第 30 讲　线性变换与换基

2019 年 12 月 25 日

定义 1. 设 $T: V \to W$ 是从线性空间 V 到线性空间 W 的映射，如果 T 满足

(a) $T(\boldsymbol{\alpha}+\boldsymbol{\beta}) = T(\boldsymbol{\alpha}) + T(\boldsymbol{\beta})$, $\forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in V$,

(b) $T(c\boldsymbol{\alpha}) = cT(\boldsymbol{\alpha})$, $\forall c \in \mathbb{R}$, $\boldsymbol{\alpha} \in V$,

则称 T 是线性变换 (linear transformation)，V, W 分别称为 T 的出发域和到达域。

注 1. 性质 *(a)(b)* 等价于

(c) $T(c\boldsymbol{\alpha}+d\boldsymbol{\beta}) = cT(\boldsymbol{\alpha}) + dT(\boldsymbol{\beta})$, $\forall c, d \in \mathbb{R}$, $\boldsymbol{\alpha}, \boldsymbol{\beta} \in V$.

在 *(c)* 中令 $c = d = 1$ 得 *(a)*, 令 $d = 0$ 得 *(b)*. 反之, 从 *(a)* 推出 $T(c\boldsymbol{\alpha}+d\boldsymbol{\beta}) = T(c\boldsymbol{\alpha})+T(d\boldsymbol{\beta})$, 再由 *(b)* 知 $T(c\boldsymbol{\alpha}) = cT(\boldsymbol{\alpha})$, $T(d\boldsymbol{\beta}) = dT(\boldsymbol{\beta})$.

命题 1. 若 T 是线性变换，则

(i) $T(\mathbf{0}) = \mathbf{0}$,

(ii) $T(c_1\boldsymbol{\alpha}_1 + \ldots + c_k\boldsymbol{\alpha}_k) = c_1T(\boldsymbol{\alpha}_1) + \ldots + c_kT(\boldsymbol{\alpha}_k)$, $\forall c_1, \ldots, c_k \in \mathbb{R}$, $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k \in V$.

证明. (i) $T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0}) + T(\mathbf{0})$, 故 $T(\mathbf{0}) = \mathbf{0}$. 注意，括号中的 $\mathbf{0}$ 是出发域中的零向量，不在括号中的 $\mathbf{0}$ 是到达域中的零向量。或者，$T(\mathbf{0}) = T(0\mathbf{0}) = 0T(\mathbf{0}) = \mathbf{0}$. (ii) 重复使用 (a)(b) 即可。 □

例 1. 设 $A \in M_{m\times n}(\mathbb{R})$, 则 $\mathscr{A}: \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto A\boldsymbol{x}$ 是线性变换。这是因为矩阵乘法满足 $A(\boldsymbol{x}_1 + \boldsymbol{x}_2) = A\boldsymbol{x}_1 + A\boldsymbol{x}_2$, $A(c\boldsymbol{x}) = c(A\boldsymbol{x})$.

命题 2. 若 $\varphi: \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto \varphi(\boldsymbol{x})$ 是线性变换，则存在唯一的矩阵 $A \in M_{m\times n}(\mathbb{R})$ 使得 $\varphi(\boldsymbol{x}) = A\boldsymbol{x}$, $\forall \boldsymbol{x} \in \mathbb{R}^n$.

证明. 由命题 1(ii) 知，若已知线性变换在某些向量处的取值，则线性 (linearity) 决定了这个线性变换在这些向量的任意线性组合处的取值。取 $\mathbb{R}^n$ 的自然基 $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$，并令 $A = \begin{bmatrix} \varphi(\boldsymbol{e}_1) & \cdots & \varphi(\boldsymbol{e}_n) \end{bmatrix}$，则 $A \in M_{m\times n}(\mathbb{R})$ 且对任意的 $\boldsymbol{x} = (x_1, \ldots, x_n)$，下式成立

$$\varphi(\boldsymbol{x}) = \varphi(x_1\boldsymbol{e}_1 + \ldots + x_n\boldsymbol{e}_n) = x_1\varphi(\boldsymbol{e}_1) + \ldots + x_n\varphi(\boldsymbol{e}_n) = \begin{bmatrix} \varphi(\boldsymbol{e}_1) & \cdots & \varphi(\boldsymbol{e}_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A\boldsymbol{x}.$$

A 的存在性证毕，还需证明满足条件的矩阵 A 唯一。若 A 满足 $\varphi(\boldsymbol{x}) = A\boldsymbol{x}$ 对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ 成立，则 A 的第 j 列等于 $A\boldsymbol{e}_j = \varphi(\boldsymbol{e}_j)$，即 A 由 φ 唯一确定。 □

注 2. *例 1 和命题 2 合在一起说明了从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换与 $m \times n$ 阶实矩阵一一对应。例如，从 $\mathbb{R}$ 到 $\mathbb{R}$ 的线性变换只有 $f(x) = cx$ ($c \in \mathbb{R}$).*

