§1.2 线性变换及其矩阵表示

定义: T 称为由 V^n 到 V^m 的**变换 (映射)**, 如果 T 将 V^n 中的向量映射到 V^m 中的向量, 写作

$$T : \alpha \in V^n \to \beta = T\alpha \in V^m$$

其中 β 为 α 在 T 下的**像**, α 称为 β 的**原像**.

变换的例子

- $T : \alpha = (x_1, x_2)^T \in \mathbb{R}^2 \to \beta = (0, x_2)^T \in \mathbb{R}^2$
- $T : \alpha = (x_1, x_2)^T \in \mathbb{R}^2 \to \beta = (2x_1, 2x_2)^T \in \mathbb{R}^2$
- $T : \alpha = p(t) \in \mathbf{P}(t) \to \beta = (p(t))^2 \in \mathbf{P}(t)$

线性变换

定义: T 为由 V^n 到 V^m 的变换, 如果对于任意的 $k \in F$, $\alpha, \beta \in V^n$, 都有

$$T(\alpha + \beta) = T\alpha + T\beta, \quad T(k\alpha) = kT\alpha,$$

即变换 T 与线性运算可交换, 则 T 是线性变换.

特别的, 当 T 是 V^n 到自身的一个线性变换, 则称 T 是 V^n 的线性变换.

例 1. 给定 $A \in F^{m\times n}$, 定义由 V^n 到 V^m 的变换 T 为

$$T : x \in F^n \to y = Ax \in F^m.$$

由矩阵运算的性质得知, 易证 T 是一个线性变换.

例 2. 给定 $P \in F^{m\times m}$ 和 $Q \in F^{n\times n}$, 定义由 V^n 到 V^m 的变换 T 为

$$T : X \in F^{m\times n} \to Y = PXQ \in F^{m\times n}.$$

由矩阵运算的性质得知, 易证 T 是一个线性变换.

例 3. 对于 $\mathbf{P}_n(t)$ 中的多项式求导运算 $\frac{d}{dt}$, 记为 D, 即

$$Dp(t) = \frac{d}{dt}p(t), \quad p(t) \in \mathbf{P}_n(t)$$

因求导运算和线性运算可以交换顺序, 可知 D 是 $\mathbf{P}_n(t)$ 的一个线性变换.

例 4. V 的

- 恒等变换 $I: I\alpha = \alpha \forall \alpha \in V$
- 零变换 $O: O\alpha = 0, \forall \alpha \in V$

都是 V 的线性变换.

线性变换的一些简单性质:

- (1). $O\alpha = 0, \forall \alpha \in V$.
  $T0 = 0$, $T(-\alpha) = -T\alpha, \forall \alpha \in V$.
- (2). $T(\sum_{i=1}^{r} k_i\alpha_i) = \sum_{i=1}^{r} k_i T\alpha_i$, 即任意一组向量的线性组合的像, 等于它们的像的线性组合.
- (3). 一组线性相关的向量 $\alpha_1, \alpha_2, ..., \alpha_r$, 它们在 T 下的像 $T\alpha_1, T\alpha_2, ..., T\alpha_r$ 也线性相关.
  注意: 线性无关的一组向量在 T 下的像可能是线性相关的, 例如零变换把线性无关的向量都映射为零向量.

线性变换的矩阵表示

设 T 是 V^n 到 V^m 的线性变换, 在 V^n 和 V^m 中分别取基$\mathcal{B}_\alpha = \{\alpha_1, \alpha_2, ..., \alpha_n\}$
和$\mathcal{B}_\beta = \{\beta_1, \beta_2, ..., \beta_m\}$, 则 α_j 的像 $T\alpha_j$ $(1 \leq j \leq n)$ 可由基 $\mathcal{B}_\beta$ 唯一线性表出:

$$T\alpha_j = \sum_{i=1}^m \beta_i a_{ij} = (\beta_1, \beta_2, ..., \beta_m) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

将 $T\alpha_j$ 按 $j = 1, 2, ..., n$ 的顺序排列, 则有

$$(T\alpha_1, .., T\alpha_n) = (\beta_1, .., \beta_m)\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

令 $T\mathcal{B}_\alpha := (T\alpha_1, T\alpha_2, ..., T\alpha_n)$, 则上式可简写为

$$T\mathcal{B}_\alpha = \mathcal{B}_\beta A$$

则这称为线性变换 T 的矩阵表示, 其中 A 称为 T 在基偶 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$ 下的矩阵.
特别的, 若 $V^n = V^m$ 且 $\mathcal{B}_\alpha = \mathcal{B}_\beta$, 则 $T\mathcal{B}_\alpha = \mathcal{B}_\alpha A$, 此时称 n 阶方阵 A 为 T 在基 $\mathcal{B}_\alpha$ 下的矩阵.

