

线性代数第三次习题课

李弢

2019 年 10 月 31 日

目录

1 线性映射的矩阵	1
1.1 复习与思考	1
1.2 习题选讲	3
2 多项式	6
2.1 习题选讲	6
3 特征值, 特征向量, 不变子空间	7
3.1 复习与思考	7
3.2 习题选讲	8

1 线性映射的矩阵

1.1 复习与思考

本节介绍了线性映射与矩阵的关系。在一般的线性代数教材中, 往往是先引入矩阵与行列式的计算, 然后稍微提及线性空间以及线性映射等概念, 这么做给人感觉缺乏动机, 不知道矩阵这个定义有什么用, 似乎完全是为了为难学生而造出来的。通过这一节的学习我们知道, 矩阵是线性映射的一种表示。什么叫做“表示”? 就是说本来有个抽象的东西在那, 我知道它在那, 也知道它是什么, 但就是说不出来。有了“表示”之后, 我就能抓住它了, 把它清晰地刻画出来。比如, 平面上有很多条直线, 每条直线上都有无穷多个点, 似乎难以把握。但是每条直线都可以写成 $ax + by = c$ 这样的形式, 只

要确定了 a, b, c , 一条直线就完全决定了, 那么这个 a, b, c 就是平面上直线的一种表示。之前我们说过, 如果有线性空间 V, W , 那么它们之间的线性映射的全体 $\mathcal{L}(V, W)$ 也是一个线性空间。有同学就觉得这个线性空间难以把握, 特别是其中的向量本身就是线性映射, 给人一种说不出的感觉。学了这一节之后我们就发现, 里面的向量可以看成是矩阵, 这样表示出来就清晰多了。具体来说, 如果 $T \in \mathcal{L}(V, W)$, V 的一组基为 v_1, v_2, \dots, v_m , W 的一组基为 w_1, w_2, \dots, w_n , 那么 $Tv_i \in W$ 可以用 w_1, w_2, \dots, w_n 线性表示出来, 如 $Tv_i = \sum_{j=1}^n a_{ji}w_j$. 我们把这些表示的系数写到一块, 就变成了一个矩阵,

$$\mathcal{M}(T) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}.$$

寻找这个矩阵的过程其实类似于填写表格, 要填写第 (j, i) 个元素, 只需要看 T 把 v_i 映射完了以后, 在 w_j 这个方向的分量。对于 V 里的每一个向量 $v = \sum_{i=1}^m x_i v_i$, 有

$$Tv = T \sum_{i=1}^m x_i v_i = \sum_{i=1}^m T(x_i v_i) = \sum_{i=1}^m x_i T v_i = \sum_{i=1}^m x_i \sum_{j=1}^n a_{ji} w_j = \sum_{j=1}^n \left(\sum_{i=1}^m a_{ji} x_i \right) w_j$$

也就是说, 在 V 里相对于 v_1, v_2, \dots, v_m 的坐标为 (x_1, x_2, \dots, x_m) 的元素, 在经过 T 作用后映射到 W , 相对于 w_1, w_2, \dots, w_n 的坐标为

$$\left(\left(\sum_{i=1}^m a_{1i} x_i \right), \left(\sum_{i=1}^m a_{2i} x_i \right), \dots, \left(\sum_{i=1}^m a_{ni} x_i \right) \right).$$

按照矩阵和向量乘法的定义, 这就是 $\mathcal{M}(T)x$.

表示的方法不是唯一的, 比如在刚刚的例子里面我们很容易看出 a, b, c 与 $\lambda a, \lambda b, \lambda c (\lambda \neq 0)$ 所代表的直线都是一样的。我们该选择哪个表示? 一般来说肯定是希望越简单越好, 比如我们可以限定 $b = 1$, 那么表示出所有不平行于 y 轴的直线是可以的。对于线性映射而言, 我们也有多种表示它的方式, 很明显一个线性空间有无数组基, 怎样选取 V 和 W 的基使得对应的矩阵 A 尽可能的简单? 这个问题要放到以后来讨论。

