第二章　矩阵与线性变换

引　言　矩阵是什么？

在线性代数中, 矩阵是表示线性方程组的一种简便形式. 对于二元或三元线性方程组 $Ax = b$来说, 其解 (集) 有非常直观的几何意义, 即若干直线或平面的交集, 多元的情形也有类似的意义. 那么, 矩阵是否也有类似的解释呢？从函数或映射的角度来理解, 线性方程组 $Ax = b$ 的解 (集) 可以看成是在 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的映射 $\sigma : x \mapsto Ax$ 下向量 $b \in \mathbb{F}^m$ 的原像. 特别地, 齐次线性方程组 $Ax = 0$ 的解 (集) 恰好是映射 $\sigma : x \mapsto Ax$ 的“零点”. 注意映射 $\sigma : x \mapsto Ax$ 满足下述两条简单性质：

(T1) (**可加性**) $\sigma(x + y) = \sigma(x) + \sigma(y)$;
(T2) (**齐次性**) $\sigma(ax) = a\sigma(x)$, $a \in \mathbb{F}$.

我们将在本章中证明, 如果映射 $\tau : \mathbb{F}^n \to \mathbb{F}^m$ 满足上述两条性质, 则必然存在矩阵 $A_{m\times n}$ 使得 $\sigma(x) = Ax, \forall x \in \mathbb{F}^n$. 于是, 求向量 $b \in \mathbb{F}^m$ 在 σ 下的原像等价于解线性方程组 $Ax = b$. 因此, 矩阵 A 和满足条件 (T1) 与 (T2) 的映射 (此即本章的核心: 线性变换或线性映射) 具有非常深刻的天然联系, 本章的目的即是研究二者之间的联系.

第一节　子空间: 直和与空间分解

复数域 $\mathbb{C}$ 是实数域 $\mathbb{R}$ 上的 2 维线性空间, 实数域作为复数域的子集自身是实数域上的 1 维线性空间, 这时我们称 $\mathbb{R}$ 是 $\mathbb{C}$ 的“子空间”. 一般地, 有下述定义.

定义 2.1.1　(子空间) 设 V 是一个线性空间, U 是 V 的一个非空子集. 如果 U 本身关于 V 的向量加法与数乘作成一个线性空间, 则称 U 是 V 的一个线性子空间或子空间.

任何非零线性空间都至少有两个子空间, 即零子空间 $\{0\}$ 与它自身, 称为平凡子空间. 其余的子空间称为真子空间. 零子空间常简记为 0 (请注意我们用同一个符号 0 表示很多不同的概念).

注意子空间的定义强调“关于 V 的向量加法与数乘”, 比如 $\mathbb{R}$ 是 $\mathbb{C}$ 的子集且本身是线性空间, 因此 $\mathbb{R}$ 是 $\mathbb{C}$ 的实线性子空间, 但 $\mathbb{R}$ 却不是复线性空间 $\mathbb{C}$ 的子空间, 因为 $\mathbb{R}$ 关于 $\mathbb{C}$ 的数乘并不封闭. 再比如, 在第一章, 例 1.4.1 中, $V = \{$全体正实数$\}$ 是 $\mathbb{R}$ 的子集, V 本身也是实线性空间, 但它并不是 $\mathbb{R}$ 的子空间, 因为 V 的两个运算均与 $\mathbb{R}$ 中的不同.

例 2.1.1　$\mathbb{F}[x]_n$ 是多项式空间 $\mathbb{F}[x]$ 的一个 n 维子空间. 若把实系数多项式也看成闭区间 $[a, b]$ 上的函数, 则 $\mathbb{R}[x]$ 与 $\mathbb{R}[x]_n$ 都是 $[a, b]$ 上的连续函数空间 $C[a, b]$ 的子空间, 其中, $C[a, b]$ 的加法和数乘都是普通的.

一个有限维线性空间 V 的子空间 U 仍是有限维的, 且 U 的维数不超过 V 的维数; 且若二者维数相等, 则它们本身也是相等的: 这是因为, 由第一章, 推论 1.4.1, n 维线性空间中任意 n 个线性无关的向量均构成 V 的一组基.

下述简单易行的子空间判别准则 (证明见习题 1) 告诉我们, 一个非空子集是子空间当且仅当它关于加法与数乘均封闭:

定理 2.1.1 (子空间判别法) 设 U 是线性空间 V 的一个非空子集. 则 U 是子空间 $\Longleftrightarrow$ 对任意 $\lambda \in \mathbb{F}$, $\alpha, \beta \in U$, 有 $\alpha + \beta \in U$ 与 $\lambda\alpha \in U$.

注意子空间必然包含 0 向量 (在平面或空间解析几何中, 就是子空间要过原点), 因此初学者不妨将此条也加入到上述准则中去.

可以直接验证子空间的下述性质 (见习题 2):

命题 2.1.1 (1) 传递性: 即若 U 是 V 的子空间, W 是 U 的子空间, 则 W 也是 V 的子空间;

(2) 任意多个 (可以无限) 子空间的交集仍是子空间, 称为这些子空间的交, 且是含于这些子空间的最大子空间; 特别, 两个子空间 U 与 W 的交 $U \cap W$ 仍是子空间.

一般来说, 两个子空间的并集 $U \cup W$ 不再是子空间 (为什么?). 但 V 是包含 U 与 W 的一个子空间, 因此要问包含 U 与 W 的最小的子空间是否存在? 显然, 这样的子空间必须包含所有形如 $\alpha + \beta$, $\alpha \in U$, $\beta \in W$ 的向量. 将全体这样的向量构成的子集合记为 $U + W$, 由 定理2.1.1 可知 $U + W$ 是子空间, 且是包含 U 与 W 的最小的子空间, 称为 U 与 W 的**和**.

子空间的和的概念可以推广到任意有限多个子空间的情形. 设 $U_1, U_2, \cdots, U_s$ 是线性空间 V 的子空间, 则集合

$$\{\alpha_1 + \alpha_2 + \cdots + \alpha_s \,|\, \alpha_j \in U_j,\ 1 \le j \le s\}$$

也是子空间 (为什么?), 称为 $U_1, U_2, \cdots,\ U_s$ 的和, 记为 $U_1 + U_2 + \cdots + U_s$ 或 $\sum\limits_{j=1}^{s} U_j$.

设 V 是线性空间, $S \subset V$. 称 V 的包含 S 的最小子空间为由 S 生成 (或张成) 的子空间, 记为 $\mathrm{Span}\, S$, S 称为 $\mathrm{Span}\, S$ 的生成元集. 显然, 当 $S = \emptyset$ 或 $S = \{0\}$ 时, $\mathrm{Span}\, S = 0$ 是零维子空间; 若 $S = \{\alpha\}$ 是一元集且 $\alpha \ne 0$, 则 $\mathrm{Span}\{\alpha\} = \{k\alpha \,|\, k \in \mathbb{F}\}$ 是一维子空间, 记为 $\mathbb{F}\alpha$; 若 $S = \{\alpha_1, \alpha_2, \cdots, \alpha_s\}$ 是有限集, 则记 $\mathrm{Span}\, S = \mathrm{Span}\{\alpha_1, \alpha_2, \cdots, \alpha_s\}$, 此时有

$$\mathrm{Span}\{\alpha_1, \alpha_2, \cdots, \alpha_s\} = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \,|\, k_j \in \mathbb{F}, 1 \le j \le s\}.$$

一般地, 直接验证可知,

$$\mathrm{Span}\, S = \{k_1\alpha_1 + k_2\alpha_2 + \cdots + k_s\alpha_s \,|\, \alpha_j \in S, k_j \in \mathbb{F}, s \ge 0\},$$

即 $\mathrm{Span}\, S$ 由 S 中元素的所有可能的 (有限) 线性组合构成. 特别地, 两个子空间 U 与 W 的和 $U + W = \mathrm{Span}\, U \cup W$.

例 2.1.2 设 $V = M_n(\mathbb{R})$ 是 n 阶实矩阵构成的线性空间. 令 $D = \{$全体对角矩阵$\}$, $U = \{$全体上三角矩阵$\}$, $W = \{$全体下三角矩阵$\}$. 则 D, U 与 W 均是 V 的子空间, 且 $U + W = V$; $U \cap W = D$; 但 $U \cup W$ 不是子空间 (加法不封闭). 其中 D 的维数等于 n; U 与 W 的维数均等于 $n(n+1)/2$; $U + W = V$ 的维数等于 n^2. 即

$$(\dim U + \dim W) - \dim (U + W) = \dim (U \cap W). \tag{2.1.1}$$

(请读者比较该公式与普通集合的计数公式: $|A \cup B| = |A| + |B| - |A \cap B|$, 其中 $|A|$ 表示 A 的元素个数.)

例 2.1.3 设 $V = \mathbb{R}[x]_n$, $U = \{f(x) \in V \mid f(1) = 0\}$, $W = \{f(x) \in V \mid f(2) = 0\}$. 则 $U + W = V$, $U \cap W = \{f(x) \in V \mid f(1) = 0 \text{ 且 } f(2) = 0\}$. 但 $U \cup W = \{f(x) \in V \mid f(1) = 0 \text{ 或 } f(2) = 0\}$ 不是子空间 (比如, $x - 1 + x - 2 \notin U \cup W$). 容易计算, $\dim U = n - 1 = \dim W$, $\dim (U \cap W) = n - 2$. 故仍有 (2.1.1) 式.

定理 2.1.2 (维数定理) 设 V 是线性空间, U 与 W 是 V 的两个子空间. 则

$$\dim (U + W) = (\dim U + \dim W) - \dim (U \cap W). \tag{2.1.2}$$

证 我们仅给出证明的思路, 细节见习题3. 设 $\dim U = s$, $\dim W = t$, $\dim (U \cap W) = r$. 任取 $U \cap W$ 的一组基 $\alpha_1, \alpha_2, \cdots, \alpha_r$. 由于 $U \cap W$ 是 U 与 W 的公共子空间, 故 $U \cap W$ 的基是 U 与 W 的线性无关的向量组, 因此可以扩充成 U 或 W 的基. 设

$$\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_{r+1}, \beta_{r+2}, \cdots, \beta_s$$

与

$$\alpha_1, \alpha_2, \cdots, \alpha_r, \gamma_{r+1}, \gamma_{r+2}, \cdots, \gamma_t$$

分别是 U 与 W 的基. 则

$$\alpha_1, \alpha_2, \cdots, \alpha_r, \beta_{r+1}, \beta_{r+2}, \cdots, \beta_s, \gamma_{r+1}, \gamma_{r+2}, \cdots, \gamma_t$$

是 $U + W$ 的一组基. (为此只需证明该向量组线性无关, 且 $U + W$ 的任何向量均可由这些向量线性表示.) $\square$

由维数定理可知, 欲使子空间 $U+W$ 的维数最大, 必要且只要 $U \cap W = 0$, 亦即 U 与 W 重合的部分最小. 这时我们称和 $U+W$ 是**直和**, 记为 $U \oplus W$. 因此 $\dim (U \oplus W) = \dim U + \dim W$.

例 2.1.4 二维平面 $\mathbb{R}^2$ 是 x 轴与 y 轴 (均是1维子空间) 的直和. 类似地, $\mathbb{R}^3$ 是 x 轴, yoz 平面 (这是一个2维子空间) 的直和.

例 2.1.5 只含奇 (偶) 次项的多项式称为奇 (偶) 多项式. 0多项式既是奇多项式也是偶多项式. 全体奇 (偶) 多项式作成多项式空间的子空间, 称为奇 (偶) 多项式子空间. 多项式空间是奇多项式子空间与偶多项式子空间的直和.

例 2.1.6 n 阶矩阵空间 $M_n(\mathbb{F})$ 是纯量矩阵子空间 $\{A \in M_n | A = \lambda I, \lambda \in \mathbb{F}\}$ 与迹0子空间 $\{A \in M_n | \mathrm{tr} A = 0\}$ 的直和.

定理 2.1.3 (直和的判定) 设 U 与 W 是线性空间 V 的两个子空间, 则下列命题等价:

(1) $U + W$ 是直和 (即 $U \cap W = 0$);

(2) 对任意 $\alpha \in U + W$, 分解式 $\alpha = u + w$, 其中 $u \in U$, $w \in W$ 是唯一的, 即若还有 $\alpha = u' + w'$, 则 $u = u'$, $w = w'$;

(3) 零向量的分解式唯一; 即若 $0 = u + w$, $u \in U$, $w \in W$, 则 $u = w = 0$;

(4) $\dim (U + W) = \dim U + \dim W$.

注: 经常将 定理 2.1.3(3) 作为直和的定义.

证 由直和的定义, (1) 与 (4) 是等价的, 而 (2) 显然蕴涵 (3). 故只需证明 (3)$\Longrightarrow$(1)$\Longrightarrow$(2).

现设 (3) 成立, 而 $\alpha \in U \cap W$, 则 $-\alpha \in U \cap W$, 因此 $0 = \alpha + (-\alpha)$, 但零向量的分解式唯一, 故 $-\alpha = \alpha = 0$, 即 $U \cap W = 0$, 故 (1) 成立.

如果 (1) 成立, 而 $\alpha \in U+W$ 有两个分解式 $\alpha = u+w = u'+w'$, 其中 $u, u' \in U$, $w, w' \in W$, 则 $u - u' = w' - w$, 但 $u - u' \in U$, $w' - w \in W$, 从而 $u - u' = w' - w \in U \cap W = 0$, 即 $u = u', w' = w$, 故 (2) 成立. □

直和的概念可以推广到任意有限多个子空间, 即归纳地定义子空间 $W_1, W_2, \cdots, W_s$ 的直和 $W_1 \oplus W_2 \oplus \cdots \oplus W_s$ 为 $(W_1 + W_2 + \cdots + W_{s-1}) \oplus W_s$. 需要注意的是此时每一个子空间与其余子空间的和的交为 0.