例 6. 求 $\mathbf{P}_n(t)$ 的线性变换 $D = \frac{d}{dt}$ 在基 $\mathcal{B} = \{1, t, ..., t^n\}$ 下的矩阵.

解: 由 $D\mathcal{B} = \mathcal{B}A$, 其中 $D\mathcal{B} = (0, 1, 2t, ..., nt^{n-1})$,
即 $D\mathcal{B} = (0, 1, 2t, ..., nt^{n-1}) = (1, t, t^2, ..., t^n)A$,
由 $jt^{j-1} = \sum_{i=1}^{n} a_{ij}t^i$, 知 $a_{j-1,j} = j$ 而其他 a_{ij} 为零.

可得 $A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & n \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$.

例 7 求 $\mathbf{P}_2(t)$ 到 $\mathbf{P}_3(t)$ 的线性变换 J

$$J(p(t)) = \int_0^t p(t)dt$$

基偶 $\{\mathcal{B}_1 = \{1, t, t^2\}, \mathcal{B}_2 = \{1, t, \frac{t^2}{2}, \frac{t^3}{3}\}\}$ 下的矩阵.

解: 由 $J\mathcal{B}_1 = \mathcal{B}_2 A$, 其中 $J\mathcal{B}_1 = (t, \frac{t^2}{2}, \frac{t^3}{3})$, 即
$J\mathcal{B}_1 = (t, \frac{t^2}{2}, \frac{t^3}{3}) = (1, t, \frac{t^2}{2}, \frac{t^3}{3})A$,

可得 $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

由 T 的矩阵表示，确定 T 的像的坐标.

设 $\alpha \in V^n$, 则可由 $\mathcal{B}_\alpha$ 线性表出,

$$\alpha = \sum_{i=1}^{n} x_i \alpha_i = \mathcal{B}_\alpha x$$

即α 在 $\mathcal{B}_\alpha$ 下的坐标为 x. 对于线性变换 T,

$$T\alpha = \sum_{i=1}^{n} x_i T\alpha_i = T\mathcal{B}_\alpha x$$

将 $T\mathcal{B}_\alpha$ 用 V^m 的基 $\mathcal{B}_\beta$ 表出, 若 $T\mathcal{B}_\alpha = \mathcal{B}_\beta A$, 则

$$T\alpha = \mathcal{B}_\beta A x,$$

即$T\alpha$ 在 $\mathcal{B}_\beta$ 下的坐标为 Ax.

定理: 设 $\mathcal{B}_\alpha = \{\alpha_1, \alpha_2, ..., \alpha_n\}$ 和 $\mathcal{B}_\beta = \{\beta_1, \beta_2, ..., \beta_n\}$ 分别是 V^n 和 V^m 的基, 对于给定的 $m \times n$ 矩阵 $A = [a_{ij}]$, 则存在唯一的从 V^n 到 V^m 的线性变换 T, 使得它在 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$ 下的矩阵是 A, 即 $T\mathcal{B}_\alpha = \mathcal{B}_\beta A$.

证明: 先证明线性变换 T 的存在性.
对于任意 $\alpha \in V^n$, 设其在基 $\mathcal{B}_\alpha$ 下的坐标为 x, 即 $\alpha = \mathcal{B}_\alpha x$. 现定义变换 T 将 α 映为 $\mathcal{B}_\beta Ax$, 即

$$T : \alpha = \mathcal{B}_\alpha x \to \beta = \mathcal{B}_\beta Ax$$

我们仍需验证 T 是线性变换.

证明 (续): 验证 T 是线性变换.

设 V^n 中的 $\alpha = \mathcal{B}_\alpha x$ 和 $\gamma = \mathcal{B}_\alpha z$, 而 $k, l \in F$, 则

$$\begin{aligned} T(k\alpha + l\gamma) &= T(k\mathcal{B}_\alpha x + l\mathcal{B}_\alpha z) \\ &= T(\mathcal{B}_\alpha(kx + lz)) \\ &= \mathcal{B}_\beta A(kx + lz) \\ &= k\mathcal{B}_\beta Ax + l\mathcal{B}_\beta Az \\ &= kT(\mathcal{B}_\beta x) + lT(\mathcal{B}_\beta z) \\ &= kT\alpha + lT\gamma \end{aligned}$$

则 T 是线性变换, 且由 T 定义易得 $T\mathcal{B}_\alpha = \mathcal{B}_\beta A$.

证明 (续): 再证明线性变换 T 的唯一性.