例 2. $f : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto x + 2y + 3z$ 是线性变换。事实上，f 由矩阵 $A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ 诱导，是 $\mathbb{R}^3$ 上的线性函数。一般的，从 V 到 $\mathbb{R}$ 的线性变换称为 V 上的线性函数。

例 3. 设 $\mathcal{C}^1(\mathbb{R})$ 是 $\mathbb{R}$ 上连续可微函数构成的线性空间，$\mathcal{C}(\mathbb{R})$ 是 $\mathbb{R}$ 上连续函数构成的线性空间，则求导运算 $D : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}(\mathbb{R})$, $F(x) \mapsto F'(x)$ 是线性变换。这是因为导数有如下性质：$(F + G)' = F' + G'$, $(cF)' = cF'$ ($c \in \mathbb{R}$).

例 4. $\int : \mathcal{C}(\mathbb{R}) \to \mathcal{C}^1(\mathbb{R})$, $f(x) \mapsto \int_0^x f(t)dt$ 是线性变换。

例 5. 设 $\mathcal{F}(\mathbb{R})$ 是实轴上 (实值) 函数构成的线性空间，a 是某固定的实数，则 $\mathrm{ev}_a : \mathcal{F}(\mathbb{R}) \mapsto \mathbb{R}$, $f(x) \mapsto f(a)$ 是线性变换。泛函分析里会把 $f(a)$ 写成 $f(x)$ 与 delta 函数 $\delta_a(x)$ 的积分，

例 6 (PS3.5, 30)**.** 设 $A = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$，则 $T : M_3(\mathbb{R}) \to M_3(\mathbb{R})$, $X \mapsto AX$ 是线性变换。事实上，$X \mapsto XA$, $X \mapsto AX - XA$, $X \mapsto AXA$ 都是 $M_3(\mathbb{R})$ 上的线性变换。

例 7. 考虑满足斐波那契递归公式的数列构成的集合

$$\mathscr{F} = \{(a_0, a_1, a_2, a_3, \ldots) : a_i \in \mathbb{R}, a_{i+2} = a_{i+1} + a_i, i \geq 0\}.$$

容易验证

(a) 若数列 $(a_0, a_1, a_2, \ldots)$ 和数列 $(b_0, b_1, b_2, \ldots)$ 都满足斐波那契递归公式，则它们的和 $(a_0 + b_0, a_1 + b_1, a_2 + b_2, \ldots)$ 满足斐波那契递归公式。

(b) 若 $(a_0, a_1, a_2, \ldots)$ 满足斐波那契递归公式，则对任意的实数 λ，数列 $(\lambda a_0, \lambda a_1, \lambda a_2, \ldots)$ 也满足斐波那契递归公式。

(c) 集合 $\mathscr{F}$ 在这两种运算下构成一个实线性空间。

(d) $\dim \mathscr{F} = 2$.

证明. 粗略的解释是，若知道了 $\mathscr{F}$ 中某数列的前两项，由递归公式就知道了数列的所有项，所以在确定 $\mathscr{F}$ 中数列时，我们有两个自由度的选择空间。

精确的解释是，数列

$$\boldsymbol{e}_0 = (1, 0, 1, 1, 2, 3, \ldots)$$

和经典的斐波那契数列

$$\boldsymbol{e}_1 = (0, 1, 1, 2, 3, 5, \ldots)$$

是 $\mathscr{F}$ 的一组基。它们是初始两项分别等于 $(1, 0)$ 和 $(0, 1)$ 且满足斐波那契递归公式的数列。对任意的 $(a_0, a_1, a_2, \ldots) \in \mathscr{F}$, 以下等式成立

$$(a_0, a_1, a_2, \ldots) = a_0(1, 0, 1, 1, 2, 3, \ldots) + a_1(0, 1, 1, 2, 3, 5, \ldots) = a_0\boldsymbol{e}_0 + a_1\boldsymbol{e}_1.$$

□

(e) $T : \mathscr{F} \to \mathscr{F}; \quad (a_0, a_1, a_2, a_3, \ldots) \mapsto (a_1, a_2, a_3, a_4, \ldots)$ 是 $\mathscr{F}$ 上的线性变换。

证明. 若 $(a_0, a_1, a_2, a_3, \ldots) \in \mathscr{F}$, 则 $(a_1, a_2, a_3, a_4, \ldots) \in \mathscr{F}$，故 T 是良定义的。再根据定义验证 T 是线性变换。 □

例 8. 以下例子都不是线性变换

(1) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto 2x + 1$, 原因：$f(0) \neq 0$. 一般的，形如 $\boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ 的变换称为 *仿射变换 (affine transformation)*.

(2) $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$, 原因：$f(x + y) = (x + y)^2 = x^2 + 2xy + y^2 \neq x^2 + y^2 = f(x) + f(y)$, $f(cx) = c^2x^2 \neq cx^2 = cf(x)$.

