

碎碎念之线性变换在给定基下的矩阵表示

问题 0.1. 在 $\mathbb{R}^3$ 中考虑以下两组向量：

$$u_1 = (1, 0, 1)^\top, u_2 = (2, 1, 0)^\top, u_3 = (1, 1, 1)^\top$$

和

$$v_1 = (1, 0, 0)^\top, v_2 = (1, 1, 0)^\top, v_3 = (1, 1, 1)^\top$$

可以验证这两组向量都是 $\mathbb{R}^3$ 的基。假定线性变换 $\mathcal{T}$ 把基 u_1, u_2, u_3 映到基 v_1, v_2, v_3 。

1. 求基 u_1, u_2, u_3 到基 v_1, v_2, v_3 的过渡矩阵；
2. 求 $\mathcal{T}$ 在基 u_1, u_2, u_3 下的矩阵；
3. 求 $\mathcal{T}$ 在基 v_1, v_2, v_3 下的矩阵；
4. 求 $\mathcal{T}^2(u_1) = \mathcal{T}(\mathcal{T}(u_1))$ ；
5. 求 $(1, 2, 3)^\top$ 在基 u_1, u_2, u_3 下的坐标。

在解答这道题之前，我们先回顾一下在给定基向量下，一个线性变换如何用矩阵表示。

教材中以及老师给出的答案是：对于映射 $\mathcal{T}: \mathbb{V} \rightarrow \mathbb{W}$,

$$\mathcal{T}(x_j) = Ax_j = \sum_{i=1}^m a_{ij} y_i \quad (0.1)$$

其中 x_i 是 $\mathbb{V}$ 空间的一组基， y_i 是 $\mathbb{W}$ 空间的一组基。

这个定义看起来是不是像天书？完全不知所云，看完也不知道 A 是干嘛的，以及怎么找。

那么我接下来给出一个较为直观的理解方式。

命题 0.1. 我们约定 $V = [v_1 | \dots | v_n]$ ，其中 $(v_1, \dots, v_n)$ 是在自然基意义下的坐标。则：

$$\mathcal{T}(V) = V \cdot A_V \quad (0.2)$$

其中 A_V 即为映射 $\mathcal{T}$ 在给定基 V 下的矩阵表示。

看着也很抽象，那我来解释一下。可能大家最习惯的是看

$$\mathcal{T}(x) = Ax$$

左乘一个矩阵非常直观对吧？ A 很容易理解成对一个输入向量进行一个作用，输出的就是变换的结果向量。

那么(0.2)这里右乘一个矩阵 A_V 要怎么理解呢？右乘的本质是对 V 的列做线性组合， A_V 记录了 V 的基底（即列向量）的组装过程，而这个组装过程恰恰就是一个结果向量在这个基下的分解。

为什么要对 $\mathcal{T}(V)$ 做分解呢？我们由定义知 $\mathcal{T}(V) = W$ ，我们想研究在这组给定基 V 下该映射 $\mathcal{T}$ 的矩阵表示，其实只需要知道 V 中的每个基在变换后得到的实体向量 W 在该 V 基下的坐标是多少就可以了。这个实体向量就是 $\mathcal{T}(V)$ ，而且是在自然基^{0.1}意义下的。 A_V 的每个列向量就是 $\mathcal{T}(V)$ 最终在 V 基下的坐标。换句话说整个参考系换到 V 后，它的标准基在变换 $\mathcal{T}$ 作用后的结果就是 A_V 的每个列向量。

承认了(0.2)的合理性后，我们开始来解答这个题。

答案. 我们先不考虑过渡矩阵的事，它会在后面自然而然地呈现出来。

1. 考虑 $\mathcal{T}$ 在自然基下的**标准矩阵表示**。由定义知：

$$\mathcal{T}(U) = V \quad (0.3)$$

我们记 $\mathcal{T}$ 在自然基下的标准矩阵为 A ，有 $AU = V$ ，故

$$A = VU^{-1} \quad (0.4)$$

2. 考虑 $\mathcal{T}$ 在 U 的**基底下矩阵表示**。由(0.2)知

$$\mathcal{T}(U) = U \cdot A_U$$

而同时有(0.3)，故

$$A_U = U^{-1}V$$

3. 考虑 $\mathcal{T}$ 在 V 的**基底下矩阵表示**。由(0.2)知

$$\mathcal{T}(V) = V \cdot A_V \quad (0.5)$$

而同时有(0.4)，故

$$\mathcal{T}(V) = A \cdot V = (VU^{-1})V = V(U^{-1}V)$$

对比(0.5)，我们得到

$$A_V = U^{-1}V$$

我们会惊讶地发现， A_U 和 A_V 居然奇迹般相等了，这是个很好的结论，但也仅限于在 U 或 V 这两个基底下的矩阵表示。

既然有这么好的一致性，那我们就自然定义过渡矩阵：

定义 0.1. 在 n 维空间 $\mathbb{V}$ 中存在两组基 $X = [x_1 | \dots | x_n], Y = [y_1 | \dots | y_n]$ ，若满足：