根据刚刚说的线性映射与矩阵的关系, 可以定义矩阵的加法和数乘: $\mathcal{M}(T) + \mathcal{M}(S)$ 应为 $\mathcal{M}(T + S)$, 即 $T + S$ 的矩阵; $\lambda \mathcal{M}(T)$ 应为 $\mathcal{M}(\lambda T)$

的矩阵。此外,最重要的是矩阵乘法, $\mathcal{M}(T)\mathcal{M}(S)$ 应该是 $\mathcal{M}(T \circ S)$, 即 T 与 S 复合的矩阵。在矩阵加法中, 两个矩阵的形状是完全一样的, 它们都是同样的两个线性空间之间的线性映射。在矩阵乘法中, 两个矩阵的形状却不一样, 记忆的要点是中间的维数要相同才能乘到一起。这里可以看出为什么矩阵乘法的定义如此奇怪, 而不是对应元素相乘。目的是为了保证这样定义的运算有几何意义, 即与映射的复合相容。

给定了 V 和 W 的基的情况下, $\mathcal{L}(V, W)$ 与 $\mathbf{F}^{n \times m}$ 之间存在一个线性的一一映射 \mathcal{M} . 这就是说, 我们可以把线性映射和矩阵给等同起来, 给了一个线性映射我就可以拿一个矩阵表示它, 给了一个矩阵我也就可以找一个对应的线性映射。

线性映射的逆类似于“反函数”, 不是所有的线性映射都存在逆, 只有当它既是单射又是满射的时候才可逆。如果原空间和像空间相同, 那么单射和满射其中之一成立即可推出可逆。如果两个线性空间之间存在可逆的线性映射, 就称它们同构。所谓的“同构”给人一种一一对应的感觉, 它们的结构是一样的。有限维线性空间同构当且仅当维数相等。特别的, 任何一个 n 维线性空间总是同构于 \mathbf{F}^n , 这一点其实很显然, 比如取 V 的随便一组基, 就可以把 V 中的元素用 n 元组 (坐标) 表示出来, 反正任何一个 n 元组, 也可以看成是基前面的系数, 线性组合出 V 中的一个向量。

1.2 习题选讲

1. 设 W 是有限维的, 并且 $T \in \mathcal{L}(V, W)$. 证明 T 是单射当且仅当有 $S \in \mathcal{L}(W, V)$ 使得 ST 是 V 上的恒等映射。

解答: \Rightarrow : 我们构造这样的 S . 由于 T 是单射, 那么将其视为像空间为 $\text{range } T$ 的线性映射 $\hat{T} \in \mathcal{L}(V, \text{range } T)$ 是可逆的, 设其逆为 $\hat{S} \in \mathcal{L}(\text{range } T, V)$. 如果 $\text{range } T = W$, 那么 \hat{S} 就是要找的映射。否则, 将像空间 W 分解为 $\text{range } T \oplus W'$, 再令 $S = \hat{S}P_{\text{range } T, W'} \in \mathcal{L}(W, V)$, 它是两个线性映射的复合, 因而也是线性映射, 并且满足 $ST = \hat{S}P_{\text{range } T, W'}T = \hat{S}\hat{T} = \text{Id}_V$ 。

\Leftarrow : 若有 $Tv_1 = Tv_2$, 那么 $v_1 = ST(v_1) = ST(v_2) = v_2$, 因此 S 单射。

2. 设 V 是有限维的, 并且 $T \in \mathcal{L}(V, W)$. 证明 T 是满射当且仅当有 $S \in \mathcal{L}(W, V)$ 使得 TS 是 W 上的恒等映射。

解答: \Rightarrow : 我们可以将 V 分解为 $T \oplus U$. 由于 T 是满射, 对于任意 $w \in W$, 总存在 $v \in V$ 使得 $Tv = w$. 这里的 v 不是唯一的, 但可以分解为 $n + u$,