类似两个子空间的直和的判定, 有下述多子空间直和的判定定理, 其证明留做习题.

定理 2.1.4 (多子空间直和的判定) 设 $W_1, W_2, \cdots, W_s$ 是线性空间 V 的子空间, 则下列命题等价:

(1) $W_1+W_2+\cdots+W_s$ 是直和即 $\dim(W_1+W_2+\cdots+W_s) = \dim W_1+\dim W_2+\cdots+\dim W_s$;

(2) $W_j \cap \sum_{k\neq j} W_k = 0, 1 \le j \le s, 1 \le k \le s$;

(3) 任意向量 $\alpha \in W_1 + W_2 + \cdots + W_s$ 的分解式唯一;

(4) 零向量的分解式唯一.

例 2.1.7 三维实空间 $\mathbb{R}^3$ 是 x 轴, y 轴与 z 轴的直和, 即 $\mathbb{R}^3 = \mathbb{R}\vec{i} \oplus \mathbb{R}\vec{j} \oplus \mathbb{R}\vec{k}$.

例 2.1.8 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 n 维线性空间 V 的一组基, 则 $V = \mathbb{F}\alpha_1 \oplus \mathbb{F}\alpha_2 \oplus \cdots \oplus \mathbb{F}\alpha_n$. 因此每个有限维线性空间均可以分解成 1 维子空间的直和.

例 2.1.9 n 阶矩阵空间是严格上三角矩阵子空间, 严格下三角矩阵子空间与对角矩阵子空间的直和.

设 V 是线性空间, U 是 V 的一个子空间. 则存在另一个子空间 W 使得 $V = U \oplus W$(此仅需将 U 的一组基扩充成 V 的一组基即可, 新扩充的部分生成的子空间即是一个 W). W 称为 U 的**补子空间**. 显然 U 的补子空间一般不是唯一的 (什么时候唯一?).

例 2.1.10 设 $V = \mathbb{R}^3$. 则 V 的 1 维子空间就是通过原点的所有直线; 2 维子空间就是通过原点的所有平面. 任何一个 1 维子空间的补子空间可以是任意一个不含该 1 维子空间的 2 维子空间. 反之亦然.

例 2.1.11 常数多项式子空间与常数项为 0 的多项式子空间互为补子空间.

例 2.1.12 设 A 是一个 3 阶实数矩阵, $V = \mathbb{R}^3$. 齐次线性方程组 $Ax = 0$ 的解空间 U 是 V 的一个维数等于 $3 - r(A)$ 的子空间. 其一个补空间恰是 A 的行空间 (由 A 的行向量所生成的子空间, 见下段) $\mathrm{Span}\{A^1, A^2, A^3\}$.

我们需要特别关注与一个矩阵 $A_{m\times n}$ 相联系的下列四个子空间:

(1) A 的零 (化) 空间$N(A)$: 即齐次线性方程组 $Ax = 0$ 的解空间;

(2) A 的列空间(或像空间, 值域)$R(A)$: 即 A 的列向量生成的子空间;

(3) A 的行空间$R(A^T)$: 即 A 的行向量生成的子空间;
(4) A 的左零 (化) 空间$N(A^T)$: 即线性方程组 $y^TA=0$ 或 $A^Tx=0$ 的解空间.

注意, $N(A)$ 与 $R(A^T)$ 是 $\mathbb{F}^n$ 的子空间; 而 $R(A)$ 与 $N(A^T)$ 是 $\mathbb{F}^m$ 的子空间, 且有

$$\dim N(A)+\dim R(A^T)=n; \qquad \dim N(A^T)+\dim R(A)=m.$$

于是, 矩阵的秩就是列空间与行空间的维数. 请注意, 行空间中的向量仍然看成是列向量.

计算矩阵的四个子空间的方法如下:

首先将 A 通过行初等变换化为其Hermite标准形 H_A. 因为 $Ax=0$ 等价于 $H_Ax=0$, 所以 $N(A)=N(H_A)$, 同时, 由于行初等变换不改变列向量之间的线性关系, 从而由 H_A 的列向量的极大线性无关组可得到 A 的列向量的极大线性无关组, 这就是 $R(A)$ 的一组基; 其次, H_A 的行是 A 的行的线性组合, 于是 A 的行空间等于 H_A 的行空间, 即 $R(A^T)=R(H_A^T)$. 但由 H_A 不能同时得出 A 的左零化空间. 为此, 需要记录将 A 化为Hermite标准形时的可逆矩阵 P(参看下例), 即 $PA=H_A$. 设 $r(A)=r$, 则 H_A 的最后 $m-r$ 行均为0, 由矩阵乘法的行向量结构可知, 矩阵 P 的最后 $m-r$ 行是齐次线性方程组 $y^TA=0$ 的线性无关解, 因此它们恰好是 A 的左零化空间 $N(A^T)$ 的一组基.

例 2.1.13 设

$$A=\begin{pmatrix}1&1&2\\0&1&1\\1&2&3\end{pmatrix},$$

求 A 的四个相关子空间.

解 将 A 化为Hermite标准形, 并记录相应的可逆矩阵 P:

$$(A,I_3)=\begin{pmatrix}1&1&2&1&0&0\\0&1&1&0&1&0\\1&2&3&0&0&1\end{pmatrix}\longrightarrow\begin{pmatrix}1&0&1&1&-1&0\\0&1&1&0&1&0\\0&0&0&-1&-1&1\end{pmatrix}=(H_A,P),$$

由此可得 $PA=H_A$, 即

$$\begin{pmatrix}1&-1&0\\0&1&0\\-1&-1&1\end{pmatrix}\begin{pmatrix}1&1&2\\0&1&1\\1&2&3\end{pmatrix}=\begin{pmatrix}1&0&1\\0&1&1\\0&0&0\end{pmatrix}.$$

从而矩阵 A 的秩 $=2$, 而 H_A 的前两列线性无关, 所以 A 的前两列线性无关, 即

$$\begin{aligned}&R(A)=\mathrm{Span}\{A_1,A_2\}=\mathrm{Span}\{(1,0,1)^T,(1,1,2)^T\};\\&R(A^T)=R(H_A^T)=\mathrm{Span}\{(H_A^1)^T,(H_A^2)^T\}=\mathrm{Span}\{(1,0,1)^T,(0,1,1)^T\};\\&N(A)=N(H_A)=\mathrm{Span}\{(-1,-1,1)^T\}.\end{aligned}$$

最后, 为了求出 $N(A^T)$, 可考察矩阵 P, 它的最后一行左乘 A 为零向量, 故它是 $xA=0$ 的解. 于是

$$N(A^T)=\mathrm{Span}\{(P^3)^T\}=\mathrm{Span}\{(-1,-1,1)^T\}.$$

思考题
1. 两个子空间的并何时是子空间?

2. 两个向量张成的子空间的几何意义是什么?

3. 两个子空间的交, 并与和的几何意义分别是什么?

4. 实数域 $\mathbb{R}$ 作为实线性空间的所有子空间是什么? 作为有理数域上的线性空间呢?

5. 复数域 $\mathbb{C}$ 作为实线性空间的所有子空间是什么? 作为复数域上的线性空间呢?

6. 解释 3 阶矩阵 A 的四个子空间的几何意义和相互位置关系.

7. 设 $\mathbb{F} = \mathbb{C}$ 或 $\mathbb{R}$. 则 $\mathbb{F}$ 上的 n 元二次型全体构成 $\mathbb{F}$ 上的线性空间 (第一章第五节思考题 5). 全体半正定二次型是否是该线性空间的子空间? 全体不定二次型呢?

8. 设 $\mathbb{F} = \mathbb{C}$ 或 $\mathbb{R}$. 则 $\mathbb{F}$ 上的 n 维双线性型全体构成 $\mathbb{F}$ 上的线性空间 (第一章第五节思考题 5). 全体 n 维对称双线性型是否是该线性空间的子空间?

第二节 矩阵与线性变换

为了比较两个线性空间, 需要建立两者之间具有一定性质的映射, 最基本的性质当然是 (T1) 与 (T2). 以实数域 $\mathbb{R}$ 为例, 设 f 是从 $\mathbb{R}$ 到 $\mathbb{R}$ 自身的满足 (T1) 与 (T2) 的映射 (此时是普通的函数). 则由 (T2) 可得, 对任意 $x \in \mathbb{R}, f(x) = xf(1)$, 于是 $f(x) = kx$, 即是某正比例函数或线性函数, 故线性变换可以看成是线性函数的推广. 现若 f 仍满足 (T1) 与 (T2), 而仅将其定义域 $\mathbb{R}$ 变为 $\mathbb{R}^2$, 值域仍为 $\mathbb{R}$, 则容易得到 (请计算!)$f(x,y) = ax + by$, 其中 $a = f(1,0)$ 而 $b = f(0,1)$. 这是二元线性函数. 进一步, 如果将 f 的定义域与值域 $\mathbb{R}$ 均变为 $\mathbb{R}^2$ 会发生什么情况? 满足条件 (T1) 与 (T2) 的对应是否仍然存在呢?

例 2.2.1 考虑平面坐标系中坐标轴的旋转. 假定坐标轴逆时针旋转了 θ 角度, 于是原坐标系中的点 $P(x,y)$ 将变为 $P'(x',y')$. 现在的问题是两种坐标间的关系, 此即所谓转轴公式:

$$\begin{aligned} x &= \quad x'\cos\theta + y'\sin\theta, \\ y &= -x'\sin\theta + y'\cos\theta. \end{aligned} \tag{2.2.1}$$

可以看出, 如果点 P 与 Q 分别被变到点 P' 与 Q', 则点 $P+Q$ 与 kP (平面可以定义向量加法与数乘从而构成线性空间!) 将被分别变到点 $P'+Q'$ 与 kP'! 现将任意一个点 $P(x,y)$ 记成 α, 旋转后的点记成 $f(\alpha)$, 则有

$$f(\alpha+\beta) = f(\alpha) + f(\beta); \quad f(k\alpha) = kf(\alpha).$$

这就是 (T1) 与 (T2). 另外, 转轴公式中的两个等式各自表示一个二元线性函数, 因此平面坐标系中的旋转变换是两个二元线性函数.

例 2.2.2 考察定义在 $\mathbb{R}$ 上的全体无限次可微函数的集合 V(这是一个无限维实线性空间) 的求导运算 $\partial : f(x) \mapsto f'(x)$, 它是 V 到自身的一个映射. 由高等数学或数学分析,

$$\partial(f+g) = \partial(f) + \partial(g); \quad \partial(kf) = k\partial(f).$$

类似地, 变上限的定积分 $f(x) \mapsto \int_a^x f(x)\,\mathrm{d}\,x$ 也给出 V 到自身的一个满足 (T1) 与 (T2) 的映射, 只不过此时“自变量”是函数, 函数值也是函数.

例 2.2.3 考察 n 阶线性常微分方程

$$y^{(n)} + f_1(x)y^{(n-1)} + f_2(x)y^{(n-2)} + \cdots + f_{n-1}(x)y' + f_n(x)y = 0, \tag{2.2.2}$$

其中 $y^{(i)} = \frac{\mathrm{d}^i y}{\mathrm{d}\,x^i}, i = 1, 2, \cdots, n$. 如果将方程的左边记为 $L(y)$, 则 $L : y \mapsto L(y)$ 给出全体 $\geq n$ 阶可导函数之集合到全体实函数之集合的映射 (二者均是无限维实线性空间), 显然 L 满足 (T1) 与 (T2). 微分方程 (2.2.2) 的解恰好是映射 L 的“零点”集 $\{y : L(y) = 0\}$.

例 2.2.4 考虑函数 $f(x)$ 的 Laplace 变换

$$L(f(x)) = \int_{-\infty}^{+\infty} f(t)e^{-ixt}\,\mathrm{d}\,t.$$

则 L 满足 (T1) 与 (T2). (读者可以自证其反演公式也满足这两个条件.)

可以看出, 满足条件 (T1) 与 (T2) 的映射不但基本而且广泛存在, 于是有下面的定义.

定义 2.2.1 设 U 与 V 是两个线性空间. U 到 V 内的一个映射 σ 如果满足可加性条件 (T1) 与齐次性条件 (T2), 则称 σ 是 U 到 V 的线性变换 (linear transformation) 或线性映射 (linear map).

一般将 U 到自身的线性变换称为线性算子 (本书将不区分线性变换、线性映射和线性算子这三个名词), U 到 V 的线性变换全体记为 $\mathrm{Hom}(U,V)$ (或更精确地, 记为 $\mathrm{Hom}_{\mathbb{F}}(U,V)$). 特别, 将 $\mathrm{Hom}(V,V)$ 记为 $\mathrm{End}\,V$, 而将 $\mathrm{Hom}(V,\mathbb{F})$ 记为 V^*, 称为 V 的**对偶空间或共轭空间**.

例 2.2.5 对偶空间的最简单例子是 $\mathrm{Hom}_{\mathbb{R}}(\mathbb{R},\mathbb{R}) = \mathbb{R}^*$, 即由全体正比例实函数构成的集合, 因为每个正比例函数的几何意义是通过原点的直线, 故对偶空间也可以理解为全体通过原点的直线. 稍后我们将看到, 这样的理解实际上恰好解释了“对偶”一词.