(3) $x^3, e^x, \sin(x)$ 都不是 $\mathbb{R}$ 上的线性变换。

定义 2. 若线性变换 $T : V \to W$ 还是双射（既单又满），则称 T 是可逆线性变换。

命题 3. *(1)* 可逆线性变换的逆变换仍是线性变换。

(2) 线性变换的复合仍是线性变换。

(3) 若 $T: V \to W$ 是线性变换，V_1, W_1 分别是 V, W 的子空间且 $T(V_1) \subseteq W_1$，则 $T_1: V_1 \to W_1$, $\boldsymbol{\alpha} \mapsto T(\boldsymbol{\alpha})$ 是线性变换。

证明. (1) 设 $T: V \to W$ 是可逆线性变换，T 的逆变换是

$$T^{-1}: W \to V, \quad \boldsymbol{\gamma} \mapsto T^{-1}(\boldsymbol{\gamma}),$$

其中 $T^{-1}(\boldsymbol{\gamma})$ 是 V 中唯一满足 $T(T^{-1}(\boldsymbol{\gamma})) = \boldsymbol{\gamma}$ 的向量。设 $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in W$, 则由 T 的线性知

$$T(T^{-1}(\boldsymbol{\gamma}_1) + T^{-1}(\boldsymbol{\gamma}_2)) = T(T^{-1}(\boldsymbol{\gamma}_1)) + T(T^{-1}(\boldsymbol{\gamma}_2)) = \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2,$$

故

$$T^{-1}(\boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2) = T^{-1}(\boldsymbol{\gamma}_1) + T^{-1}(\boldsymbol{\gamma}_2).$$

类似可证明 $T^{-1}(c\boldsymbol{\gamma}) = cT^{-1}(\boldsymbol{\gamma})$.

(2) 若 $S: U \to V$, $T: V \to W$ 都是线性变换，T 与 S 的复合 $T \circ S$ 定义如下

$$(T \circ S)(\boldsymbol{\alpha}) = T(S(\boldsymbol{\alpha})), \quad \boldsymbol{\alpha} \in U,$$

故对任意的 $\boldsymbol{\alpha}, \boldsymbol{\beta} \in U$,

$$(T \circ S)(\boldsymbol{\alpha} + \boldsymbol{\beta}) = T(S(\boldsymbol{\alpha} + \boldsymbol{\beta})) = T(S(\boldsymbol{\alpha}) + S(\boldsymbol{\beta})) = T(S(\boldsymbol{\alpha})) + T(S(\boldsymbol{\beta})) = (T \circ S)(\boldsymbol{\alpha}) + (T \circ S)(\boldsymbol{\beta}).$$

类似可以证明 $(T \circ S)(c\boldsymbol{\alpha}) = c(T \circ S)(\boldsymbol{\alpha})$, 故 $T \circ S$ 是线性变换。

(3) 显然。 □

定理 4. 设 $T: V \to W$ 是线性变换，令

$$ker(T) = \{\boldsymbol{\alpha} \in V \mid T(\boldsymbol{\alpha}) = \mathbf{0}\},$$

$$im(T) = \{T(\boldsymbol{\alpha}) \mid \boldsymbol{\alpha} \in V\},$$

则

(1) $ker(T)$ 是 V 的线性子空间，称为 T 的核。

(2) $im(T)$ 是 W 的线性子空间，称为 T 的像。

(3) T 是单射当且仅当 $ker(T) = 0$.

(4) T 是满射当且仅当 $im(T) = W$.

(5) 若 $\dim V < \infty$, 则 $\dim V = \dim ker(T) + \dim im(T)$.

证明. 作业题做过。 □

例 9. (1) 矩阵 A 诱导的线性变换 $\mathscr{A}$ 的核是 $N(A)$, 像是 $C(A)$.

(2) 例 2 中线性函数 f 的核是 $\mathbb{R}^3$ 中的平面 $x+2y+3z=0$，像是 $\mathbb{R}$.

(3) 例 5 中 ev_a 的核是满足 $f(a)=0$ 的函数构成的子空间，像是 $\mathbb{R}$.