其中 $n \in T$ 且 $u \in U$. 这样的分解中 u 是唯一的: 设 $Tv_1 = Tv_2 = w, v_1 = n_1 + u_1, v_2 = n_2 + u_2$, 那么 $T(u_1 - u_2) = T(n_2 - n_1) = 0$, 而由 $u_1 - u_2 \in U$ 和 $T \cap U = 0$ 可知 $u_1 = u_2$. 我们定义 $Sw = u$. 容易验证 S 是线性的且 TS 是 W 上的恒等映射。

\Leftarrow : 对于任意 $w \in W$, 都有 $T(S(w)) = w$, 因此 T 是满射。

3. 设 U 和 V 都是有限维向量空间, 并且 $S \in \mathcal{L}(V, W), T \in \mathcal{L}(U, V)$, 证明

$$\dim \text{null } ST \leq \dim \text{null } S + \dim \text{null } T.$$

解答: 显然, $\text{null } T$ 是 $\text{null } ST$ 的子空间. 取 $\text{null } T$ 的一组基 u_1, u_2, \dots, u_m , 可将其扩充为 $\text{null } ST$ 的一组基 $u_1, \dots, u_m, v_1, \dots, v_l$. 显然有 $S(Tv_i) = 0, i = 1, 2, \dots, l$, 即 $Tv_i \in \text{null } S, i = 1, 2, \dots, l$. 现说明 Tv_1, Tv_2, \dots, Tv_l 线性无关, 若否, 存在不全为零的 $\lambda_1, \lambda_2, \dots, \lambda_l$ 使得

$$\lambda_1 Tv_1 + \lambda_2 Tv_2 + \dots + \lambda_l Tv_l = 0,$$

由此 $T(\lambda_1 v_1 + \dots + \lambda_l v_l) = 0$, 即 $\lambda_1 v_1 + \dots + \lambda_l v_l \in \text{null } T$, 这与 v_1, \dots, v_l 的取法矛盾. 因此, $\text{null } S$ 包含一个长度为 l 的线性无关组 Tv_1, Tv_2, \dots, Tv_l , 其维数必然大于等于 l .

$$\dim \text{null } ST = m + l = \dim \text{null } T + l \leq \dim \text{null } T + \dim \text{null } S.$$

4. 证明矩阵加法和乘法的分配性质成立, 即设 A, B, C 都是矩阵, 并且 $A(B + C)$ 有意义, 那么 $A(B + C) = AB + AC$.

解答: 由于矩阵和线性映射之间一一对应, 我们可以将矩阵看成是线性映射, 即存在线性映射 S, T, R 使得 $\mathcal{M}(S) = A, \mathcal{M}(T) = B, \mathcal{M}(R) = C$, 那么

$$\begin{aligned} A(B + C) &= \mathcal{M}(S)(\mathcal{M}(T) + \mathcal{M}(R)) \\ &= \mathcal{M}(S)\mathcal{M}(T + R) = \mathcal{M}(S \circ (T + R)) \\ &= \mathcal{M}(S \circ T + S \circ R) \\ &= \mathcal{M}(S \circ T) + \mathcal{M}(S \circ R) \\ &= AB + AC. \end{aligned}$$

5. 设 V 是有限维的, 并且 $S, T \in \mathcal{L}(V)$, 证明 ST 可逆当且仅当 S 和 T 都可逆。

解答: \Rightarrow : ST 可逆等价于 ST 既是单射又是满射, 等价于 $\text{null } ST = 0, \text{range } ST = V$. 由于 $\text{range } ST \subset \text{range } S, \text{null } T \subset \text{null } ST$, 容易看出 $\text{range } S = V, \text{null } T = 0$, 即 S 是满射而 T 是单射。由于 S 和 T 都是 V 上的线性变换, 仅从单射或满射即可推出可逆, 故 S 和 T 都可逆。

\Leftarrow : 记 S, T 的逆分别为 S^{-1}, T^{-1} , 那么

$$T^{-1}S^{-1}ST = T^{-1}\text{Id}_V T = \text{Id}_V,$$

$$STT^{-1}S^{-1} = S\text{Id}_V S^{-1} = \text{Id}_V.$$

由此可见, ST 是可逆的, 且其逆为 $T^{-1}S^{-1}$.