设 $\sigma \in \mathrm{Hom}(U,V)$. 则当 σ 作为映射是单的 (或满的) 时, 称 σ 是单变换 (或满变换). 既单又满的变换称为同构. 如果存在同构 $\sigma \in \mathrm{Hom}(U,V)$, 则 U 与 V 称为是同构的线性空间, 记为 $U \cong V$.

注 1. 可加性条件 (T1) 与齐次性同阶 (T2) 等价于:

$$\sigma(a\alpha + b\beta) = a\sigma(\alpha) + b\sigma(\beta), \quad a,b \in \mathbb{F},\ \alpha,\beta \in U. \tag{2.2.3}$$

满足上述公式的映射称为“保持线性性质”(因此线性变换就是“保持线性性质”的映射). 重复使用上述公式, 即可导出计算上非常有用的“线性叠加原理”

$$\sigma(a_1\alpha_1+\cdots+a_s\alpha_s) = a_1\sigma(\alpha_1)+\cdots+a_s\sigma(\alpha_s), a_j \in \mathbb{F},\ \alpha_j \in U. \tag{2.2.4}$$

一般来说, 称函数或映射 f 是 r- 齐次的是指存在固定的常数 r 使得对任意的 x 均有 $f(kx) = k^r(x)$. 但如果 f 还满足可加性条件, 则 r- 齐次性必是齐次性, 即有 $r = 1$, 请读者自证.

注 2. 如果 $\mathbb{F}$ 是有理数域$\mathbb{Q}$, 则容易证明可加性蕴涵齐次性. 因为由 (T1) 可知, 对任意正整数 n,m 有

$$f(2x) = f(x+x) = 2f(x),\ f(nx) = f(x+(n-1)x) = nf(x).$$

故有

$$f(nx) = f(m \times (\frac{n}{m}x)) = mf(\frac{n}{m}x),$$

即可得

$$f(\frac{n}{m}x) = \frac{n}{m}f(x),$$

即对任意正有理数 r 有

$$f(rx) = rf(x).$$

但由 $f(0) = f(0+0) = f(0) + f(0)$, 故有 $f(0) = 0$, 从而 $0 = f(x + (-x)) = f(x) + f(-x)$, 即有 $f(-x) = -f(x)$, 故知条件 T(2) 对所有有理数成立.

请读者举例说明, 即使 $\mathbb{F}$ 是有理数域 $\mathbb{Q}$, 齐次性条件 (T2) 也不必蕴涵可加性条件 (T1).

一般将满足可加性的映射称为加性映射, 对非有理数域的其它数域 $\mathbb{F} \subseteq \mathbb{C}$, 加性映射未必是齐次的, 如下例.

例 2.2.6 设 $\mathbb{F} = \mathbb{Q}(\sqrt{2})$, 则数域 $\mathbb{F}$ 是其自身上的 1 维线性空间. 对任意 $\alpha = a + b\sqrt{2} \in \mathbb{F}, a, b \in \mathbb{Q}$, 定义

$$\sigma(\alpha) = \sigma(a + b\sqrt{2}) = a,$$

则 σ 显然是 $\mathbb{F}$ 到自身的一个加性映射, 但 σ 不满足齐次性条件, 因为

$$\sigma(\sqrt{2} \bullet \sqrt{2}) = \sigma(2) = 2 \text{ 而 } \sqrt{2}\sigma(\sqrt{2}) = 0.$$

例 2.2.7 (非线性变换的例子) 平面上的极坐标变换

$$\begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \tag{2.2.5}$$

不是线性变换. 类似地, 空间中的柱面坐标变换与球面坐标变换也不是线性变换.

由定义立即可得线性变换如下的简单性质.

命题 2.2.1 (1) $\sigma(0) = 0$; $\sigma(-\alpha) = -\sigma(\alpha)$;
(2) 若 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 线性相关, 则 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s)$ 也线性相关;
(3) 若 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s)$ 线性无关, 则 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 也线性无关.
(结论 (2) 与 (3) 的逆命题见习题 15.)

注意性质 $\sigma(0) = 0$ 的几何意义就是线性变换必须保持原点不动, 因此平面或空间解析几何中的平移变换一般不是线性变换 (何时是线性变换?).

定理 2.2.1 (线性变换的构造) 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是线性空间 U 的一组基, $\beta_1, \beta_2, \cdots, \beta_n$ 是线性空间 V 的任意 n 个向量, 则唯一地存在一个线性变换 σ 使得 $\sigma(\alpha_j) = \beta_j$, $1 \leq j \leq n$.

证明思路. 如果线性变换 σ 满足条件, 则 σ 必须将 U 中任意向量 $\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n$ 映到

$$\sigma(\alpha) = k_1\sigma(\alpha_1) + k_2\sigma(\alpha_2) + \cdots + k_n\sigma(\alpha_n) = k_1\beta_1 + k_2\beta_2 + \cdots + k_n\beta_n.$$

容易验证, 如上定义的映射 σ 确是一个 U 到 V 的满足条件 $\sigma(\alpha_j) = \beta_j$, $1 \leq j \leq n$ 的线性变换, 且是唯一可能满足这些条件的线性变换. □

注 1. 上述证明中构造线性变换 σ 的方法是基本的, 称为线性扩展法, 即确定一组基的像, 再对任意的向量, 定义其像就是相应于基元素的像的线性组合.

注 2. 定理 2.2.1 告诉我们, 要确定一个定义域为 n 维线性空间的线性变换, 只需要知道该线性变换在一组基下的像就足够了, 因此, 表达一个线性变换, 只需要列出其在一组基下的像即可. 这有些类似于确定一个一元 n 次多项式只需要知道该多项式在 $n+1$ 个不同点的值就行了.

例 2.2.8 (线性函数) 为了确定一个从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的线性变换 σ, 我们只需要知道它在任意一组基下的值, 为简单起见, 假设我们知道了 σ 在标准基下的值, 设

$$\sigma(e_i) = a_i,\, 1 \le i \le n\,.$$

则对任意 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$, 有

$$\sigma(x) = \sigma(\sum_{i=1}^{n} x_i e_i) = \sum_{i=1}^{n} x_i \sigma(e_i) = \sum_{i=1}^{n} a_i x_i.$$

因此从 $\mathbb{R}^n$ 到 $\mathbb{R}$ 的线性变换 σ 恰好就是 n 元线性函数. 换句话说, $\mathbb{R}^n$ 的对偶空间 $(\mathbb{R}^n)^*$ 就是 n 元线性函数的集合. 容易看出, $(\mathbb{R}^n)^*$ 在函数的自然加法和数乘下构成线性空间.

实际上, 任何线性空间 V 的对偶空间 V^* 都有一个自然的线性空间结构. 更一般地, 稍后我们将看到 $\mathbb{F}$ 上的线性空间 U 到 V 的全体线性变换之集合 $\mathrm{Hom}_{\mathbb{F}}(U,V)$ 都有一个自然的 $\mathbb{F}$ 线性空间结构.

下面介绍一些特殊的线性变换:

(1) 零变换: 将线性空间 V 的所有向量均变为 $0\,(\in V)$ 向量的变换, 称为零变换, 记为 0 (更确切地, 0_V); 即对任意 $\alpha \in V$, 有 $0(\alpha) = 0$.

(2) **恒等变换**: 将线性空间 V 中任意向量均变为自己的变换, 称为恒等变换或单位变换, 记为 I_V(或简单地记为 I 或 1); 即对任意 $\alpha \in V$, $I(\alpha) = \alpha$. 恒等变换显然是 V 到自身的同构 (称为 V 的**自同构**).

(3) **位似**: 设 $k \in \mathbb{F}$. 将线性空间 V 的任意向量 α 变为 $k\alpha$ 的变换 σ 称为 (伸缩) 系数为 k 的位似, 即 $\sigma(\alpha) = k\alpha$. 零变换与恒等变换分别是 $k = 0$ 与 $k = 1$ 时的位似. 非零位似均是自同构.

(4) **可逆变换**: 设线性变换 $\sigma \in \mathrm{Hom}(U, V)$, 如果存在线性变换 $\tau \in \mathrm{Hom}(V, U)$ 使得对任意 $\alpha \in U$, $\beta \in V$, 均有 $\tau(\sigma(\alpha)) = \alpha$, $\sigma(\tau(\beta)) = \beta$, 则称 σ 是可逆线性变换, τ 称为其逆变换. 可以证明, 如果 σ 的逆变换存在则必定唯一, 其逆记为 σ^{-1}, 且 $(\sigma^{-1})^{-1} = \sigma$(习题 16).

比如, 任何系数为 $k(k \neq 0)$ 的位似均是可逆的, 且其逆变换是系数为 k^{-1} 的位似. 又如, 例 2.2.1 中的旋转变换就是线性空间 $\mathbb{R}^2$ 的可逆线性变换, 其逆变换就是顺时针旋转 θ 角度或逆时针旋转 $-\theta$ 角度.

例 2.2.9 由于 $\mathbb{R}$ 上的线性变换就是线性函数 (即正比例函数), 故 $\mathbb{R}$ 上的可逆线性变换 f 的逆变换就是 f 的反函数 f^{-1}.

例 2.2.10 设 $A \in M_n(\mathbb{F})$. 考虑 $\mathbb{F}^n$ 上的由 A 定义的线性变换 $\sigma_A : x \mapsto Ax$, 则 σ_A 可逆 $\Longleftrightarrow$ 矩阵 A 可逆, 且此时 $\sigma^{-1} : x \mapsto A^{-1}x$, 即 $(\sigma_A)^{-1} = \sigma_{A^{-1}}$.

例 2.2.11 (相似变换) 设 P 是 n 阶可逆矩阵, 则

$$X \mapsto P^{-1}XP \tag{2.2.6}$$

是矩阵空间 $M_n(\mathbb{F})$ 的一个自同构, 称为由 P 诱导的相似变换或共轭变换. 任何矩阵 A 在相似变换下的像是与 A 相似的矩阵.

和线性变换 $\sigma \in \mathrm{Hom}(U,V)$ 密切相关的有两个集合, 一个是其“零点”集 $\{\alpha \in U \mid \sigma(\alpha) = 0\}$, 称为 σ 的核, 记为 $\mathrm{Ker}(\sigma)$ 或 $\sigma^{-1}(0)$; 另一个是其“函数值”的集合, 即 $\{\alpha \in V \mid \exists \beta \in U 使得 \alpha = \sigma(\beta)\}$, 称为 σ 的像, 记为 $\mathrm{Im}(\sigma)$ 或 $\sigma(U)$. 易证, $\mathrm{Ker}\sigma$ 与 $\mathrm{Im}\sigma$ 分别是 U 与 V 的子空间 (见习题 17), 其维数分别记为 $\eta(\sigma)$ 与 $r(\sigma)$, 称为 σ 的**零度**与秩.

例 2.2.12 设 $U = \mathbb{F}^n$, $V = \mathbb{F}^m$, A 是一个 $m \times n$ 矩阵. 对任意 $x \in U$, 定义 U 到 V 的线性变换 σ 为 $\sigma(x) = Ax$. 则 σ 的核就是 A 的零空间, σ 的零度恰好等于 $n - r(A)$ (故此数又称为矩阵 A 的零度); σ 的像空间就是 A 的列空间 $R(A)$, σ 的秩就是 A 的秩 $r(A)$.

例 2.2.13 设 U 是 V 的子空间. 对任意 $x \in U$, 定义 U 到 V 的映射 ι 为 $\iota(x) = x$. 则 ι 是线性变换, 常称为**包含映射** . 显然, 包含映射是单的.

例 2.2.14 设 U, V 是 $\mathbb{F}$ 上的线性空间 (不必有限维), $\sigma \in \mathrm{Hom}(U,V)$. 定义 U 到线性空间 $\mathrm{Im}(\sigma)$ 的映射 $\widetilde{\sigma}$ 为 $\widetilde{\sigma}(x) = \sigma(x)$. 则 $\widetilde{\sigma}$ 是线性变换. 显然, $\widetilde{\sigma}$ 是满的. 于是, 由一个线性变换 σ 可以诱导出如下三个线性变换

$$\mathrm{Ker}(\sigma) \xrightarrow{\iota} U \xrightarrow{\widetilde{\sigma}} \mathrm{Im}(\sigma) \xrightarrow{\iota} V$$

其中两个 ι 均为包含映射 (同一符号不同映射).

利用核空间与像空间可以给出单变换与满变换的简洁刻画.

定理 2.2.2 设 U, V 是 $\mathbb{F}$ 上的线性空间 (不必有限维), $\sigma \in \mathrm{Hom}(U,V)$. 则
(1) σ 是单的 $\iff$ $\mathrm{Ker}(\sigma) = 0$;
(2) σ 是满的 $\iff$ $\mathrm{Im}(\sigma) = V$;
(3) σ 是同构 $\iff$ σ 可逆.
特别地, 如果 $U = V$ 是有限维线性空间, $\sigma \in \mathrm{End}V$, 则 σ 是单的 $\iff$ σ 是满的 $\iff$ σ 可逆.