如果存在两个线性变换 T_1 和 T_2, 满足

$$T_1\mathcal{B}_\alpha = T_2\mathcal{B}_\alpha = \mathcal{B}_\beta A.$$

那么线性变换 $\tilde{T} = T_1 - T_2$ 满足对于某个 $\tilde{A}$ 有

$$\tilde{T}\mathcal{B}_\alpha = \{0, 0, ..., 0\} = \mathcal{B}_\beta \tilde{A}$$

这里 $\tilde{A}$ 只能为零, 即 $\tilde{T}$ 将任何向量映为 0. 从而

$$T_1\alpha = (T_2 + \tilde{T})\alpha = T_2\alpha + \tilde{T}\alpha = T_2\alpha$$

对于任意 α 成立, 因此映射唯一.

V^n 到 V^m 的线性变换 T, 在给定的基偶 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$ 下, 对应一个矩阵 A. 反之, 对于同样的空间和基偶, 如果给定矩阵 A, 则对应一个线性变换 T.

接下来的问题:

- 如果取 V^n 和 V^m 的另一组基 $\{\mathcal{B}_{\alpha'}, \mathcal{B}_{\beta'}\}$, 那么 T 对应另外一个矩阵 B, 那么矩阵 A 和 B 之间有什么关系?
- 怎样选取 V^n 和 V^m 的基, 使得 T 的矩阵表示尽可能简单?

设 n 阶方阵 P 是基 $\mathcal{B}_\alpha$ 到 $\mathcal{B}_{\alpha'}$ 的变换矩阵, 而 n 阶方阵 Q 是基 $\mathcal{B}_\beta$ 到 $\mathcal{B}_{\beta'}$ 的变换矩阵, $m \times n$ 矩阵 A, B 分别是 T 在基偶 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$ 和 $\{\mathcal{B}_{\alpha'}, \mathcal{B}_{\beta'}\}$ 下的矩阵, 那么由关系式

$$\mathcal{B}_{\alpha'} = \mathcal{B}_\alpha P, \quad \mathcal{B}_{\beta'} = \mathcal{B}_\beta Q,$$
$$T\mathcal{B}_\alpha = \mathcal{B}_\beta A, \quad T\mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'} B$$

可以推出 $\mathcal{B}_\beta AP = T\mathcal{B}_\alpha P = T\mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'} B = \mathcal{B}_\beta QB$

$$\mathcal{B}_\beta(AP - QB) = 0$$

因 $\mathcal{B}_\beta$ 是基, 则 $AP = QB, A = QBP^{-1}, B = Q^{-1}AP$

如果$A = QBP^{-1}$ 或 $B = Q^{-1}AP$, 其中 P, Q 为可逆方阵, 那么称 A 和 B 是相抵 (或等价)的. 如上证明了, 一个 V^n 到 V^m 的线性变换在不同基偶下的矩阵是相抵的.

假如 $V^m = V^n$, $\mathcal{B}_\alpha = \mathcal{B}_\beta$, $\mathcal{B}_{\alpha'} = \mathcal{B}_{\beta'}$, 那么 $Q = P$, 则有$A = PBP^{-1}$. 此时方阵 A 与 B 是相似的.
即一个 V^n 到自身的线性变换在不同基偶下的矩阵是相似的.

之前的第二个问题等价于: 与 A 相抵 (或相似) 的最简单的矩阵是什么?

定义: 设 T 是从 V^n 到 V^m 的线性变换, 则

$$\begin{aligned} N(T) &:= \{\alpha \in V^n | T\alpha = 0\} \\ R(T) &:= \{\beta \in V^m | \beta = T\alpha, \alpha \in V^n\} \end{aligned}$$

分别称为 T 的核和 T 的值域.

- $N(T)$ 是 V^n 的一个子空间, 也被称为 T 的零空间, 其维数称为 T 的零度, 记作null T ;
- $R(T)$ 是 V^m 的一个子空间, 也被称为 T 的值空间, 其维数称为 T 的秩, 记作rank T .

一个线性变换 T 将 V^n 中的任一向量 α 映射为 V^m 的向量 $T\alpha$. V^n 的每一个向量都有在 T 下的像, 但未必 V^m 的每一个向量都有原像. 这发生在$\mathrm{rank}\, T < m$的时候, 例如:

- $T : \alpha = (x_1, x_2)^T \in \mathbb{R}^2 \to \beta = (0, x_2)^T \in \mathbb{R}^2$
- $T : \alpha = p(t) \in \mathbf{P}_n(t) \to \beta = \frac{d}{dt} p(t) \in \mathbf{P}_n(t)$

但可以通过重新定义合适的空间 $T : V^n \to R(T)$, 使 $R(T)$ 中每一个向量都有原像

- $T : \alpha = (x_1, x_2)^T \in \mathbb{R}^2 \to \beta = (0, x_2)^T \in \{0\} \times \mathbb{R}$
- $T : \alpha = p(t) \in \mathbf{P}_n(t) \to \beta = \frac{d}{dt} p(t) \in \mathbf{P}_{n-1}(t)$