Remark: 这里的结论告诉我们 $(ST)^{-1} = T^{-1}S^{-1}$, 这叫做“穿脱原理”, 即取逆的时候顺序要反过来。

6. 设 n 是整数, 且 $a_{ij} \in \mathbf{F}, i, j = 1, 2, \dots, n$, 证明下面的 (a) 和 (b) 等价:

- 齐次线性方程组

$$\sum_{k=1}^n a_{1k}x_k = 0,$$

$$\vdots,$$

$$\sum_{k=1}^n a_{nk}x_k = 0,$$

只有平凡解 $x_1 = x_2 = \dots = x_n = 0$.

- 对于每组 $c_1, c_2, \dots, c_n \in \mathbf{F}$, 方程组

$$\sum_{k=1}^n a_{1k}x_k = c_1,$$

$$\vdots,$$

$$\sum_{k=1}^n a_{nk}x_k = c_n,$$

都有解。

解答 1: 构造线性映射 $T: \mathbf{F}^n \rightarrow \mathbf{F}^n$,

$$T((x_1, x_2, \dots, x_n)) = \left(\sum_{k=1}^n a_{1k}x_k, \dots, \sum_{k=1}^n a_{nk}x_k \right).$$

那么 (a) 即是说 T 是单射, (b) 即是说 T 是满射, 由于 T 是 \mathbf{F}^n 上的线性变换, 这二者都等价于 T 可逆。

解答 2: 记 $v_i = (a_{i1}, a_{i2}, \dots, a_{in}) \in \mathbf{F}^n, i = 1, 2, \dots, n$, 那么 (a) 即是说 v_1, v_2, \dots, v_n 线性无关, (b) 即是说 v_1, v_2, \dots, v_n 是 \mathbf{F}^n 的张成组, 由于向量组的长度等于 \mathbf{F}^n 维数, 线性无关与张成组也是等价的, 都等价于该向量组是一组基。

2 多项式

2.1 习题选讲

1. 设 z_1, z_2, \dots, z_{m+1} 是 \mathbf{F} 中的不同元素, 并且 $w_1, w_2, \dots, w_{m+1} \in F$, 证明存在唯一一个多项式 $p \in \mathcal{P}_m(\mathbf{F})$ 使得对 $j = 1, 2, \dots, m+1$ 都有 $p(z_j) = w_j$.

解答: 定义一个映射 $T: \mathcal{P}_m(\mathbf{F}) \rightarrow \mathbf{F}^{m+1}, T(p) = (p(z_1), p(z_2), \dots, p(z_{m+1}))$, 显然 T 是线性映射, 我们来说明 T 是可逆的, 因此对于任意 $(w_1, w_2, \dots, w_{m+1}) \in \mathbf{F}^{m+1}$ 都存唯一 $p \in \mathcal{P}_m(\mathbf{F})$ 使得 $p(z_j) = w_j$.

- 为说明 T 是单射, 只需证明 $\text{null } T = \{p \in \mathcal{P}_m(\mathbf{F}) | p(z_1) = p(z_2) = \dots = p(z_{m+1}) = 0\} = 0$. 由代数学基本定理可知, 一个不超过 m 次的非零多项式最多有 m 个根 (计重数), 但这里有 $m+1$ 个根, 因此 p 只能恒为 0. 这就证得了 T 是单射。
- 还需要说明 T 是满射。由于 $\dim \mathcal{P}_m(\mathbf{F}) = m+1$, 故存在可逆映射 $S: \mathbf{F}^{m+1} \rightarrow \text{cal } \mathcal{P}_m(\mathbf{F})$. 考虑 $TS \in \mathcal{L}(\mathbf{F}^{m+1})$, 由于 T 是单射, 且 S 是单射, TS 也是单射。而 TS 是 \mathbf{F}^{m+1} 上的线性映射, 由单射即可推知是可逆映射。再由前一节的习题可知, T 也可逆, 因而满。(这么看来, 似乎不用单独证明 T 是单射了)