证 我们仅给出 (3) 的证明框架, 其余见习题 16. 显然, 可逆线性变换必然既是单映射也是满映射, 即可逆线性变换必是同构. 反过来, 定义同构变换 $\sigma \in \mathrm{Hom}(U,V)$ 的逆变换 τ 如下: 对任意 $\beta \in V$, 由于 σ 是满的, 故存在 $\alpha \in U$ 使得 $\sigma(\alpha) = \beta$, 令

$$\tau : \beta \mapsto \alpha.$$

由于 σ 是单的, 故上面的定义是合理的. 验证 $\tau\sigma = I_U, \quad \sigma\tau = I_V$ 以及 τ 是线性变换即可. □

设 $\sigma \in \mathrm{Hom}(U,V)$, 我们知道, σ 由其在任何一组基下的像唯一确定. 因此我们需要了解两件事, 第一是这组基的像是什么? 第二是每个向量 $\alpha \in U$ 的像是什么?

设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 U 的一组基, $\alpha_1', \alpha_2', \cdots, \alpha_m'$ 是 V 的一组基. 设

$$\begin{gathered}
\sigma(\alpha_1) = a_{11}\alpha_1' + a_{21}\alpha_2' + \cdots + a_{m1}\alpha_m', \\
\sigma(\alpha_2) = a_{12}\alpha_1' + a_{22}\alpha_2' + \cdots + a_{m2}\alpha_m', \\
\cdots\cdots\cdots \\
\sigma(\alpha_n) = a_{1n}\alpha_1' + a_{2n}\alpha_2' + \cdots + a_{mn}\alpha_m',
\end{gathered}$$

或用矩阵形式表达为

$$(\sigma(\alpha_1),\sigma(\alpha_2),\cdots,\sigma(\alpha_n))=(\alpha'_1,\alpha'_2,\cdots,\alpha'_m)A, \tag{2.2.7}$$

其中 $A=(a_{ij})\in\mathbb{F}^{m\times n}$ 称为 σ 关于 $\alpha-$ 基和 $\alpha'-$ 基的矩阵. 特别地, 如果 $U=V$, $\alpha-$ 基等于 $\alpha'-$ 基, 则 σ 关于 $\alpha-$ 基和 $\alpha-$ 基的矩阵简称为 σ 关于 $\alpha-$ 基的矩阵.

公式 (2.2.7) 的左端一般简记为 $\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)$, 即有

$$\sigma(\alpha_1,\alpha_2,\cdots,\alpha_n)=(\alpha'_1,\alpha'_2,\cdots,\alpha'_m)A.$$

(这样的简写有两个含义, 其一是简化记号, 其二是强调“线性变换”是作用在“向量”上而非系数上 (线性变换的齐次性)! 从而系数或坐标或坐标形成的矩阵都可以拿到线性变换的外面!)

公式 (2.2.7) 描述了一组基在线性变换下的像, 由此立即可得任何向量在该线性变换下的像, 即有下述

定理 2.2.3 (线性变换下的坐标变换) 设线性变换 $\sigma\in\mathrm{Hom}(U,V)$ 在 $\alpha-$ 基和 $\alpha'-$ 基下的矩阵为 A, 向量 $\alpha\in U$ 在 $\alpha-$ 基下的坐标为 x, 则 $\sigma(\alpha)$ 在 $\alpha'-$ 基下的坐标为 Ax.

例 2.2.15 为计算 例 2.2.1 的旋转变换 σ 在标准基 $e_1=(1,0)^T, e_2=(0,1)^T$ 下的矩阵 A, 需要计算 $\sigma(e_1)$, $\sigma(e_2)$. 直接计算可得, $\sigma(e_1)=(\cos\theta,\sin\theta)^T$, $\sigma(e_2)=(-\sin\theta,\cos\theta)^T$. 故

$$A=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}.$$

例 2.2.16 设 $V=\mathbb{F}^n$, σ 是 V 的线性变换, 其在标准基 $e_1, e_2,\cdots, e_n$ 下的矩阵为 A. 则对任意 $\alpha=(x_1,x_2,\cdots,x_n)^T=(e_1,e_2,\cdots,e_n)(x_1,x_2,\cdots,x_n)^T$, 有

$$\begin{aligned}\sigma(\alpha)&=\sigma(e_1,e_2,\cdots,e_n)(x_1,x_2,\cdots,x_n)^T\\&=(e_1,e_2,\cdots,e_n)A(x_1,x_2,\cdots,x_n)^T=A(x_1,x_2,\cdots,x_n)^T.\end{aligned}$$

即 $\sigma:\alpha\mapsto A\alpha$. 换句话说, $\mathbb{F}^n$ 的线性变换几乎就是左乘一个矩阵 (只要选取标准基). 进一步, 设 $A=(a_{ij})$, 则

$$\begin{aligned}\sigma(\alpha)&=A(x_1,x_2,\cdots,x_n)^T\\&=(a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n,\cdots,a_{n1}x_1+a_{n1}x_2+\cdots+a_{nn}x_n)^T\\&=(\sigma_1(\alpha),\cdots,\sigma_n(\alpha)),\end{aligned}$$

其中每个 $\sigma_j(\alpha)=a_{j1}x_1+a_{j2}x_2+\cdots+a_{jn}x_n$, $1\le j\le n$ 恰好是 n 元线性函数; 换句话说, 线性变换不过是多个多元线性函数而已.

上例具有普遍意义, 即对有限维线性空间而言, 线性变换与矩阵是一回事. 稍后我们将确切地描述这个事实.

由于线性空间有不同的基, 自然要问线性变换在不同基下的矩阵有何联系? 为此, 设

$$\{\alpha_1,\,\alpha_2,\cdots,\,\alpha_n\},\quad \{\beta_1,\,\beta_2,\cdots,\,\beta_n\}$$

是 U 的两组基, P 是由 α- 基到 β- 基的过渡矩阵; 设

$$\{\alpha_1', \alpha_2', \cdots, \alpha_m'\}, \quad \{\beta_1', \beta_2', \cdots, \beta_m'\}$$

是 V 的两组基, Q 是由 α'– 基到 β'- 基的过渡矩阵. 设 $\sigma \in \mathrm{Hom}(U, V)$ 关于 α– 基和 α'- 基的矩阵为 A, 关于 β– 基和 β'– 基的矩阵为 B. 设 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)x$, 问 $\sigma(\alpha)$ 关于 β'– 基的坐标是什么?

首先, α 关于 β– 基的坐标为 $P^{-1}x$; 再按公式 (2.2.7), $\sigma(\alpha)$ 关于 α'– 基的坐标为 Ax, 关于 β'– 基的坐标为 $BP^{-1}x$. 故由坐标变换公式得, $Q^{-1}Ax = BP^{-1}x$. 故 $Q^{-1}A = BP^{-1}$, 即 $QB = AP$ 或

$$B = Q^{-1}AP. \tag{2.2.8}$$

实际上, 线性空间的由 α- 基到 α'- 基的过渡矩阵也是一个线性变换 σ 在 α-α'- 基下的矩阵, 只要定义 σ 将 α- 基的第 i 个基向量对应到 α'- 基的第 i 个基向量即可. 因此公式 (2.2.8) 不过是说下面的图是交换图 (即两组用矩阵表达的线性变换的合成相等)

$$\begin{tikzcd}
(\alpha - \text{基})U \arrow[r, "\sigma", "A"'] \arrow[d, "P^{-1}"'] & V(\beta - \text{基}) \arrow[d, "Q^{-1}"] \\
(\alpha' - \text{基})U \arrow[r, "B", "\sigma"'] & V(\beta' - \text{基})
\end{tikzcd}$$

图 2.2.1

注 公式 (2.2.8) 说明线性变换在 α-α'- 基下的矩阵与在 β-β'- 基下的矩阵是等价的; 反过来, 如果两个同阶矩阵 A 与 B 是等价的, 即公式 (2.2.8) 成立, 则可以构造相应的线性变换 σ, 使得 A 与 B 分别是 σ 在 α-α'- 基与 β-β'- 基下的矩阵, 见本节思考题 5.

现假定 $U = V$ 且 α- 基等于 β- 基, α'- 基等于 β'- 基, 于是图 2.2.1 中的矩阵 P 与 Q 相等, 则公式 (2.2.8) 变为

$$B = P^{-1}AP. \tag{2.2.9}$$

这就是下面的

定理 2.2.4 设 V 是 n 维线性空间, σ 是 V 的一个线性变换. 设 $\alpha_1, \cdots, \alpha_n$ 与 $\beta_1, \cdots, \beta_n$ 是 V 的两组基, A 与 B 分别是 σ 关于该两组基的矩阵. 则 A 与 B 相似.

该定理揭示了相似矩阵 $B = P^{-1}AP$ 的实质, 即它们不过是同一线性变换在不同基下的矩阵而已, 而矩阵 P 不过是两个基之间的过渡矩阵. 因此相似矩阵具有相同的特征值就不难理解了. 由于相似矩阵具有相同的行列式和迹, 因此我们将线性变换 σ 在任意一组基下的矩阵的行列式和迹称为 σ 的行列式与迹, 分别记为 $|\sigma|$ 与 $\mathrm{tr}(\sigma)$.

例 2.2.17 设 σ 是平面 $\mathbb{R}^2$ 的旋转 $\pi/4$ 的变换. 则 σ 在标准基下的矩阵为

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

易知 σ 在基 $(1,0)^T,(1,1)^T$ 下的矩阵为

$$B=\begin{bmatrix} 0 & -\sqrt{2} \\ \frac{1}{\sqrt{2}} & \sqrt{2} \end{bmatrix}.$$

由于标准基到基 $(1,0)^T,(1,1)^T$ 的过渡矩阵为

$$T=\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

所以 $AT=TB$, 或 $B=T^{-1}AT$. 因此 $|\sigma|=1$, $\mathrm{tr}(\sigma)=\sqrt{2}$.

什么是相似矩阵? 公式 (2.2.9) 告诉我们, 同一线性变换在不同基下的矩阵是相似矩阵. 反过来, 如果矩阵 A 与 B 相似, 即满足公式 (2.2.9), 则取 $\sigma\in\mathrm{End}V$ 使得 σ 在 α- 基下的矩阵为 A; 现如下选取 V 的另一组 α'- 基, 使由 α- 基到 α'- 基的过渡矩阵为 P, 则可断言 σ 在 α'- 基下的矩阵为 B! 从而"相似矩阵就是同一线性变换在不同基下的矩阵!"可以反过来提上面的问题, 即是否有这样的两个线性变换 σ 与 τ, 它们在不同基下的矩阵相同? 见习题 18. (当然, 它们在同一组基下的矩阵必然不同.)

例 2.2.18 设 $\sigma\in\mathrm{End}\mathbb{R}^2$ 是向 x- 轴的投影:

$$\sigma((x,y)^T)=(x,0)^T.$$

则 σ 在标准基 (α- 基) 下的矩阵为

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

而 σ 在基 (α'- 基)$\{\beta_1=(1,1)^T,\beta_2=(1,-1)^T\}$ 下的矩阵为

$$B=\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

由公式 (2.2.9) 可知, $B=P^{-1}AP$, 且矩阵 P 正是由 α- 基到 α'- 基的过渡矩阵

$$P=\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

(由于 α'- 基是标准基, 故过渡矩阵 P 就是矩阵 (β_1,β_2).)

请注意, 矩阵 A 与 B 均满足条件 $A^2=A$, $B^2=B$. 这样的矩阵称为**幂等矩阵**. 显然, σ 在任意一组基下的矩阵均是幂等矩阵.

下面我们将建立线性空间的分类定理和线性变换与矩阵联系的基本定理.

定理 2.2.5 (同构定理) 域 $\mathbb{F}$ 上的两个线性空间 U 与 V 同构 $\iff \dim_{\mathbb{F}}U=\dim_{\mathbb{F}}V$.

证明思路. 必要性是显然的, 因为同构变换必然将 U 的一组基变为 V 的一组基, 因此 U 与 V 的维数相同. 反之, 若 U 与 V 的维数相同, 则将 U 的一组基变为 V 的一组基的线性变换必然是同构. 证明细节见习题 19. $\square$

定理2.2.5 是线性空间理论中最重要的定理之一, 它表明, 域 $\mathbb{F}$ 上的任意 n 维线性空间 V 均与 $\mathbb{F}^n$ 同构, 即它们具有完全相同的线性空间的结构, 可以认为它们的唯一差别仅是元素命名不同, 因此对 V 的研究可以归结到对 $\mathbb{F}^n$ 的研究, 而后者不仅表达最为简单, 也是我们最熟悉的. 见下面的几个例子.

例 2.2.19 设 $U = \mathbb{F}[x]_n, V = \mathbb{F}^n$(为简便起见, 本例 $\mathbb{F}^n$ 中的向量记为行向量). 因为它们的维数均为 n, 故由 定理2.2.5 知 U 与 V 同构. 这就是说, U 与 V 的差别仅是 U 中的向量被称为“多项式”而已. 考虑 U 与 V 的如下同构映射 σ:

$$\sigma: f(x) = \sum_{i=0}^{n-1} a_i x^i \mapsto (a_0, a_1, \cdots, a_{n-1}).$$

则 σ 可以看成是一个“重起名字”的操作, 即将 U 中的向量 $f(x) = \sum\limits_{i=0}^{n-1} a_i x^i$ 重新命名为 (或者“重新记为”) $(a_0, a_1, \cdots, a_{n-1})$. 这种“重新命名”实际上没有改变 U 的任何结构, 即 U 中的加法与数乘得以完全保留 (比如, 多项式 $1 + x$ 与 $2 - x + x^2$ 的和现在变为向量 $(1, 1, 0, \cdots 0)$ 与 $(2, -1, 1, 0, \cdots, 0)$ 的和, 前者的结果是 $3 + x^2$, 后者是 $(3, 0, 1, 0, \cdots, 0)$. 但我们知道, 后者就是多项式 $3 + x^2$ 的“新名字”). 这就是说, U 与 V 作为线性空间实际上是一回事! (在我们心目中多项式与 n 元数组应该有很大差别, 那是因为我们给多项式附加了许多线性空间以外的东西, 诸如多项式的根, 一元二次多项式的图像是抛物线等等, 但多项式的这些性质都不能由多项式的加法和数乘导出.)