(4) 例 6 中 T 的核是

$$\{X = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{bmatrix} \mid \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3 \in N(A)\},$$

像是

$$\{Y = \begin{bmatrix} \boldsymbol{y}_1 & \boldsymbol{y}_2 & \boldsymbol{y}_3 \end{bmatrix} \mid \boldsymbol{y}_1, \boldsymbol{y}_2, \boldsymbol{y}_3 \in C(A)\}.$$

例 10. 例 3 中求导算子的核是常值函数构成的子空间 $\mathbb{R}$，像是 $\mathcal{C}(\mathbb{R})$. 例 4 中积分算子的核是零，像是满足 $F(0)=0$ 的连续可微函数构成的子空间，记为 Z. 事实上，微积分基本定理告诉我们，$\frac{d}{dx}\int_0^x f(t)dt = f(x)$，换言之

$$D \circ \int = \mathbf{1}_{\mathcal{C}(\mathbb{R})},$$

从这个等式可以推出 D 是满射，$\int$ 是单射。实轴上导数处处为 0 的函数只有常数函数，所以

$$\ker(D) = \{c : c \in \mathbb{R}\}.$$

又因为 $\int_0^x f(t)dt$ 在 0 处的取值是 0，所以

$$\mathrm{im} \int \subseteq \{F(x) \in \mathcal{C}^1(\mathbb{R}) \mid F(0) = 0\}.$$

反之，若连续可微函数 $F(x)$ 满足 $F(0)=0$, 则 $F(x) = \int_0^x F'(t)dt \in \mathrm{im}\int$. 所以

$$\mathrm{im} \int = \{F(x) \in \mathcal{C}^1(\mathbb{R}) \mid F(0) = 0\} = Z.$$

$\mathcal{C}^1(\mathbb{R}) = \mathbb{R} \oplus Z$ 是常值函数 $\mathbb{R}$ 和子空间 Z 的直和，即任意函数 $F(x)$ 可以唯一的写成如下形式

$$\underbrace{F(x)}_{\in \mathcal{C}^1(\mathbb{R})} = \underbrace{F(0)}_{\in \mathbb{R}} + \underbrace{(F(x) - F(0))}_{\in Z}.$$

容易验证求导算子 D 在 Z 上的限制

$$D|_Z : Z \to \mathcal{C}(\mathbb{R})$$

和

$$\int : \mathcal{C}(\mathbb{R}) \to Z$$

互为逆线性变换，而原来的 $D : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}(\mathbb{R})$ 和 $\int : \mathcal{C}(\mathbb{R}) \to \mathcal{C}^1(\mathbb{R})$ 的复合是

$$(\int \circ D)(F(x)) = \int_0^x F'(t)dt = F(x) - F(0),$$

可以说 $\int \circ D$ 是到 Z 的投影，而另一个方向的复合是 $D \circ \int = \mathbf{1}_{\mathcal{C}(\mathbb{R})}$. 虽然这里涉及到的线性空间是无限维的，但 $\int$ "像"是 D 的广义逆。

0.1 线性变换的矩阵

从 $\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的线性变换都可以用矩阵表示，那么一般的线性变换是否可以用矩阵表示呢？可以！先在出发域和到达域中选基。

定义 3. 设 V, W 是有限维线性空间，$T : V \to W$ 是线性变换。设 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 是 V 的一组基，$\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$ 是 W 的一组基，$T(\boldsymbol{\alpha}_j)$ 在基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$ 下的坐标是 $(a_{1j}, \ldots, a_{mj})$, 即

$$T(\boldsymbol{\alpha}_j) = a_{1j}\boldsymbol{\beta}_1 + \ldots + a_{mj}\boldsymbol{\beta}_m = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix},$$

则称 $A = (a_{ij})_{1\le i\le m, 1\le j\le n}$ 是线性变换 T 在 V 的基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 和 W 的基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$ 下的矩阵，我们用如下记号表示

$$T(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n) = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)A = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m)\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix}.$$

命题 5. 若 $\boldsymbol{v}$ 在基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_n$ 下的坐标是 $\boldsymbol{x} = (x_1, \ldots, x_n)$, 即 $\boldsymbol{v} = x_1\boldsymbol{\alpha}_1 + \ldots + x_n\boldsymbol{\alpha}_n$, 则 $T(\boldsymbol{v})$ 在基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m$ 下的坐标是 $A\boldsymbol{x}$, 即若 $A\boldsymbol{x} = \boldsymbol{y} = (y_1, \ldots, y_m)$, 则 $T(\boldsymbol{x}) = y_1\boldsymbol{\beta}_1 + \ldots + y_m\boldsymbol{\beta}_m$.

证明. 由 T 的线性知

$$\begin{aligned} T(\boldsymbol{v}) &= T(x_1\boldsymbol{\alpha}_1 + \ldots + x_n\boldsymbol{\alpha}_n) \\ &= x_1T(\boldsymbol{\alpha}_1) + \ldots + x_nT(\boldsymbol{\alpha}_n) \\ &= x_1(a_{11}\boldsymbol{\beta}_1 + \ldots + a_{m1}\boldsymbol{\beta}_m) + \ldots + x_n(a_{1n}\boldsymbol{\beta}_1 + \ldots + a_{mn}\boldsymbol{\beta}_m) \\ &= (\underbrace{a_{11}x_1 + \ldots + a_{1n}x_n}_{y_1})\boldsymbol{\beta}_1 + \ldots + (\underbrace{a_{m1}x_1 + \ldots + a_{mn}x_n}_{y_m})\boldsymbol{\beta}_m. \end{aligned}$$

□

命题 6. 设 $\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n'$ 是 V 的另一组基，且

$$(\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n') = (\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)P,$$ [1]

$\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m'$ 是 W 的另一组基，且

$$(\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m') = (\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m)Q,$$ [2]

则 T 在基 $\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n'$ 和基 $\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m'$ 下的矩阵是 $Q^{-1}AP$.