Remark 1: 一个构造性的证明已经在第二次讲义 (3.2.6) 中给出, 但是那里只说明存在并未说明唯一。

Remark 2: 定理 3.21 其实不一定非要原空间和像空间相同, 实际上只要它们的维数相同就可以了, 上面的证明过程就体现了这一点, 维数相同就可以用一个可逆映射同构过去。此外, 这个定理与推论 2.16/2.17 有相似之处。

3 特征值, 特征向量, 不变子空间

3.1 复习与思考

这一章里面考虑的都是线性变换, 即原空间与像空间是同一个空间。特征值与特征向量的定义具有清晰的几何意义, $Tu = \lambda u$ 意即 T 作用在 u 这个向量上, 就等于是把它拉伸了一个倍数而方向不变。为什么我们需要考虑不变子空间与特征子空间呢? 其实是为了对一个特定的线性变换寻找一个简单的表示。设 $T \in \mathcal{L}(V)$, 并且 V 能够写成 T 的不变子空间的直和 $U_1 \oplus U_2 \oplus \cdots \oplus U_k$, 选取 U_i 的基 $u_{i1}, u_{i2}, \dots, u_{in_i}$ 并合在一起组成 V 的基, 那么在这组基下 T 的矩阵表示一定是一个对角阵

$$\mathcal{M}(T) = \begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{kk} \end{bmatrix}.$$

这种表示已经相对很简单了, 称为分块对角阵。每一块内部是否还可以简化? 这涉及到不变子空间内部如何选取基。换言之, 我们可以把 T 给限制在不变子空间 U_i 上, 考虑一个更小空间上的线性映射如何简单地表示出来, 这就把一个较难的问题分解为了几个较简单的问题, 每个子问题的规模都变小了。如果我们能把 V 写成 T 的特征子空间的直和 $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$, 那么这样表示的矩阵不仅仅是分块对角阵, 直接就是对角阵了

$$\mathcal{M}(T) = \begin{bmatrix} \lambda_1 \mathbf{I}_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 \mathbf{I}_{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k \mathbf{I}_{n_k} \end{bmatrix}.$$

这种情况非常理想, 但可惜的是一般情况下不一定能这么分解 V . 在什么条件下能把空间分解为特征子空间的直和将是后面研究的重点问题。我们已经证明过一个向量不可能对应于两个不同的特征值, 因此 $V_{\lambda_i} \cap V_{\lambda_j} = 0, \lambda_i \neq \lambda_j$, 这说明 $V_{\lambda_i} + V_{\lambda_j}$ 必然是直和, 但两两直和不能推出全体加在一起也是直和。更一般地, 其实有 $V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_k}$ 总是直和 (目前不太容易说明), 但它们的和可能是 V 的一个真子空间, 无法包含 V 中所有的向量。

3.2 习题选讲

1. 设 $T \in \mathcal{L}(V)$, 证明若 V 的子空间 U_1, U_2, \dots, U_m 在 T 的作用下都是不变的, 那么 $U_1 + U_2 + \dots + U_m$ 在 T 下也是不变的。

解答: 对于任意 $u \in U_1 + U_2 + \dots + U_m$, 存在 $u_i \in U_i, i = 1, 2, \dots, m$ 使得 $u = u_1 + u_2 + \dots + u_m$. 那么 $Tu = Tu_1 + \dots + Tu_m$. 由于 U_i 在 T 下都是不变的, $Tu_i \in U_i$, 也因此 $Tu \in U_1 + U_2 + \dots + U_m$.

2. 设 $T \in \mathcal{L}(V)$, 证明任意一族在 T 下不变的子空间的交也是在 T 下不变的。

解答: 设 $U_\alpha, \alpha \in I$ 是 T 的一组不变子空间。对于任意 $v \in \cap_{\alpha \in I} U_\alpha$, 有 $Tu \in U_\alpha, \forall \alpha \in I$, 因此 $Tu \in \cap_{\alpha \in I} U_\alpha$, 即说明了 $\cap_{\alpha \in I} U_\alpha$ 仍然在 T 下不变。

Remark: 这里仍然要注意“任意”不是“可数”。

3. 证明或举反例: 若 V 的子空间 U 在 V 的每个算子下都是不变的, 则 $U = \{0\}$ 或 $U = V$.