$$Y = XP$$

则称 P 为基变换 X 到 Y 的过渡矩阵。

0.1. 自然基即一组标准正交基，这里在 $\mathbb{R}^3$ 中的自然基是 $(1, 0, 0), (0, 1, 0), (0, 0, 1)$

我们先不急着解题目中的最后两问。上述推导 A_U 和 A_V 还是需要一定的洞察力，但其实还有更简单的方法。

这里需要引入一个强大的工具：线性变换的换基公式！

定理 0.1. 对于映射 $T: \mathbb{V} \rightarrow \mathbb{W}$ ，其在基 V 到基 W 意义下的矩阵表示为 A （意味着 A 的输入是一个 V 基底的坐标，输出是映射后的 W 基底的坐标）。现对于新基 $V' = VP$ ，新基 $W' = WQ$ ，则在基 V' 和 W' 意义下的矩阵表示 B 为：

$$B = Q^{-1}AP \quad (0.6)$$

利用换基公式(0.6)，我们直接对(0.4)中得到的 A 作变换。

答案. 注意 A 是自然基到自然基意义下变换 $\mathcal{T}$ 的矩阵表示。

1. 考虑 A_U ，即基 U 到基 U 意义下变换 $\mathcal{T}$ 的矩阵表示。换基的两个过渡矩阵都是 U ，则：

$$A_U = U^{-1}AU$$

将(0.4)带入，得到

$$\begin{aligned} A_U &= U^{-1}VU^{-1}U \\ &= U^{-1}V \end{aligned}$$

结果一致。

2. 考虑 A_V ，即基 V 到基 V 意义下变换 $\mathcal{T}$ 的矩阵表示。换基的两个过渡矩阵都是 V ，则：

$$A_V = V^{-1}AV$$

将(0.4)带入，得到

$$\begin{aligned} A_V &= V^{-1}VU^{-1}V \\ &= U^{-1}V \end{aligned}$$

结果一致。

3. 另外考虑 $A_{U \rightarrow V}$ ，即基 U 到基 V 意义下变换 $\mathcal{T}$ 的矩阵表示。换基的两个过渡矩阵分别是 U 和 V 。这个变换很有意思，由于 $\mathcal{T}$ 本身就把基 U 映到基 V ，考虑 U 意义下的一个单位向量变换后是 V 意义下的单位向量，其坐标不会发生改变，因此 $A_{U \rightarrow V} = I$ ，我们验证一下：

$$\begin{aligned} A_{U \rightarrow V} &= V^{-1}VU^{-1}U \\ &= I \end{aligned}$$

结果一致。

上述推导更为简洁直观。下面我们带入数据完整地解答这道题。

答案.

1. 求过渡矩阵 P :

$$P = U^{-1}V = \frac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 1 \end{bmatrix}$$

2. 求 A_U :

$$A_U = U^{-1}V = \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 1 \end{bmatrix}$$

3. 求 A_V :

$$A_V = U^{-1}V = \begin{bmatrix} 1/2 & -1/2 & 0 \\ 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 1 \end{bmatrix}$$

4. 求 $\mathcal{T}^2(u_1) = \mathcal{T}(\mathcal{T}(u_1))$:

由题意知 $\mathcal{T}(u_1) = v_1$, 故 $\mathcal{T}^2(u_1) = \mathcal{T}(v_1)$

由(0.2)知

$$\mathcal{T}(v_1) = Av_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$$

故

$$\mathcal{T}^2(u_1) = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$$

5. 求新坐标 $[x]_U$:

类似于(0.2), 我们有

$$x = U \cdot [x]_U$$

注意 x 是自然基意义下的坐标向量, 故

$$[x]_U = U^{-1}x = \frac{1}{2} \begin{bmatrix} 1 & -2 & 1 \\ 1 & 0 & -1 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 3 \end{bmatrix}$$

注记 0.1. 2025.4.16 补充：

根据很多学长的反馈，事实上(0.2)的写法也非常的反人类。这里作出三点补充：

1. (0.2)的推法不应该先从矩阵开始，最直接的是从向量角度考虑：

对于一个实体向量 v ，其在给定基底 V 下的坐标 x 满足

$$v = V \cdot x$$

之后考虑映射 T 在给定基底 V 下的矩阵表示，即只需求基向量经过映射后在 V 下的坐标。

$$T(v_j) = V \cdot A_j$$

其中 v_j 为第 j 个基向量， A_j 为矩阵 A 的第 j 列。

这些坐标按列拼成的矩阵 A 即为映射 T 在给定基底 V 下的矩阵表示。

$$T(V) = V \cdot A$$

故有(0.2)。

2. 更一般地，不局限于 V 基映到 V 基，对于任意基底 P 下表示的向量，如果知道映射规则（也即知道 $T(P)$ ，这是一个实体向量），那么其在另一个基底 Q 下输出的矩阵表示 A' （输入的是 P 基下的坐标，输出的是映射 T 后在 Q 基下的坐标）即为

$$T(P) = Q \cdot A$$

其中 A 的第 j 列是基底 P 下的第 j 个基向量经映射 T 后在基底 Q 下坐标。

3. 对于(0.1)，实际上也是可以理解的：

$$Ax_j = \sum_{i=1}^m a_{ij} y_i$$

其中 x_j 为输入基 X 意义下的坐标， y_i 为输出基 Y 下的第 i 个基向量。

a_{ij} 即为基 Y 下的每个坐标分量，一列拼起来得到的一个列向量 A_j 是 x_j 在基 Y 下的坐标，也即矩阵 A 的第 j 列。

从考试的角度来说，求一个线性映射 T 在某一个基底 V 下的矩阵表示 A ，其实就是把基向量按照规则 T 映射后得到的向量，求这个向量在基 V 下的坐标。把这些坐标按列拼成的矩阵即为所求。