证明. 根据定义验证即可

$$\begin{aligned}
T(\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n') &= T((\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)P) &&= (T(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n))P \\
&= ((\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m)A)P &&= (\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m)(AP) \\
&= ((\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m')Q^{-1})AP &&= (\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m')(Q^{-1}AP).
\end{aligned}$$

第二个等号用到了 T 是线性变换的性质。 □

定义 4. 设 $A,B\in M_{m\times n}(\mathbb{R})$. 若存在可逆阵 P,Q 使得 $B=PAQ$，则称 A 与 B 相抵。

命题 7. *(1)* 矩阵的相抵是等价关系。

(2) A 与 B 相抵的充要条件是 $r(A)=r(B)$. [3]

证明. (1) 容易。(2) 回顾：存在可逆阵 P 使得 $PA=R=\mathrm{rref}(A)$. 用 R 的主元列的线性组合可以把 R 的自由列都消为 0，然后把 R 的主元列依次换到最左边，得到形如 $\begin{bmatrix} I_r & 0 \\ 0 & 0\end{bmatrix}$ 的矩阵，其中 $r=r(A)$. 对矩阵做初等列变换等价于在矩阵的右边乘初等方阵，故存在可逆阵 Q 使得

$$PAQ = RQ = \begin{bmatrix} I_r & 0 \\ 0 & 0\end{bmatrix}.$$

等式右边的矩阵称为 A 的相抵标准型。因为与可逆阵相乘不改变矩阵的秩，故若 A,B 相抵，则 $r(A)=r(B)$. 反之，若 $r(A)=r(B)=r$, 则 A,B 都与 $\begin{bmatrix} I_r & 0 \\ 0 & 0\end{bmatrix}$ 相抵，再由相抵的对称性和传递性知 A,B 相抵。 □

命题 8. *(1)* 线性变换在不同基下的矩阵是相抵的。

[1] P 是 n 阶可逆阵，称为从 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 到 $\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n'$ 的过渡矩阵。

[2] Q 是 m 阶可逆阵，是从 $\boldsymbol{\beta}_1,\dots,\boldsymbol{\beta}_m$ 到 $\boldsymbol{\beta}_1',\dots,\boldsymbol{\beta}_m'$ 的过渡矩阵。

[3] 这里默认 A,B 大小一样。

(2) 若 A 是线性变换 T 的矩阵，则任一与 A 相抵的矩阵也是 T 的矩阵。[4]

(3) 设 $T: V \to W$ 是线性变换，V, W 都是有限维线性空间，则存在 V 的基和 W 的基使得 T 在这两组基下的矩阵是 $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$.

证明. (1) 已证。

(2) 把命题 6 的证明反过去。

(3) 用 T 的任一矩阵的相抵标准形。或者，任取由 $\ker(T)$ 的基 $\boldsymbol{\alpha}_{r+1}, \dots, \boldsymbol{\alpha}_n$ 扩张得来的 V 的一组基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \dots, \boldsymbol{\alpha}_n$. 在作业中已经证明了 $\boldsymbol{\beta}_1 = T(\boldsymbol{\alpha}_1), \dots, \boldsymbol{\beta}_r = T(\boldsymbol{\alpha}_r)$ 是 $\mathrm{im}(T)$ 的一组基。我们把它扩充为 W 的一组基 $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_r, \boldsymbol{\beta}_{r+1}, \dots, \boldsymbol{\beta}_m$, 则

$$\begin{aligned} T(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_r, \boldsymbol{\alpha}_{r+1}, \dots, \boldsymbol{\alpha}_n) &= (T(\boldsymbol{\alpha}_1), \dots, T(\boldsymbol{\alpha}_r), \mathbf{0}, \dots, \mathbf{0}) \\ &= (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_r, \mathbf{0}, \dots, \mathbf{0}) \\ &= (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_r, \boldsymbol{\beta}_{r+1}, \dots, \boldsymbol{\beta}_m) \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

□

例 11. 设 $A \in M_{m\times n}(\mathbb{R})$, $\mathscr{A}: \mathbb{R}^n \to \mathbb{R}^m$ 是 A 诱导的线性变换，则

(1) $\mathscr{A}$ 在 $\mathbb{R}^n$ 的自然基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ 和 $\mathbb{R}^m$ 的自然基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ 下的矩阵是 A.