例 2.2.20 设 A 是非空有限集合, $|A| = n$. 则 $\mathbb{F}^A \cong \mathbb{F}^n$; 故函数空间 $\mathbb{F}^A$ 的 (线性) 结构与我们熟悉的线性空间 $\mathbb{F}^n$ 完全一致 (参照第一章习题24).

例 2.2.21 第一章例1.4.2中的实线性空间 $V = \{$ 全体正实数 $\}$ 是几维的? 注意, 由于 V 不是 $\mathbb{R}$ 的子空间, 故我们不能直接推出 $\dim V \leq 1$. 但容易证明映射 $x \mapsto \log x$ 是 V 到 $\mathbb{R}$ 的一个同构变换, 故由 定理2.2.5 知 $\dim V = \dim \mathbb{R} = 1$. 另外, 我们还知道, 全体正实数的乘法运算就是全体实数的加法运算, 而正实数的幂就是普通的实数乘法 (自然, 此时每一个正实数 x 都有一个新名字 $\log x$).

以下我们考察线性变换的整体结构. 首先我们将建立 $\mathrm{Hom}(U, V)$ 的线性空间结构. 对于 $\mathrm{Hom}(U, V)$ 中的两个线性变换, 可以自然地定义加法以及数乘两个运算.

设 $\sigma, \tau \in \mathrm{Hom}\,(U, V)$, $\alpha \in V$, $k \in \mathbb{F}$, 定义

$$\sigma + \tau : \alpha \mapsto \sigma(\alpha) + \tau(\alpha); \quad k\sigma : \alpha \mapsto k\sigma(\alpha). \tag{2.2.10}$$

显然, 如上 (2.2.10) 定义的 $\sigma + \tau$ 与 $k\sigma$ 仍是 U 到 V 的线性变换, 分别称为 σ 与 τ 的和以及数 k 与 σ 的数乘. 如此, 我们在 $\mathrm{Hom}\,(U, V)$ 中定义了向量的加法与数乘. 易证, $\mathrm{Hom}\,(U, V)$ 在上述加法与数乘下, 作成 $\mathbb{F}$ 上的一个线性空间. 设 $\dim_{\mathbb{F}} U = n, \dim_{\mathbb{F}} V = m$, 则 (证明见习题21)

$$\dim_{\mathbb{F}} \mathrm{Hom}(U, V) = mn = (\dim_{\mathbb{F}} U)(\dim_{\mathbb{F}} V). \tag{2.2.11}$$

进一步, 如果 $U = V$, 则还可以定义 $\mathrm{End}\, V$ 中两个线性变换 σ 与 τ 的乘法

$$\sigma\tau : \alpha \mapsto \sigma(\tau(\alpha)), \quad \forall\, \alpha \in V \tag{2.2.12}$$

需要注意, 与线性变换的加法不同, 由公式 (2.2.12) 定义的线性变换的乘法实际上是映射的合成, 因此不具有交换性. 线性变换 σ 与自己的乘积 $\sigma\sigma$ 记为 σ^2. 归纳地, 对任意自然数 k, 可以定义 σ 的 k 次幂 $\sigma^k = \sigma^{k-1}\sigma$ (为方便记, 规定 $\sigma^0 = I$). 继而, 对 $\mathbb{F}$ 上的任意多项式 $f(x)$, 可以定义线性变换 σ 的多项式 $f(\sigma)$. 可以证明 (请读者自证), 它是一个可以和 σ 交换的线性变换, 即 $\sigma f(\sigma) = f(\sigma)\sigma$.

类似于矩阵的情形, 满足 $\sigma^2 = \sigma$ 的线性变换称为幂等变换. 满足 $\sigma^k = 0$ 的线性变换称为**幂零变换**(且使此式成立的最小自然数称为 σ 的**幂零指数**). 零变换与恒等变换都是幂等变换.

例 2.2.22 设 $V = \mathbb{F}^n$. 对 $\alpha = (x_1, x_2, \cdots, x_n)^T$, 定义

$$\sigma(\alpha) = (x_1, 0, \cdots, 0)^T;\ \tau(\alpha) = (x_2, x_3, \cdots, x_n, 0)^T.$$

则 σ 与 τ 均是 V 的线性变换, 且 σ 是幂等变换, τ 是幂零变换 (幂零指数是多少?).

例 2.2.23 零变换在任何基下的矩阵都是零矩阵. 恒等变换在任何基下的矩阵都是单位矩阵. 位似变换在任何基下的矩阵都是同一纯量矩阵. 幂等变换在任何基下的矩阵都是幂等矩阵. 幂零变换在任何基下的矩阵都是幂零矩阵.

例 2.2.24 幂等变换与投影变换等价. 首先, 投影变换显然是幂等的; 反之, 如果线性空间 V 的线性变换 σ 是幂等的, 则 $V = \mathrm{Im}(\sigma) \oplus \mathrm{Ker}(\sigma)$ (请证明!), 因此 σ 就是 V(沿子空间 $\mathrm{Ker}(\sigma)$) 向子空间 $\mathrm{Im}(\sigma)$ 上的投影变换.

定理 2.2.6 (矩阵与线性变换) 设 V 是 $\mathbb{F}$ 上的 n 维线性空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一组基. 设 $M_n(\mathbb{F})$ 是 $\mathbb{F}$ 上全体 n 阶矩阵组成的线性空间. 对任意 $\sigma \in \mathrm{End}\,V$, 记 $A(\sigma)$ 是 σ 在该基下的矩阵. 定义 $\mathrm{End}\,V$ 到 $M_n(\mathbb{F})$ 的映射 ψ 为

$$\begin{array}{cccc} \psi: & \mathrm{End}\,V & \longrightarrow & M_n(\mathbb{F}) \\ & \sigma & \longmapsto & A(\sigma) \end{array}$$

则 ψ 是一个保持运算 (加法, 数乘与乘法) 的一一映射, 即满足下列条件:

(1) $A(\sigma + \tau) = A(\sigma) + A(\tau)$;

(2) $A(k\sigma) = kA(\sigma), \forall\, k \in \mathbb{F}$;

(3) $A(\sigma\tau) = A(\sigma)A(\tau)$;

(4) σ 可逆 $\iff$ $A(\sigma)$ 可逆; 且此时 $A(\sigma)^{-1} = A(\sigma^{-1})$;

(5) $A(0) = 0$; $A(I) = I$.

证 此处仅证明 ψ 是一一映射, 其余见习题 22. 由前两条可知 ψ 是线性变换. 为证 ψ 是单的, 只需证明 $\mathrm{Ker}(\psi) = 0$. 设 $\psi(\sigma) = 0$, 即 $A(\sigma) = 0$. 对任意 $\alpha \in V$, 设 x 是 α 在该基下的坐标, 则由 定理 2.2.3 可知 $\sigma(\alpha)$ 的坐标为 $A(\sigma)x = 0x = 0$. 于是 $\sigma = 0$. 再证 ψ 是满的. 设 B 是任意 n 阶矩阵. 令 $\sigma(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_n)Bx$. 则显然 σ 是 V 的线性变换, 且 σ 在该基下的矩阵是 B, 从而 $\psi(\sigma) = B$, 即 ψ 是满的. □

例 2.2.25 设 $\sigma \in \mathrm{End}\mathbb{R}^2$. 则对任意 $\alpha = (x\,, y)^T \in \mathbb{R}^2$, 有

$$\sigma(\alpha) = \sigma(xe_1) + \sigma(ye_2) = x\sigma(e_1) + y\sigma(e_2).$$

记 $\sigma(e_1)=(a_{11},\,a_{12})^T$, $\sigma(e_2)=(a_{21},\,a_{22})^T$. 则

$$\sigma(\alpha)=x\sigma(e_1)+y\sigma(e_2)=\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix} x\\ y\end{pmatrix}.$$

于是 σ 与矩阵

$$\begin{pmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

确实是一回事.

定义了适当乘法 (即自身满足结合律并且关于加法满足分配律, 关于数乘满足结合律) 的 ($\mathbb{F}$) 线性空间称为 ($\mathbb{F}$)**代数**. 而保持两个代数所有运算的映射称为**同态**, 既单又满的同态称为同构. 因此 $\operatorname{End} V$ 被称为是 V 的自同态代数. 于是 定理2.2.6 中的映射 ψ 是两个代数 $\operatorname{End} V$ 与 $M_n(\mathbb{F})$ 之间的同构映射, 而 $\operatorname{End} V$ 与 $M_n(\mathbb{F})$ 是同构的代数. 换句话说, 有限维线性空间的线性变换就是矩阵! 我们知道 $\operatorname{End} V$ 中的乘法乃是映射的合成, 因此矩阵的乘法实际上是线性变换的合成, 这就解释了矩阵乘法的许多"怪异", 诸如无交换性, 有零因子等等.

设 V 是 n 维线性空间, 其自同态代数 $\operatorname{End} V$ 中的可逆线性变换具有重要意义, 其全体构成的集合记为 $\operatorname{Aut} V$, 称为 V 的自同构群, 其中的群运算是线性变换的乘积 (即合成). 由 定理2.2.6 可知, $\operatorname{Aut} V$ 与全体 n 阶可逆矩阵构成的集合 $GL_n(\mathbb{F})$ 相对应, 后者按矩阵的乘法也做成群, 称为 ($\mathbb{F}$ 上的) 一般线性群, 该群无论在理论上还是应用上都具有无比重要的作用. 类似于线性空间, 可以定义群的同构, 则知 $\operatorname{Aut} V$ 与 $GL_n(\mathbb{F})$ 是同构的群.

对 定理2.2.6 略加修改即可得到如下的线性变换基本定理 (证明见习题23):

定理 2.2.7 (线性变换基本定理) 设 U 与 V 分别为 n 维与 m 维 $\mathbb{F}$ 线性空间. 则 $\operatorname{Hom}_{\mathbb{F}}(U,V)\cong\mathbb{F}^{m\times n}$. 特别地, V^* 与 V 同构.

下面的例子解释了 V^* 与 V 同构的本质, 请比较 例2.2.5.

例 2.2.26 设 $V=\mathbb{R}^n$. 则 $V^*=(\mathbb{R}^n)^*$ 是全体 n 元线性函数形成的线性空间. 我们如下建立这两个线性空间之间的一个同构 $*$. 对 $\mathbb{R}^n$ 中的每一个向量 (点)$v=(v_1,v_2,\cdots,v_n)$, 定义 $*(v)$ 是如下的线性函数: 对任意 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$, 规定

$$*(v):(x_1,x_2,\cdots,x_n)\mapsto v^Tx=v_1x_1+v_2x_2+\cdots+v_nx_n$$

如果我们将线性函数 $*(v)$ 记作 v^*, 则 $\mathbb{R}^n$ 与 $(\mathbb{R}^n)^*$ 之间的这个同构 $*$ 把 $\mathbb{R}^n$ 中的每个向量 v 对应到 $(\mathbb{R}^n)^*$ 的一个以向量 $(-1,v)$ 为法向量的超平面 v^*. 如下图:

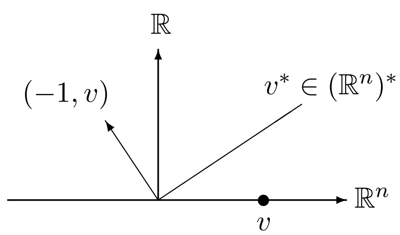

图2.2.2

因此, 线性空间与其对偶空间的同构实际上是将点看作 (超) 平面. 比如, $\mathbb{R}^2$ 与 $(\mathbb{R}^2)^*$ 之间的这个同构 $*$ 将平面上的点 (向量)(A,B) 对应到 $(\mathbb{R}^2)^*$ 中的线性函数 $z=Ax+By$, 这恰好是 $\mathbb{R}^3$ 中以向量 $(-1,A,B)$ 为法向量的 (超) 平面.

例 2.2.27 (对偶空间的**对偶基**) 设 n 维线性空间 V 的一组基为 $\alpha_1,\alpha_2,\cdots,\alpha_n$, 则可定义其对偶空间 V^* 的一组基 $\alpha_1^*,\alpha_2^*,\cdots,\alpha_n^*$ 如下:

$$\alpha_i^*(\alpha_j)=\delta_{ij}, 1\leq i,\, j\leq n. \tag{2.2.13}$$

这组基称为 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 的对偶基. 例如, 关于 $\mathbb{R}^2$ 的标准基 e_1, e_2 的对偶基为 e_1^*, e_2^*, 其中

$$e_1^*(x,\,y)=x,\quad e_2^*(x,\,y)=y.$$

由 例2.2.8, 设 $f(x,\,y)=a_1x+a_2y$ 是 $(\mathbb{R}^2)^*$ 中的任意一个向量 (即 $\mathbb{R}^2$ 上的线性函数), 则 $f=a_1e_1^*+a_2e_2^*$.