解答: 假设 U 既不等于 $\{0\}$ 也不等于 V , 设 U 的一组基为 u_1, u_2, \dots, u_m , 那么显然有 $1 \leq m < \dim V$. 我们将其扩展为 V 的一组基 $u_1, \dots, u_m, v_1, \dots, v_l$, 定义一个线性算子 $T \in \mathcal{L}(V)$, $T(u_i) = v_1, i = 1, 2, \dots, m, T(v_j) = 0, j = 1, 2, \dots, l$, 那么在这个 T 的作用下 U 不再是不变子空间。因此, 若 V 的子空间 U 在 V 的每个算子下都是不变的, 则 $U = \{0\}$ 或 $U = V$.

4. 设 $S, T \in \mathcal{L}(V)$ 使得 $TS = ST$, 证明对每个 $\lambda \in \mathbf{F}$, $\text{null}(T - \lambda I)$ 在 S 下都是不变的。

解答: 只需证明若 $(T - \lambda I)v = 0$, 则 $(T - \lambda I)Sv = 0$. 而 $(T - \lambda I)Sv = TSv - \lambda Sv = STv - \lambda Sv = S(T - \lambda I)v = 0$, 故证得。

5. 定义 $T \in \mathcal{L}(\mathbf{F}^2)$ 如下

$$T(w, z) = (z, w).$$

求 T 的所有本征值和本征向量。

解答: 设 λ 是 T 的特征值, (a, b) 是对应的特征向量, 那么应有 $T(a, b) = \lambda(a, b)$, 即

$$b = \lambda a, a = \lambda b.$$

由此得 $b = \lambda^2 b$. 若 $b = 0$, a 也必须为 0, (a, b) 不是特征向量, 故只考虑 $b \neq 0$ 的情况, 这时可得 $\lambda = 1$ 或 -1 .

- 当 $\lambda = 1$ 时, $b = a$, 所有形如 $\{(x, x) | x \in \mathbf{F}, x \neq 0\}$ 的均为对应的特征向量;
- 当 $\lambda = -1$ 时, $b = -a$, 所有形如 $\{(x, -x) | x \in \mathbf{F}, x \neq 0\}$ 的均为对应的特征向量。

6. 设 n 是整数, 并设 $T \in \mathcal{L}(\mathbf{F}^n)$ 定义如下:

$$T(x_1, x_2, \dots, x_n) = (x_1 + \dots + x_n, \dots, x_1 + \dots + x_n).$$

求 T 的所有本征值和本征向量。

解答: 设 λ 是 T 的特征值, (x_1, \dots, x_n) 是对应的特征向量, 应有

$$\begin{cases} x_1 + \dots + x_n = \lambda x_1 \\ x_1 + \dots + x_n = \lambda x_2 \\ \vdots \\ x_1 + \dots + x_n = \lambda x_n \end{cases} \Rightarrow \begin{cases} (1 - \lambda)x_1 + \dots + x_n = 0 \\ x_1 + \dots + x_n = 0 \\ \vdots \\ x_1 + \dots + (1 - \lambda)x_n = 0 \end{cases}$$

先考虑何时该方程有非零解, 这可以等价地去考虑

$$(1 - \lambda, 1, \dots, 1), (1, 1 - \lambda, \dots, 1), \dots, (1, 1, \dots, 1 - \lambda)$$

何时线性相关。显然当 $\lambda = 0$ 时线性相关, 此外当 $\lambda = n$ 时, 它们的和为 0, 因而也线性相关。我们先考虑这两种情况。

- 如果 $\lambda = 0$, 只要 $x_1 + x_2 + \dots + x_n = 0$ 即可;
- 如果 $\lambda = n$, 那么所有形如 $\{(x, x, \dots, x) | x \in \mathbf{F}, x \neq 0\}$ 的均为对应的特征向量。

由于 $\dim V_0 = n - 1, \dim V_n = 1, V_0 \cap V_n = \{0\}$, 可见 $V = V_0 \oplus V_n$, 因此没有其他的特征值、特征向量了。