(2) 若 $A = U\Sigma V^T$ 是 A 的奇异值分解，$U = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_m \end{bmatrix}$, $V = \begin{bmatrix} \boldsymbol{v}_1 & \dots & \boldsymbol{v}_n \end{bmatrix}$, 则 $\mathscr{A}$ 在 A 的右奇异向量 $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ 和 A 的左奇异向量 $\boldsymbol{u}_1, \dots, \boldsymbol{u}_m$ 下的矩阵是 Σ.

(3) 若 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 是 $\mathbb{R}^n$ 的一组基，$\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m$ 是 $\mathbb{R}^m$ 的一组基，且令

$$P = \begin{bmatrix} \boldsymbol{\alpha}_1 & \dots & \boldsymbol{\alpha}_n \end{bmatrix}, \qquad Q = \begin{bmatrix} \boldsymbol{\beta}_1 & \dots & \boldsymbol{\beta}_m \end{bmatrix},$$

则 $\mathscr{A}$ 在基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 和基 $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m$ 下的矩阵是 $Q^{-1}AP$.

证明. (1) $\mathscr{A}(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) = AI_n = A = I_m A = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)A$.

(2) $\mathscr{A}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) = AV = U\Sigma = (\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)\Sigma$.

(3) $\mathscr{A}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n) = AP = Q(Q^{-1}AP) = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m)(Q^{-1}AP)$. 这和命题 6 中的公式一致，因为从 $\mathbb{R}^n$ 的基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ 到基 $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_n$ 的过渡矩阵是 P, 从 $\mathbb{R}^m$ 的基 $\boldsymbol{e}_1, \dots, \boldsymbol{e}_m$ 到基 $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m$ 的过渡矩阵是 Q, $\mathscr{A}$ 在两组自然基下的矩阵是 A. □

[4] 1,2 合在一起说明，线性变换 T 的矩阵恰是一个相抵等价类中的矩阵。

定义 5. 设 $T:V\to V$ 是有限维线性空间 V 上的线性变换，$\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是 V 的一组基。我们在出发域和到达域中都使用这组基

$$T(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)A,$$

则 T 在这两组相同的基下的矩阵 A 称为 T 在基 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵。

命题 9. 若 $\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n'$ 是 V 的另一组基，且 $(\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n')=(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n)P$, 则 T 在基 $\boldsymbol{\alpha}_1',\dots,\boldsymbol{\alpha}_n'$ 下的矩阵是 $P^{-1}AP$, 故 T 在不同基下的矩阵是相似的，而且任一与 A 相似的矩阵 B 均是 T 在某组基下的矩阵。

证明. 在命题 6 中令 $Q=P$ 即可。 □

例 12. 若 $A\in M_n(\mathbb{R})$, 则 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^n$ 在 $\mathbb{R}^n$ 的自然基 $\boldsymbol{e}_1,\dots,\boldsymbol{e}_n$ 下的矩阵是 A. 若 A 可对角化且 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 是 A 的 n 个线性无关的特征向量，特征值分别是 $\lambda_1,\dots,\lambda_n$，则 A 在 $\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_n$ 下的矩阵是 $\Lambda=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$.

例 13. 设 $A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$, $T=T_A:M_2(\mathbb{R})\to M_2(\mathbb{R})$, $X\mapsto AX$. 在 $M_2(\mathbb{R})$ 中取基

$$E_{11}=\begin{bmatrix}1&0\\0&0\end{bmatrix},E_{12}=\begin{bmatrix}0&1\\0&0\end{bmatrix},E_{21}=\begin{bmatrix}0&0\\1&0\end{bmatrix},E_{22}=\begin{bmatrix}0&0\\0&1\end{bmatrix}.$$

对任意的 $X=\begin{bmatrix}a&b\\c&d\end{bmatrix}$, 计算得 $AX=\begin{bmatrix}1&1\\0&1\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}a+c&b+d\\c&d\end{bmatrix}$. X 在 $E_{11},E_{12},E_{21},E_{22}$ 下的坐标是 (a,b,c,d), AX 在 $E_{11},E_{12},E_{21},E_{22}$ 下的坐标是 $(a+c,b+d,c,d)$. 因为

$$\begin{bmatrix}a+c\\b+d\\c\\d\end{bmatrix}=\underbrace{\begin{bmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{bmatrix}}_{M}\begin{bmatrix}a\\b\\c\\d\end{bmatrix},$$

故 T 在 $E_{11},E_{12},E_{21},E_{22}$ 下的矩阵是 M. 或者，从上面的公式可以看出 $T(E_{11})=E_{11}$, $T(E_{12})=E_{12}$, $T(E_{21})=E_{11}+E_{21}$, $T(E_{22})=E_{12}+E_{22}$, 故

$$T(E_{11},E_{12},E_{21},E_{22})=(E_{11},E_{12},E_{11}+E_{21},E_{12}+E_{22})=(E_{11},E_{12},E_{21},E_{22})\begin{bmatrix}1&0&1&0\\0&1&0&1\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