例 2.2.28 (双重对偶) 线性空间 V 的对偶空间 V^* 的对偶空间 $(V^*)^*$ 称为 V 的**双重对偶**或二次对偶, 记为 V^{**}. 这实际上是由线性函数的线性函数 (即所谓线性泛函) 构成的线性空间. 由线性变换基本定理, 如果 V 是有限维线性空间, 则 $V\cong V^*\cong V^{**}$. 但实际上 V 与 V^{**} 间有更为自然的同构, 即不依赖于基的选取. 对任意 $v\in V$, $f\in V^*$, 定义 V 到 V^{**} 的映射 ϕ 如下:

$$\phi:\, v\mapsto\, \phi(v):f\mapsto\, f(v).$$

则 ϕ 是 V 到 V^{**} 的同构变换, 证明见习题24.

思考题

1. 是否有可加性与齐次性等价的情形?

2. 平面 (即 $\mathbb{R}^2$) 上的线性变换能否将直线变为抛物线或者椭圆? 能否将抛物线或者椭圆变为直线? 空间 (即 $\mathbb{R}^3$) 中的线性变换能否将平面变为直线? 能否将抛物线变为直线或者椭圆?

3. 如何建立空间中的过原点的直线和平面上的过原点的直线之间的同构映射?

4. 解释关于线性变换的公式 (2.2.8) 与矩阵的等价之间的关系?

5. 以线性变换的观点解释列满秩与行满秩矩阵以及矩阵的满秩分解.

6. 设 V 是1维线性空间, 则 $\mathrm{End}V$ 与 $\mathrm{Aut}V$ 是什么? 特别, 什么是 $\mathrm{End}\mathbb{R}$, $\mathrm{Aut}\mathbb{R}$, $\mathrm{End}\mathbb{C}$, $\mathrm{Aut}\mathbb{C}$?

7. 有限维线性空间上的单线性变换就是满线性变换, 此结论对无限维线性空间成立吗?

8. 同构 $\mathbb{R}\cong\mathbb{R}^*$ 与 $\mathbb{R}^3\cong(\mathbb{R}^3)^*$ 有何自然含义?

9. 设 V 是空间中满足 $x+y+z=0$ 的子空间, V 的对偶空间是什么?

第三节 内积空间的正交分解

在内积空间中, 每个子空间除去有无穷多个补子空间外, 还有唯一一个“好的”补子空间, 即所谓**正交补**. 设 V 是 n 维内积空间, U 是 V 的子空间. 令 $W=\{\alpha\in V\,|\,(\alpha,\beta)=0,\forall\beta\in U\}$. 容易证明 (习题30), W 是 V 的一个子空间, 称为 U 的正交补, 记为 $U^\perp$. 显然有下面的

定理 2.3.1 设 V 是 n 维内积空间, U 是 V 的子空间. 则 $V=U\oplus U^\perp$. 特别, $\dim U^\perp=n-\dim U$.

实际上, 对内积空间 V 的任何子空间 U, 正交补 $U^{\perp}$ 总是存在的: 为此只须将 U 的一个标准正交基 (极大线性无关标准正交向量组)$\alpha_1, \alpha_2, \cdots, \alpha_s$ 扩充为 V 的一个标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n$ 即可知,

$$U^{\perp} = \mathrm{Span}\{\alpha_{s+1}, \alpha_{s+2}, \cdots, \alpha_n\}.$$

另外, 若 $U^{\perp} = W$, 则 $W^{\perp} = U$, 即正交补还是对称的: $(U^{\perp})^{\perp} = U$.

显然, 可以对内积空间 V 中的任何非空子集定义其正交补. 特别, 单个非零向量 $v \in V$ 的正交补记为 $v^{\perp}$, 由 定理2.3.1 可知 $\dim v^{\perp} = \dim V - 1$, 即 $v^{\perp}$ 是 V 的一个超平面.

回顾任意 $m \times n$ 矩阵 A 的四个相关子空间:

$N(A)$ 与 $R(A^*)$ 是 $\mathbb{F}^n$ 的子空间; 而 $R(A)$ 与 $N(A^*)$ 是 $\mathbb{F}^m$ 的子空间. 并且, $N(A)$ 与 $R(A^*)$ 互补; $R(A)$ 与 $N(A^*)$ 互补; 实际上, 它们还是正交 (垂直) 的. 即有

定理 2.3.2 (1) $N(A) = R(A^*)^{\perp}$; (2) $R(A) = N(A^*)^{\perp}$.

证 设 $A \in \mathbb{F}^{m\times n}$. 我们只证 (1), (2) 的证明类似, 见习题31. 设 $x \in N(A)$, 即 $Ax = 0$, 亦即 $A^{(i)}x = 0\,, i = 1, 2, \cdots, m$. 由 $\mathbb{C}^n$ 中的 (普通) 内积定义, 这就是 $\bar{A}$(注意: 不是 A!) 的每个行向量 $\overline{A^{(i)}}$ 与 x 正交. 从而 $\bar{A}$ 的行空间的每个向量也与 x 正交. 故 $x \in R(A^*)^{\perp}$. 反之亦然. □

例 2.3.1 (Fourier 系数) 设 $V = \{f(x)\,|\,f(x) = \sum\limits_{n=0}^{m}(a_n\cos nx + b_n \sin nx), \forall a_n, b_n \in \mathbb{R}\}$. 则 V 是一个 $2m+1$ 维实线性空间. 定义其上的内积为:

$$(f(x), g(x)) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)g(x)\,\mathrm{d}\,x.$$

则 $\frac{1}{2}$, $\cos x$, $\sin x$, $\cos 2x$, $\sin 2x$, $\cdots$, $\cos mx$, $\sin mx$ 构成 V 的一组标准正交基 (这是内积定义中乘以系数 $\frac{1}{\pi}$ 的唯一理由). 对任何非零常函数 c 有 (见习题34)

$$c^{\perp} = 1^{\perp} = \mathrm{Span}\{\cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cdots,\ \cos mx,\ \sin mx\}.$$

我们在数学分析或高等数学课程中已经知道

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,\mathrm{d}\,x, \quad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,\mathrm{d}\,x.$$

实际上, $a_n = (f(x), \cos nx)$ 与 $b_n = (f(x), \sin nx)$ 恰好是 $f(x)$ 与该组标准正交基的每个基向量的内积.

正交补空间有众多应用, 此处我们仅介绍“最佳近似定理”与其在“广义相对论”中的应用. 先引入一个定义.

定义 2.3.1 (最佳近似) 设 U 是欧氏空间 V 的子空间, $\beta \in V$. 如果 $\alpha \in U$ 满足下面的条件:

$$\|\beta - \alpha\| \le \|\beta - \gamma\|, \quad \forall\, \gamma \in U.$$

则称 α 是 β 在 U 中的**最佳近似 (向量)**, 常记为 $\hat{\beta}$(如果不需要强调子空间 U).

由上述定义可知, β 在子空间 U 中的最佳近似 (向量) 就是 U 中与 β 最近 (距离最短) 的向量, 这样的向量一定存在而且唯一 (为什么?).

例 2.3.2 (几何意义) 设 U 是 $\mathbb{R}^3$ 的一个二维子空间即 U 是一个通过原点的平面, 则由三垂线定理, 任何向量 $\alpha = (x, y, z)$ 在 U 上的最佳近似就是 α 在平面 U 上的投影向量, 即 $\hat{\beta} = \mathrm{Proj}_U\alpha$.

例 2.3.3 在 $\mathbb{R}^3$ 中, 任何向量 $\alpha = (x, y, z)$ 在子空间 x- 轴上的最佳近似为 $(x\,, 0, 0)$.

定理 2.3.3 (最佳近似定理) 设 U 是欧氏空间 V 的子空间, $\beta \in V, \alpha \in U$. 则 α 是 β 在 U 上的最佳近似向量 $\iff \beta - \alpha \in U^\perp$.

证 设 γ 是 U 中任一向量, 注意 $\beta - \gamma = (\beta - \alpha) + (\alpha - \gamma)$, 且 $\beta - \alpha \in U^\perp$, $\alpha - \gamma \in U$, 所以 $\|\beta - \gamma\|^2 = \|\beta - \alpha\|^2 + \|\alpha - \gamma\|^2 \geq \|\beta - \alpha\|^2$. 因此, 对任意 $\gamma \in U$, 有

$$\|\beta - \alpha\| \leq \|\beta - \gamma\|,$$

即 α 是 β 在 U 上的最佳近似向量. □

由最佳近似定理, 求向量 β 在子空间 U 中的最佳近似向量, 等价于求 β 在直和分解 $V = U \oplus U^\perp$ 下的表达式, 即若 $\beta = \alpha + \gamma$, 其中 $\alpha \in U, \gamma \in U^\perp$, 则 α 就是 β 在 U 上的最佳近似向量 $\hat{\beta}$. 那么, 如何求得这样的分解呢? 先考察 $U = \mathbb{R}v$ 是 1 维子空间的情形. 设向量 β 在子空间 U 中的最佳近似为 $\hat{\beta} = xv$, 则由最佳近似定理, 内积 $(\beta - xv, v) = 0$, 故

$$x = \frac{(\beta, v)}{(v, v)} \tag{2.3.1}$$

因此 $\hat{\beta} = \frac{(\beta,v)}{(v,v)}v$, 这恰好是 β 在 v 上的投影向量 $\mathrm{Proj}_v\beta$, 而 $\gamma = \beta - \mathrm{Proj}_v\beta$ 恰是 β 关于子空间 U 的正交投影向量 (该向量在 U 的正交补中). 如果 $U = \mathbb{R}v_1 \oplus \mathbb{R}v_2$ 是 2 维子空间, 则

$$\hat{\beta} = x_1v_1 + x_2v_2 \tag{2.3.2}$$

且内积 $(\beta - \hat{\beta}, v_i) = 0,\ i = 1, 2$, 即有下面的方程组

$$(\beta - x_1v_1 - x_2v_2, v_i) = 0,\ i = 1, 2 \tag{2.3.3}$$

由于 v_1, v_2 线性无关, 故方程组 (2.3.3) 有唯一解, 由此即可求得 $\hat{\beta}$. 特别, 若 v_1, v_2 还是正交的, 则

$$x_1 = \frac{(\beta, v_1)}{(v_1, v_1)},\ x_2 = \frac{(\beta, v_2)}{(v_2, v_2)} \tag{2.3.4}$$

因此 $\hat{\beta}$ 在表达式 (2.3.2) 中的两个分量 x_1v_1 与 x_2v_2 恰好分别是 β 在 v_1 与 v_2 上的投影向量 $\mathrm{Proj}_{v_1}\beta$ 与 $\mathrm{Proj}_{v_2}\beta$. 故此时 β 在子空间 U 中的最佳近似就是 β 在 U 的各个正交基向量的投影向量之和 $\mathrm{Proj}_{v_1}\beta + \mathrm{Proj}_{v_2}\beta$. 一般地, 若 U 是任何非 0 子空间, 则有下述计算最佳近似的结论 (证明见习题 35)

命题 2.3.1 设 β 是欧氏空间 V 的一个向量, U 是 V 的一个子空间, $\alpha_1, \cdots, \alpha_s$ 是 U 的一个正交基, 则 β 在 U 上的最佳近似向量为

$$\hat{\beta} = \mathrm{Proj}_{v_1}\beta + \cdots + \mathrm{Proj}_{v_s}\beta = \frac{(\beta, \alpha_1)}{(\alpha_1, \alpha_1)}\alpha_1 + \cdots + \frac{(\beta, \alpha_s)}{(\alpha_s, \alpha_s)}\alpha_s. \tag{2.3.5}$$

特别, 若 $\alpha_1, \cdots, \alpha_s$ 还是 U 的一个标准正交基, 则 β 在 U 上的最佳近似向量为

$$\hat{\beta} = (\beta, \alpha_1)\alpha_1 + (\beta, \alpha_2)\alpha_2 + \cdots + (\beta, \alpha_s)\alpha_s. \tag{2.3.6}$$

根据 命题2.3.1 即可知欧氏空间中的向量 β 在任何非零子空间 U 中的投影向量 $\mathrm{Proj}_U\beta$ 为其最佳近似向量, 即

$$\mathrm{Proj}_U\beta = \hat{\beta}.$$

而 命题2.3.1 可以重新表述为"向量 β 在子空间 U 中的最佳近似就是其在 U 上的投影向量, 就是其在 U 的一组正交基的各个基向量上的投影向量之和".

例 2.3.4 在 $\mathbb{R}^3$ 中, 任何向量 $\alpha = (x, y, z)$ 在子空间 xoy 平面中的最佳近似为 $(x, y, 0)$, 此恰好是向量 α 在 xoy 平面的标准正交基 (即坐标轴)$\{(1,0,0),(0,1,0)\}$ 上的投影向量 $(x,0,0)$ 与 $(0,y,0)$ 之和.

最佳近似向量的一个应用是求方程组的近似解. 设 $Ax = b$ 是一个矛盾方程, 即对任何 $x = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$ 均有 $Ax \neq b$. 此时, 如果存在向量 $x^0 \in \mathbb{R}^n$ 使得对任意向量 $x \in \mathbb{R}^n$, 均有 $\|Ax^0 - b\| \leq \|Ax - b\|$, 则称 x^0 是方程的一个**最优解**. (注: 该解即是函数 $y = \|Ax - b\|$ 的最小值点, 而此时的函数值恰好是"点"b 到"直线"或"平面"(或更一般地"子空间")$R(A)$ 的最小距离. 由于函数 $y = \|Ax - b\|$ 与函数 $y = \|Ax - b\|^2 = (Ax - b)^T(Ax - b) = x^TA^TAx - 2x^TAb + b^Tb$ 的最小值点相同, 而后者更方便求解, 因此, 最优解也称为"最小二乘解".)

