。现取  $(v_1, v_2, \dots, v_n)$  是  $V$  中的一组基, 那么一定有:

$$c_1 v_1 + c_2 v_2 + \dots + c_n v_n = 0 \quad (3.4)$$

当且仅当  $c_j$  同时为 0。现在对上式做  $T$  变换

$$T(c_1 v_1 + c_2 v_2 + \dots + c_n v_n) = 0 \quad (3.5)$$

$$c_1 T v_1 + c_2 T v_2 + \dots + c_n T v_n = 0 \quad (3.6)$$



由于  $Tv_j \neq 0$  ( $v_j$  不在 null space 中), 所以得到这  $n$  个  $Tv_j$  线性独立。因为所有  $V$  中的向量都可以由  $(v_1, v_2, \dots, v_n)$  表示, 像空间中的向量:

$$Tv = T(b_1v_1 + b_2v_2 + \dots + b_nv_n) \quad (3.7)$$

$$= b_1Tv_1 + b_2Tv_2 + \dots + b_nTv_n \quad (3.8)$$

所有像都可以由  $Tv_j$  来表示, 于是我们肯定像空间的维度为  $n$ 。即有  $\dim V = 0 + n = n$   $\square$

我们从上面的证明中可以看到, 如果有一个  $T$ , 它的 null space 中只有一个零元素, 那么原空间中的基也就可以仍然通过线性映射转化到像空间中, 成为像空间的基。这时候我们发现在这种  $T$  的作用下, 原空间与像空间的维度没有发生变化!

### 3.5.2 null space 还有非零向量

现在假设有一个线性映射  $T$ , 它的 null space 不再只有一个零向量了, 假设里面还有许多非零向量。<sup>1</sup>, 但是我们不确定维数。没有关系, 我们假设 null space 其中一个非零向量  $v^*$ , 我们知道有一组不全为 0 的系数  $a_j$  使得<sup>2</sup>:

$$Tv^* = 0 \quad (3.9)$$

$$T(a_1v_1 + a_2v_2 + \dots + a_nv_n) = 0 \quad (3.10)$$

$$a_1Tv_1 + a_2Tv_2 + \dots + a_nTv_n = 0 \quad (3.11)$$

$a_j$  不全为 0, 那么我们就知道  $Tv_j$  必然不是线性独立的。即像空间的维度一定会小于  $n$ 。我们现在假设  $a_j$  中最多只有一个不为 0, 比如  $a_1 \neq 0$ , 其他系数必须满足  $a_2 = \dots = a_n = 0$ , 则可以得到

$$a_1Tv_1 = 0, a_1 \neq 0$$

$$a_2Tv_2 + \dots + a_nTv_n = 0, a_2 = \dots = a_n = 0$$

那么我们推知  $\{Tv_2, \dots, Tv_n\}$  线性独立, 并且

$$Tv_1 = 0$$

即我们选择的一组基中的向量  $v_1$  是属于这个 null space 的,  $v_1$  这个基中的向量被压缩到了 kernel 里了! 那么此时像空间中的元素怎么表示呢?

$$Tv = T(b_1v_1 + b_2v_2 + \dots + b_nv_n) \quad (3.12)$$

$$= b_1Tv_1 + b_2Tv_2 + \dots + b_nTv_n \quad (3.13)$$

$$= b_2Tv_2 + \dots + b_nTv_n \quad (3.14)$$

<sup>1</sup>因为向量可以缩放, 所以实际上 null space 里的元素要么只有一个, 要么就是无穷多个

<sup>2</sup>因为  $v^*$  是非零向量

由于  $v_1$  被压缩为零向量了，我们发现所有像空间的元素都可以由  $\{Tv_2, \dots, Tv_n\}$  表示，并且由于它们线性独立，所以**像空间的维度**  $\dim \text{range}(T) = n - 1$ 。

那么对应 null space 的维度呢？由于我们假设只有  $a_1 \neq 0$ ，实际上此时，任何一个 null space 中的非零元素必然满足那么  $v^* = a_1 v_1$ 。所以此时 null space 的上线性独立的向量就只有  $v_1$  或者与它成正比的向量<sup>3</sup>。我们直接取  $v_1$  作为基，那么显然有  $\dim \text{null}(T) = 1$ 。于是  $\dim V = \dim \text{null}(T) + \dim \text{range}(T) = 1 + (n - 1) = n$ 。定理结论再次得到证实。

我们可以在上面的过程中任意选择几个  $a_j$  不为 0，而其他系数全为 0，则我们会对应产生 2 维的 null space， $n-2$  维的像空间。依次类推，如果  $T$  这个线性映射是个零映射，那么所有的元素都被压缩到了 kernel 里，null space 就是  $n$  维，对应的像空间就变成了 0 维。所以我们发现零空间与像空间的维度在  $T$  的作用下是一个**此消彼长的过程**：像空间的维度会随着零空间的维度的增大而减小。

### 3.5.3 矩阵表述

上面所有的结论我们都可以用矩阵来重新表述一遍。定义零化度 (Nullity)  $\text{Nullity}(A) = \dim \ker(A)$  首先定理可以表述为

#### 定理 3.4. rank-nullity theorem

$V$  是一个有限维空间， $T$  是一个从  $V$  到另一个向量空间  $W$  的线性映射。 $A$  是由  $T$ ，以及  $V$  和  $W$  的一组基确定的矩阵。那么  $\text{Im}(A)$  是一个有限维空间，并且

$$\dim V = \text{Nullity}(A) + \text{Rank}(A) \quad (3.15)$$

意思还是一样的，矩阵  $A$  的 null space 维度与 column space 的维度之和为  $V$  的维度。回忆  $T$  对应的矩阵的定义方式，矩阵的第  $k$  列就是

$$Tv_k = c_1 w_1 + c_2 w_2 + \dots + c_m w_m \quad (3.16)$$

如果矩阵的零空间中只有零向量，说明各列是线性独立的，那么  $Tv_k$  也就是线性独立的，故而列秩为  $n$ ，像空间维度为  $n$ 。如果有非零向量，则各列线性相关，我们选择踢出相关的列，这些列对应的基向量  $v_{i_1}, v_{i_2} \dots v_{i_p}$  就被压缩到了零空间，于是产生了  $n-p$  维的零空间。剩下的  $p$  列线性独立向量，张成了  $p$  维像空间。

## 重点

- $Ax=b$  的两种理解

<sup>3</sup>这里表达的意思是向量可以比例伸缩

- 核空间与像空间的定义
- 秩的回顾
- 线性映射基本定理的理解

## Reference

- 全文参考了 Sheldon Axler 的 Linear Algebra Done Right(3rd Edition) 第三章 3B
- 关于秩-零化度定理的直观解释, 参考了 stack exchange 的一个[回答](#)。我把比较完整的一个思路用数学语言给补全了。