例 14. 设 $V=\mathbb{R}[x]_{\deg\le 2}$ 是次数小于等于 2 的一元实系数多项式构成的线性空间, $D:V\to V$, $p(x)\mapsto p'(x)$ 是 V 上的求导变换。因为 $D(1)=0$, $D(x)=1$, $D(x^2)=2x$, 即

$$D(1,x,x^2)=(0,1,2x)=(1,x,x^2)\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix},$$

所以 D 在基 $1,x,x^2$ 下的矩阵是 $\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}$. 若 $p(x)=a+bx+cx^2$, 则 $p(x)$ 在 $1,x,x^2$ 下的坐标是 (a,b,c), 又 $p'(x)=b+2cx$, 故 $p'(x)$ 在 $1,x,x^2$ 下的坐标是 $(b,2c,0)$. 如命题 5 所料, $p(x)$ 和 $D(p(x))=p'(x)$ 在基 $1,x,x^2$ 下的坐标以及 D 在基 $1,x,x^2$ 下的矩阵满足以下等式

$$\begin{bmatrix}b\\2c\\0\end{bmatrix}=\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix}.$$

若取 $1,x,\frac{x^2}{2}$ 为 V 的基, 则

$$D(1,x,\frac{x^2}{2})=(0,1,x)=(1,x,\frac{x^2}{2})\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix},$$

因此 $\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}$ 和 $\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}$ 是 D 在不同基下的矩阵, 故它们相似。具体的, 我们有

$$\begin{bmatrix}0&1&0\\0&0&1\\0&0&0\end{bmatrix}=\underbrace{\begin{bmatrix}1&&\\&1&\\&&2\end{bmatrix}}_{P^{-1}}\begin{bmatrix}0&1&0\\0&0&2\\0&0&0\end{bmatrix}\underbrace{\begin{bmatrix}1&&\\&1&\\&&\frac{1}{2}\end{bmatrix}}_{P},$$

这里的 P 是从基 $1,x,x^2$ 到基 $1,x,\frac{x^2}{2}$ 的过渡矩阵。

0.2 线性变换的复合与矩阵乘法

若 $A\in M_{m\times n}(\mathbb{R})$, $B\in M_{n\times k}(\mathbb{R})$ 诱导的线性变换分别是 $\mathscr{A}:\mathbb{R}^n\to\mathbb{R}^m$, $\mathscr{B}:\mathbb{R}^k\to\mathbb{R}^n$, 则 $(\mathscr{A}\circ\mathscr{B})(\boldsymbol{x})=\mathscr{A}(\mathscr{B}(\boldsymbol{x}))=\mathscr{A}(B\boldsymbol{x})=A(B\boldsymbol{x})=(AB)\boldsymbol{x}$, 即 $\mathscr{A}$ 与 $\mathscr{B}$ 的复合 $\mathscr{A}\circ\mathscr{B}$ 是 AB 诱导的从 $\mathbb{R}^k$ 到 $\mathbb{R}^m$ 的线性变换。

命题 10. 设 U, V, W 是有限维线性空间, $T: V \to W$, $S: U \to V$ 是线性变换, $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ 是 U 的一组基, $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ 是 V 的一组基, $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_m$ 是 W 的一组基。若 S 在 U 的基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ 和 V 的基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ 下的矩阵是 B, T 在 V 的基 $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n$ 和 W 的基 $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_m$ 下的矩阵是 A, 则 $T \circ S$ 在 U 的基 $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k$ 和 W 的基 $\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_m$ 下的矩阵是 AB.

证明.

$$
\begin{aligned}
(T \circ S)(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k) &= T(S(\boldsymbol{\alpha}_1), \ldots, S(\boldsymbol{\alpha}_k)) \\
&= T((\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n)B) \\
&= (T(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_n))B \\
&= ((\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_m)A)B \\
&= (\boldsymbol{\gamma}_1, \ldots, \boldsymbol{\gamma}_m)(AB).
\end{aligned}
$$

□

0.3 线性变换的特征值与特征向量

定义 6. 设 V 是有限维线性空间, $T: V \to V$ 是 V 上的线性变换, $\lambda \in \mathbb{R}$. 如果存在 $\mathbf{0} \neq \boldsymbol{\alpha} \in V$ 使得

$$
T(\boldsymbol{\alpha}) = \lambda \boldsymbol{\alpha}, \tag{1}
$$

则称 λ 是线性变换 T 的特征值, 满足 *(1)* 的非零向量 $\boldsymbol{\alpha}$ 称为 T 的属于 λ 的特征向量。

若 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ 是 V 的基且 T 在这组基下的矩阵是 A, 则 $T(\boldsymbol{\alpha}) = \lambda\boldsymbol{\alpha}$ 等价于 $A\boldsymbol{x} = \lambda\boldsymbol{x}$, 其中 $\boldsymbol{x}$ 是 $\boldsymbol{\alpha}$ 在 $\boldsymbol{\epsilon}_1, \ldots, \boldsymbol{\epsilon}_n$ 下的坐标。