对任何向量 $x \in \mathbb{R}^n$, 向量 Ax 总是属于矩阵 A 的列空间 $R(A)$. 因此矛盾方程 $Ax = b$ 的最优解 x^0 必须使 Ax^0 为向量 b 在子空间 $R(A)$ 上的最佳近似向量. 由最佳近似定理可知, $Ax^0 - b \in R(A)^\perp = N(A^T)$. 所以, 向量 $Ax^0 - b$ 应满足方程 $A^T(Ax^0 - b) = 0$, 故向量 x^0 应满足方程

$$A^TAx = A^Tb. \tag{2.3.7}$$

此方程称为矛盾方程 $Ax = b$ 的**正规化方程**. 由第一章习题5易知, 方程 (2.3.7) 总是相容的! 因此, 最优解总是存在的, 且当方程组 $Ax = b$ 相容时, 最优解就是该方程组的解! 故有下述

命题 2.3.2 方程组 $A^TAx = A^Tb$ 与方程组 $Ax = \mathrm{Proj}_{R(A)}b$ 同解, 即方程组 $Ax = b$ 的最优解就是方程组 $Ax = \mathrm{Proj}_{R(A)}b$ 的解.

所以问题最终归结为如何表示投影向量 $\mathrm{Proj}_{R(A)}b$. 我们将在第六章第一节利用矩阵的广义逆讨论此问题.

例 2.3.5 已知变量 b 是变量 x_1, x_2 的函数. 现有观测数据为:

b	x_1	x_2
-1	1	-6
2	1	-2
1	1	1
6	1	7

求线性近似公式: $b = a_1x_1 + a_2x_2$.

解 按上述数据, 可得方程组

$$\begin{cases} a_1 & -6a_2 & = -1 \\ a_1 & -2a_2 & = \ 2 \\ a_1 & +\ a_2 & = \ 1 \\ a_1 & +7a_2 & = \ 6 \end{cases}$$

用矩阵表示为 $Ay = \beta$, 其中

$$A = \begin{pmatrix} 1 & -6 \\ 1 & -2 \\ 1 & 1 \\ 1 & 7 \end{pmatrix}, \quad \beta = \begin{pmatrix} -1 \\ 2 \\ 1 \\ 6 \end{pmatrix}, \quad y = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix},$$

易知这是一个矛盾方程, 故可求其最优解. 直接的方法是求其正规化方程的解. 但注意到 A 的两列 A_1, A_2 是正交的, 故由公式 (2.3.5) 知 β 在 $R(A)$ 上的最佳近似为

$$\hat{\beta} = \frac{(\beta, A_1)}{(A_1, A_1)}A_1 + \frac{(\beta, A_2)}{(A_2, A_2)}A_2 = \frac{8}{4}A_1 + \frac{45}{90}A_2.$$

故由 命题 2.3.2 知方程组 $Ay = \beta$ 的唯一 (为什么唯一?) 最优解就是组合系数

$$a_1 = 2, \quad a_2 = 1/2.$$

故所求近似公式为 $b = 2x_1 + \frac{1}{2}x_2$.

以下我们介绍正交补在广义相对论中的应用.

对于实线性空间 V, 如果在其内积的定义中去掉正定性而代之以非退化性即

若 $(\alpha\,, v) = 0, \forall v \in V$, 则 $\alpha = 0$

则称 V 是一个 **Minkowski** [18] **空间**, 而该内积称为 **Minkowski 内积**. 注意非退化性只是防止出现平凡的内积:

$$(\alpha\,, \beta) = 0, \forall\, \alpha\,, \beta \in V.$$

欧氏空间当然是 Minkowski 空间, 因为内积的正定性蕴涵非退化性. 由于没有正定性, Minkowski 空间中的向量与自身的内积 (称为该向量的范数) 可以为负数. 广义相对论使用的是 4 维 Minkowski 空间 V (称为 Minkowski 4- 空间), 其一组正交基记为 e_0, e_1, e_2, e_3, 但

$$-(e_0)^2 = (e_1)^2 = (e_2)^2 = (e_3)^2 = 1.$$

因此, 两个向量 $\alpha = (x_0, x_1, x_2\,, x_3)^T$ 与 $\beta = (y_0, y_1, y_2\,, y_3)^T$ 的内积为

$$(\alpha, \beta) = -x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3,$$

[18] Hermann Minkowski(1864-1909), 著名俄裔德国数学家, 创造数的几何 (geometry of numbers), 为相对论奠定了数学基础, 是 Albert Einstein (爱因斯坦) 的老师.

从而向量 $\alpha=(x_0,\,x_1,\,x_2\,,x_3)^T$ 的范数为

$$\alpha^2=-x_0^2+x_1^2+x_2^2+x_3^2.$$

根据其范数的符号, 向量被分成三类:

(1) 类空的 (spacelike): $\alpha^2>0$;

(2) 类时的 (timelike): $\alpha^2<0$;

(3) 类光的 (lightlike): $\alpha^2=0$.

所有类光向量的集合称为 (一个事件的) 光锥 (light cone).

类似于向量, V 的子空间 W 也被分为上述三类:

(a) 类空的 (spacelike): $\alpha^2>0, \forall\,\alpha\in W$;

(b) 类光的 (lightlike): $\alpha^2\geq 0, \forall\,\alpha\in W$ 且存在 $0\neq\alpha\in W$ 使得 $\alpha^2=0$;

(c) 类时的 (timelike): 其它.

命题 2.3.3 (正交性与相对论) 设 W 是 Minkowski 4- 空间 V 的子空间. 则
(1) W 是类时的 $\iff$ $W^\perp$ 是类空的, 反之亦然;
(2) W 是类光的 $\iff$ $W\cap W^\perp\neq 0$ $\iff$ $W^\perp$ 是类光的.

下面是极有趣的一条:

命题 2.3.4 (平行 = 垂直) 两个类光向量正交 $\iff$ 它们平行 (即成比例).

思考题

1. 试用正交分解理论解释勾股定理.

2. 试利用正交分解理论在空间中建立关于面积的勾股定理. 能否建立更高维的勾股定理?

3. 最优解何时唯一?

4. 如果在 $\mathbb{R}^3$ 中定义“广义内积” $(x,y)=x_1y_1+x_2y_2-x_3y_3$, 则正交性有何变化? 是否存在非零向量 x 与自己正交?

第四节 内积空间中的线性变换

在内积空间中, 线性变换的性质更加丰富有趣, 比如可以问线性变换将三角形变成了什么图形? 是否仍然将抛物线变为抛物线? 等等. 此时最本质的问题实际上是线性变换是否保持长度, 角度和距离 (内积空间中的距离概念见第一章习题 38). 于是有下面的定义.

定义 2.4.1 设 V 是内积空间, $\sigma\in\operatorname{End}V$. 如果 σ 保持向量间的距离, 即对任意 $\alpha,\beta\in V$, 均有 $d(\sigma(\alpha),\sigma(\beta))=d(\alpha,\beta)$, 则称 σ 是**等距变换** 或保距变换.

等距变换必然是可逆变换. 恒等变换是等距变换; $\mathbb{R}^2$ 中的旋转也是等距变换; 但比例系数的模不等于 1 的位似变换均非等距变换.

定理 2.4.1 设 V 是内积空间, $\sigma\in\operatorname{End}V$. 则 σ 是等距变换 $\iff$ σ 保持向量的长度 $\iff$ σ 保持内积.

证 对任意 $\alpha \in V$, 由于 $\|\alpha\| = \sqrt{(\alpha,\alpha)} = d(\alpha, 0)$, 故知保持距离等价于保持长度. 另外, 保持内积显然蕴含保持长度. 现设 σ 保持长度, 则

$$\begin{aligned}(\sigma(\alpha+\beta), \sigma(\alpha+\beta)) &= (\sigma(\alpha),\sigma(\alpha)) + (\sigma(\alpha),\sigma(\beta)) + (\sigma(\beta),\sigma(\alpha)) + (\sigma(\beta),\sigma(\beta)) \\ &= (\alpha,\alpha) + (\sigma(\alpha),\sigma(\beta)) + \overline{(\sigma(\beta),\sigma(\alpha))} + (\beta,\beta).\end{aligned}$$

而

$$\begin{aligned}(\alpha+\beta, \alpha+\beta) &= (\alpha, \alpha) + (\alpha,\beta) + (\beta, \alpha) + (\beta, \beta) \\ &= (\alpha,\alpha) + (\alpha, \beta) + \overline{(\beta, \alpha)} + (\beta,\beta).\end{aligned}$$

故

$$(\sigma(\alpha), \sigma(\beta)) + \overline{(\sigma(\alpha), \sigma(\beta))} = (\alpha, \beta) + \overline{(\alpha, \beta)},$$

即 $(\sigma(\alpha), \sigma(\beta))$ 与 (α, β) 具有相同的实部. 但容易验证 (请验证!) $\mathrm{Re}(\alpha, \mathrm{i}\beta) = \mathrm{Im}(\alpha, \beta)$, 即 $(\alpha, \mathrm{i}\beta)$ 的实部即是 (α, β) 的虚部, 从而 $(\sigma(\alpha), \sigma(\beta))$ 与 (α, β) 也具有相同的虚部, 因此 $(\sigma(\alpha), \sigma(\beta)) = (\alpha, \beta)$, 故定理成立. □

定理 2.4.2 设 V 是 n 维内积空间, $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的一组标准正交基, $\sigma \in \mathrm{End}\, V$, A 是 σ 在该组基下的矩阵. 则 σ 是等距变换 $\iff$ A 是酉矩阵.

证 对任意 $\alpha \in V$, 设 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)x$, $x \in \mathbb{C}^n$, 则因 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是标准正交基, 故其长度为 $\|\alpha\| = \sqrt{(\alpha,\alpha)} = \sqrt{x^*x}$. 设 $\sigma(\alpha) = \beta = (\alpha_1, \alpha_2, \cdots, \alpha_n)Ax$, 则 $\|\beta\| = \sqrt{(Ax)^*(Ax)} = \sqrt{x^*(A^*A)x}$. 故若 $A^*A = I$, 则 $\|\sigma(\alpha)\| = \|\beta\| = \sqrt{x^*x} = \|\alpha\|$, 即 σ 保持长度, 故由 定理 2.4.1 知 σ 是等距变换.

反过来, 设 σ 是等距变换, 则 σ 保持长度, 故对任意向量 $x \in \mathbb{C}^n$(或 $\mathbb{R}^n$), 有

$$(\beta, \beta) = x^*(A^*A)x = (\alpha,\alpha) = x^*x.$$

故 Hermite 二次型 $x^*(I - A^*A)x = 0$ 对一切向量 x 成立, 故知 $I - A^*A = 0$(见第一章习题 46), 即 $A^*A = I$. 所以 A 是酉矩阵. □

需要注意, 定理 2.4.2 中的条件“标准正交”不可去, 即等距变换在一般基下的矩阵未必是酉矩阵, 见下例.

例 2.4.1 设 σ 为实平面上逆时针旋转 $\pi/2$, 则它在基 $\{e_1, e_1 + e_2\}$ 下的矩阵为

$$\begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}.$$

这显然不是酉矩阵.

由于 定理 2.4.2, 欧氏空间的等距变换又称为**正交变换**；而复内积空间的等距变换也称为**酉变换**.

显然等距变换之积仍是等距变换, 等距变换之逆仍是等距变换. 从而等距变换的全体组成 $\mathrm{Aut}\, V$ 的一个子群, 称为等距变换群. 特别, 欧氏空间的全体正交变换构成正交变换群. 相应地, 全体正交矩阵组成 $GL_n(\mathbb{R})$ 的一个子群, 称为正交矩阵群.

定理 2.4.3 设 V 是内积空间, $\sigma \in \mathrm{End}\, V$. 则 σ 是等距变换 $\iff$ σ 将标准正交基变为标准正交基.

证 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 V 的任意一组标准正交基, A 是 σ 在该组基下的矩阵, $\beta_j = \sigma(\alpha_j)$, $1 \leq j \leq n$. 若 σ 是等距变换, 则由 定理2.4.1, σ 保持内积与长度, 故 $\beta_1, \beta_2, \cdots, \beta_n$ 仍是一个标准正交基.

反过来, 如果 σ 将标准正交基变为标准正交基, 则 σ 关于任何标准正交基的矩阵是标准正交基到标准正交基的过渡矩阵, 而由第一章第五节思考题4, 可知这样的过渡矩阵是酉矩阵, 从而由 定理2.4.2 知 σ 是等距变换. □

例 2.4.2 (二维正交矩阵) 实直线 $\mathbb{R}$ 上的正交变换只有 $\pm I$, 即恒等变换及其负变换 (将 x 变为 $-x$). 实平面 $\mathbb{R}^2$ 上的正交变换可以计算如下: 选定标准基 e_1, e_2. 设正交变换 σ 在标准基下的矩阵为 A, 则由 定理2.4.3 知 A 是正交矩阵, 于是

$$Q = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{2.4.1}$$

或

$$P = \begin{pmatrix} c & s \\ s & -c \end{pmatrix} \tag{2.4.2}$$

其中 $c^2 + s^2 = 1$. 矩阵 Q 对应的正交变换是我们熟悉的旋转, 而矩阵 P 既是正交矩阵又是对称矩阵, 其对应的正交变换是一种反射(对称轴为 $y = \frac{1-c}{s}x$, 见习题39), 故称为反射矩阵.