定理: 设 T 是从 V^n 到 V^m 的线性变换, 则

$$\operatorname{null} T + \operatorname{rank} T = n$$

证明: 若 $\operatorname{null} T = k$, 并设 $\mathcal{B}_0 = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$ 是 $N(T)$ 的一组基. 因为 $N(T) \subset V^n$, 所以可以把这 k 个向量扩充为 V^n 的基
$\mathcal{B}_1 = \{\alpha_1, \alpha_2, \ldots, \alpha_k, \alpha_{k+1}, \ldots, \alpha_n\}$.
我们将证明 $\{T\alpha_{k+1}, T\alpha_{k_2}, \ldots, T\alpha_n\}$ 是 $R(T)$ 的一组基, 从而得到 $\operatorname{rank} T = n - k$.

证明 (续): 任取 $\alpha \in V^n$, 则 α 可由 $\mathcal{B}_1$ 线性表出: $\alpha = \sum_{i=1}^n b_i\alpha_i$. 又因为 $\alpha_i \in N(T)$ $(1 \leq i \leq k)$, 则

$$T\alpha = \sum_{i=1}^{n} b_i T\alpha_i = \sum_{i=k+1}^{n} b_i T\alpha_i.$$

即 $R(T)$ 中的向量 $T\alpha$ 可由 $\{T\alpha_{k+1}, T\alpha_{k_2}, \ldots, T\alpha_n\}$ 线性表出.
仍需证明 $\{T\alpha_{k+1}, T\alpha_{k_2}, \ldots, T\alpha_n\}$ 线性无关.

证明 (续): 假如存在 $c_{k+1}, c_{k_2}, \ldots, c_n \in F$ 使得

$$\sum_{i=k+1}^{n} c_i T\alpha_i = 0.$$

则有 $T(\sum_{i=k+1}^{n} c_i\alpha_i) = 0$, 从而 $\sum_{i=k+1}^{n} c_i\alpha_i \in N(T)$.

于是 $\sum_{i=k+1}^{n} c_i\alpha_i$ 可由 $\mathcal{B}_0$ 线性表出, 同时也由 $\alpha_{k+1}, \ldots, \alpha_n$ 线性表出. 但 $\mathcal{B}_0$ 与 $\alpha_{k+1}, \ldots, \alpha_n$ 线性无关, 故 $\sum_{i=k+1}^{n} c_i\alpha_i = 0$, 即 $c_{k+1} = \ldots = c_n = 0$. 于是 $\{T\alpha_{k+1}, T\alpha_{k_2}, \ldots, T\alpha_n\}$ 线性无关. 综上, 这是 $R(T)$ 的一组基, 则定理得证.

注:
若 $\{\alpha_1, \alpha_2, ..., \alpha_k\}$ 是 $N(T)$ 的一组基, 并被扩充为 V^n 的基 $\mathcal{B}_\alpha = \{\alpha_{k+1}, \alpha_{k+2}, ..., \alpha_n, \alpha_1, \alpha_2, ..., \alpha_k, \}$.
由定理知 $\{T\alpha_{k+1}, T\alpha_{k+2}, ..., T\alpha_n\}$ 是 $R(T)$ 的一组基, 也可以被扩充为 V^m 中的一组基
$\mathcal{B}_\beta = \{T\alpha_{k+1}, T\alpha_{k+2}, ..., T\alpha_n, \beta_1, \beta_2, ..., \beta_{k+m-n}\}$.

则 T 在基偶 $\{\mathcal{B}_\alpha, \mathcal{B}_\beta\}$ 下的矩阵是

$$\begin{pmatrix} I_{(n-k)} & O_{(n-k)\times k} \\ O_{(m+k-n)\times(n-k)} & O_{(m+k-n)\times k} \end{pmatrix}.$$

其中 I_{n-k} 为 $(n-k)$ 阶单位阵, O 为零矩阵.

再次回顾例子

- $T : \alpha = (x_1, x_2)^T \in \mathbb{R}^2 \to \beta = (0, x_2)^T \in \mathbb{R}^2$
- $T : \alpha = p(t) \in \mathbf{P}_n(t) \to \beta = \frac{d}{dt}p(t) \in \mathbf{P}_n(t)$

第一个例子中 $N(T) = \mathrm{span}\{(1,0)^T\}$, $\mathrm{null}\, T = 1$. 对应的 $\mathrm{rank}\, T = 2 - 1 = 1 < 2$.
第二个例子中 $N(T) = \mathrm{span}\{1\}$, $\mathrm{null}\, T = 1$. 对应的 $\mathrm{rank}\, T = (n+1) - 1 = n < n+1$.

另, 虽 $\dim N(T) + \dim R(T) = \mathrm{null}\, T + \mathrm{rank}\, T = n$
但 $N(T) + R(T)$ 不一定等于 V^n.