定义 7. T 在不同基下的矩阵相似, 相似的矩阵有相同的特征多项式, 故我们定义 T 的特征多项式 $\det(\lambda\mathbf{1} - T)$ 是 T 的矩阵的特征多项式。

例 15. 回顾满足斐波那契递归公式的数列构成的线性空间

$$
\mathscr{F} = \{(a_0, a_1, a_2, a_3, \ldots) : a_i \in \mathbb{R}, a_{i+2} = a_{i+1} + a_i, i \geq 0\}
$$

和其上的线性变换

$$
T : \mathscr{F} \to \mathscr{F}; \quad (a_0, a_1, a_2, \ldots) \mapsto (a_1, a_2, a_3, \ldots).
$$

如下数列是 $\mathscr{F}$ 的一组基

$$
\boldsymbol{e}_0 = (1, 0, 1, 1, 2, 3, \ldots),
$$

$$\boldsymbol{e}_1 = (0, 1, 1, 2, 3, 5, \ldots).$$

计算发现

$$T(\boldsymbol{e}_0) = \boldsymbol{e}_1, \quad T(\boldsymbol{e}_1) = \boldsymbol{e}_0 + \boldsymbol{e}_1,$$

故 T 在基 $\boldsymbol{e}_0, \boldsymbol{e}_1$ 下的矩阵是 $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$

接下来，我们在 $\mathscr{F}$ 中寻一组“更好”的基。根据定义，$(a_0, a_1, a_2, \ldots)(\neq 0)$ 是 T 的特征向量当且仅当存在 $\lambda \in \mathbb{R}$ 使得

$$T(a_0, a_1, a_2, \ldots) \quad = \quad \lambda(a_0, a_1, a_2, \ldots). \tag{2}$$

因为

$$\begin{aligned} T(a_0, a_1, a_2, \ldots) &= (a_1, a_2, a_3, \ldots), \\ \lambda(a_0, a_1, a_2, \ldots) &= (\lambda a_0, \lambda a_1, \lambda a_2, \ldots), \end{aligned}$$

所以 (2) 式成立当且仅当 $a_{i+1} = \lambda a_i$ $(i \geq 0)$, 也就是说

$$(a_0, a_1, a_2, \ldots) = (a_0, a_0\lambda, a_0\lambda^2, \ldots)$$

是公比为 λ 的等比数列。因为 $(a_0, a_0\lambda, a_0\lambda^2, \ldots)$ 还需满足 $a_{i+2} = a_{i+1} + a_i$, 故 λ 必须满足 $\lambda^2 = \lambda + 1$, 也就是

$$\lambda^2 - \lambda - 1 = 0.$$

这个 2 次方程有两个互异实根 $\lambda_1 = \frac{1+\sqrt{5}}{2}$, $\lambda_2 = \frac{1-\sqrt{5}}{2}$. 于是

$$\begin{aligned} \boldsymbol{\epsilon}_1 &= (1, \lambda_1, \lambda_1^2, \ldots) \\ \boldsymbol{\epsilon}_2 &= (1, \lambda_2, \lambda_2^2, \ldots) \end{aligned}$$

是 T 的分别属于特征值 λ_1, λ_2 的特征向量,T 在特征向量 $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2$ 构成的基下的矩阵是 $\mathrm{diag}(\lambda_1, \lambda_2)$.

容易看出

$$(\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2) = (\boldsymbol{e}_0, \boldsymbol{e}_1) \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix},$$

即从基 $\boldsymbol{e}_0, \boldsymbol{e}_1$ 到基 $\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2$ 的过渡矩阵是 $P = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix}$. 反之,

$$(\boldsymbol{e}_0, \boldsymbol{e}_1) = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)P^{-1} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2)\frac{1}{\lambda_2 - \lambda_1} \begin{bmatrix} \lambda_2 & -1 \\ -\lambda_1 & 1 \end{bmatrix},$$

特别的

$$\boldsymbol{e}_1 = \frac{1}{\lambda_1 - \lambda_2}\boldsymbol{\epsilon}_1 - \frac{1}{\lambda_1 - \lambda_2}\boldsymbol{\epsilon}_2 = \frac{(1, \lambda_1, \lambda_1^2, \ldots) - (1, \lambda_2, \lambda_2^2, \ldots)}{\lambda_1 - \lambda_2},$$

于是我们又一次得到经典的斐波那契数列 $\boldsymbol{e}_1$ 的通项公式 $\frac{\lambda_1^k - \lambda_2^k}{\lambda_1 - \lambda_2}$.

0.4 换基

线性变换 $T : V \to W$ 在不同基下的矩阵相抵。

线性变换 $T : V \to V$ 在不同基下的矩阵相似。

二次型 $q : V \to \mathbb{R}$ 在不同基下的矩阵相合。

Gram-Schmidt 正交化把子空间的基变成正交基。