例 2.4.3 (三维正交矩阵) 对 $\mathbb{R}^3$, 正交矩阵必正交相似于下面6种矩阵 (见习题40):

$$\begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & \\ & 1 & \\ & & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix},$$

$$\begin{pmatrix} -1 & & \\ & -1 & \\ & & -1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta \end{pmatrix}, \quad \begin{pmatrix} -1 & & \\ & \cos\theta & \sin\theta \\ & -\sin\theta & \cos\theta \end{pmatrix}.$$

一般地, 将三维空间保持长度的映射称为刚体运动. 上述结论表明, 三维空间的刚体运动是平移和上述6种变换的合成. 如果上面的矩阵有复数特征值, 则意味着相应的正交变换是: 绕一固定轴的旋转, 且旋转的平面与轴正交.

例 2.4.4 设

$$A = \begin{pmatrix} \frac{1}{4} & \frac{\sqrt{3}}{4} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ \frac{\sqrt{3}}{4} & \frac{3}{4} & -\frac{1}{2} \end{pmatrix}$$

是正交矩阵, 求相应的正交变换的旋转轴与旋转的角度.

解 A 的一个特征值为1, 设其对应的特征向量为 $\alpha = (x_1, x_2, x_3)^T$, 则由方程组 $A\alpha = \alpha$ 可得 $\alpha = (\sqrt{3}, 1, 1)^T$, 故旋转轴为过点 $(0,0,0)$ 与 $(\sqrt{3}, 1, 1)$ 的直线.

为求 A 的旋转角, 可利用矩阵的相似不变量 – 迹, 即设

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix} = Q^{-1}AQ,$$

则 $\operatorname{tr} B = \operatorname{tr} A$, 故 $1 + 2\cos\theta = \frac{1}{4} - \frac{1}{2} - \frac{1}{2} = -\frac{3}{4}$, 所以 $\cos\theta = -\frac{7}{8}$.

现在可以回答本节开始的几个问题了. 由于保距变换保持长度和内积, 从而也保持向量之间的角度, 因此保距变换不改变图形的形状, 比如保距变换将圆仍变为圆, 将三角形仍变为三角形等等. 但一般的线性变换则可能将图形变得面目全非. 例如, 向 x 轴的投影变换把圆心在原点的单位圆变成了 x 轴上的线段, 而将抛物线 $y = x^2$ 变为整个 x 轴. 其它一些简单图形在线性变换下的形状, 见习题 40-42.

最著名的正交变换当属 **Householder** [19] **变换**与 **Givens** [20] **旋转**, 分别介绍如下.

例 2.4.5 (Householder 变换) 设 $v \in \mathbb{C}^n$ 是非零向量, 定义 n 阶复矩阵 H 如下:

$$H = I - \frac{2vv^*}{v^*v}. \tag{2.4.3}$$

矩阵 H 称为 **Householder 矩阵**. 由矩阵 H 定义的线性变换 H_v 称为 Householder 变换, 即对任意 $x \in \mathbb{C}^n$

$$H_v x = x - \frac{2vv^*}{v^*v} x. \tag{2.4.4}$$

容易看出, Householder 变换 H_v 实际上是关于超平面$v^\perp$ 的镜面反射(即将 $v^\perp$ 的向量固定, 而将 v 变为 $-v$, 详见习题 43). 如果将向量 x 按直和 $\mathbb{C}^n = \operatorname{Span}\{v\} \oplus v^\perp$ 正交分解为

$$x = P_v x + P_v^\perp x,$$

则

$$H_v x = P_v^\perp x - P_v x.$$

如下图.

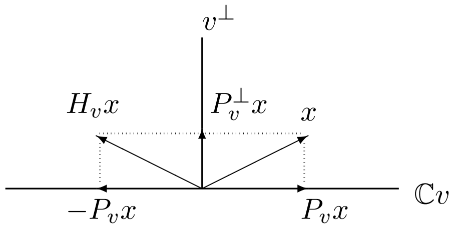

图 2.4.1

例 2.4.6 设 $v = e_3 \in \mathbb{R}^3$ 是标准向量, 则 $v^*v = 1$, 故相应的Householder 变换 H_v 变为 $H_v x = x - 2vv^*x$. 此时 H_v 正是关于 xoy 平面的反射.

例 2.4.7 (Givens 变换) 由矩阵 (2.4.1) 决定的旋转称为 Jacobi [21] 旋转或 (2 维) Givens 旋转. 一般地, 设 $c^2 + s^2 = 1, \theta = \arctan \frac{s}{c}$, 则将 n 阶实矩阵

$$G(i, j, \theta) = I_n - (1-c)(E_{ii} + E_{jj}) + s(E_{ij} - E_{ji}) \tag{2.4.5}$$

[19] Alston Scott Householder(1904-1993), 著名美国数值分析学家和生物数学家, 他在数值计算方面的开创性研究奠定了计算机科学的基础. 国际数值代数会议自 1996 年开始改称为 Householder 会议. 曾任美国数学会主席, 美国工业与应用数学会主席和 Association for Computing Machinery 主席.

[20] James Wallace Givens(1910-1993), 美国数学家, 计算机科学的先驱. 曾任美国工业与应用数学会主席.

[21] Carl Gustav Jacob Jacobi(1804-1851), 著名德国数学家, 数学上有著名的 Jacobian Conjecture (雅可比猜想), 见本书第五章.

称为 Givens 旋转矩阵 (感兴趣的读者可写出该矩阵的形状). 这是一个正交矩阵 (习题 44).

Householder 变换与 Givens 变换都可以用来将矩阵中某些特定的元素变为 0, 是矩阵快速计算中不可或缺的方法.

作为本节的结束, 我们讨论实对称矩阵对应的欧氏空间的线性变换 – 对称变换, 酉空间的情形见习题 46.

定义 2.4.2 设 σ 是欧氏空间 V 的线性变换. 如果对任意 $\alpha, \beta \in V$, 均有

$$(\sigma(\alpha), \beta) = (\alpha, \sigma(\beta)) \tag{2.4.6}$$

则称 σ 是一个**对称变换**.

位似变换都是对称变换. 反射变换也是对称变换, 这可以直接验证, 也可以由下面的定理得到.

定理 2.4.4 欧氏空间的线性变换 σ 是对称变换 $\Longleftrightarrow \sigma$ 在一组标准正交基下的矩阵是对称矩阵.

证 充分性. 设 σ 在一组标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵 A 是对称矩阵. 设 α, β 在该标准正交基下的坐标分别为 x, y. 则 $\sigma(\alpha) = (\alpha_1, \alpha_2, \cdots, \alpha_n)Ax$, $\sigma(\beta) = (\alpha_1, \alpha_2, \cdots, \alpha_n)Ay$, 因此

$$(\sigma(\alpha), \beta) = (Ax, y) = y^T Ax = (A^T y)^T x = (Ay)^T x = (x, Ay) = (\alpha, \sigma(\beta)).$$

必要性. 设 σ 是对称变换, 其在标准正交基 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 下的矩阵记为 $A = (a_{ij})$. 则 $\sigma(\alpha_j) = \sum_{i=1}^n a_{ij}\alpha_i$. 于是

$$a_{ij} = (\sigma(\alpha_j), \alpha_i) = (\alpha_i, \sigma(a_j)) = (\sigma(\alpha_i), \alpha_j) = a_{ji},$$

上式第二个等号用到欧氏空间内积的对称性, 而第三个等号用到 σ 是对称变换. □

例 2.4.8 由二维正交矩阵的分类可知, 如果一个线性变换既是对称变换又是正交变换, 则它必然是反射. 因此, 平面上非反射变换的对称变换不是等距变换, 从而对称变换和我们通常的理解相去甚远, 比如线性变换

$$(x, y)^T \mapsto (x + y, x)^T$$

是平面上的对称变换, 它将单位正方形 $0 \leq x, y \leq 1$ 变为如下的平行四边形

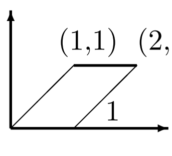

图 2.4.2

由对称变换可以自然导出另一类重要的线性变换, 称为一个线性变换的伴随, 其定义如下.

定义 2.4.3 设 σ, τ 是欧氏空间 V 的两个线性变换. 如果对任意 $\alpha, \beta \in V$, 均有

$$(\sigma(\alpha), \beta) = (\alpha, \tau(\beta)) \tag{2.4.7}$$

则称 τ 是 σ 的一个**伴随变换**.

容易证明 (见习题 45), 任何线性变换 σ 的伴随变换存在且唯一, 记为 σ^*. 如果 $\sigma = \sigma^*$, 则称 σ 是**自伴随的**或自伴的. 因此, 欧氏空间的对称变换就是自伴变换. 为了理解一般线性变换的伴随变换, 我们先考察一个例子.

例 2.4.9 设 σ 是平面上的线性变换

$$\sigma : (x, y)^T \mapsto (a_1x + b_1y, a_2x + b_2y)^T.$$

则容易验证 (见习题 45) σ 的伴随变换 σ^* 是

$$\sigma^* : (x, y)^T \mapsto (a_1x + a_2y, b_1x + b_2y)^T.$$

因此 σ^* 在标准基下的矩阵恰好是 σ 在标准基下的矩阵的转置! 因此, 伴随变换可看作是转置矩阵的几何意义.

一般地, 我们有下面的定理, 证明见习题 45.

定理 2.4.5 设 σ, τ 是欧氏空间 V 的两个线性变换, A, B 分别是 σ, τ 在某组标准正交基下的矩阵. 则

(1) $(\sigma^*)^* = \sigma$;
(2) $(\sigma + \tau)^* = \sigma^* + \tau^*$;
(3) $(\lambda\sigma)^* = \lambda\sigma^*, \forall \lambda \in \mathbb{R}$;
(4) $(\sigma\tau)^* = \tau^*\sigma^*$;
(5) $\tau = \sigma^* \iff B = A^T$.

自伴变换的最重要的例子是正交投影变换.

例 2.4.10 (正交分解) 由 定理 2.3.1, 任何向量 $x \in V$ 可以唯一分解为如下形式

$$x = P_Ux + P_U^{\perp}x,$$

其中 $P_Ux \in U$, $P_U^{\perp}x \in U^{\perp}$ 分别为 x 在 U 上的**投影向量** 和**正交投影向量**. 因此, 可将 P_U 与 $P_U^{\perp}$ 分别理解为内积空间 V 的向子空间 U 与 $U^{\perp}$ 的**正交投影变换**, 即

$$P_U : x \mapsto P_Ux$$

而

$$P_U^{\perp} : x \mapsto P_U^{\perp}x.$$

显然有 $P_U^{\perp} = P_{U^{\perp}}$ 以及 $P_U + P_U^{\perp} = I$.

一般, 如果 $U = \mathrm{Span}\{v\}$ 是 1 维子空间, 则将下标 U 改记为 v, 如 P_Ux 记为 P_vx 等.

定理 2.4.6 设 σ 是内积空间 V 的一个线性变换. 则 σ 是 V 向某子空间 U 上的正交投影变换 $\iff$ σ 是自伴的幂等变换.

证 必要性. 设 σ 是 V 向某子空间 U 上的正交投影变换, 即 $\sigma = P_U$. 则 σ 显然是幂等变换. 对任意 $x, y \in V$,

$$(P_U x, y) = (P_U x, P_U y + P_U^{\perp} y) = (P_U x, P_U y) = (x - P_U^{\perp} x, P_U y) = (x, P_U y),$$

因此 σ 是自伴的.

充分性. 设 σ 是自伴的幂等变换. 则由 例2.2.24 知 σ 是 V 向 $\mathrm{Im}(\sigma)$ 上的投影变换. 由于 σ 还是自伴的, 故 $\mathrm{Im}(\sigma)^{\perp} = \mathrm{Ker}(\sigma)$, 即 σ 还是正交投影变换. □

例 2.4.11 由矩阵 $A \in M_n(\mathbb{C})$ 定义的线性变换 $x \mapsto Ax$ 是正交投影变换 $\iff A^2 = A, A^* = A$. 此时, A 实际上是 $\mathbb{C}^n$ 向 A 的列空间 $R(A)$ 上的正交投影变换.

思考题

1. 平面 (即 $\mathbb{R}^2$) 上的非等距线性变换不能保持所有向量的长度, 但可否保持所有角度?
2. 空间 (即 $\mathbb{R}^3$) 中的非等距线性变换能否保持一些向量的长度? 能否将某个半径为 1 的圆还变为半径为 1 的圆? 特别, 空间中的幂零变换能否保持一些向量的长度? 幂等变换保持哪些向量的长度?
3. 平面上的反射变换能否由旋转实现? 反过来呢?
4. 按 例2.4.8, 对称变换并不保持图形的对称性, 如何为"对称"一词找一个恰当的几何解释?
5. 反对称矩阵对应的线性变换有何特点?
6. 对称变换是否在任何一组基下的矩阵均为对称矩阵? 在某组基下的矩阵为对称矩阵的线性变换是否一定是对称变换?
7. 定理2.4.5 中的"某组标准正交基"是否可以减弱为"某组基"?

第五节 * 张量积与商空间: 构造新线性空间

子空间的交与和都可以看成是由已知线性空间构造新线性空间的方法, 但这些方法局限在一个已知空间的内部. 本节我们研究如何由两个给定的 $\mathbb{F}$ 空间 U 与 V 构造新的线性空间. 最简单的办法是仿照子空间的直和, 构造所谓的"直和", 其定义如下.

考虑集合 U 与 V 的卡氏积(也称笛卡尔积)$U \times V = \{(u, v) \,|\, u \in U, v \in V\}$; 按分量定义 $U \times V$ 的加法与数乘运算, 即对 $(u, v), (u_1, v_1), (u_2, v_2) \in U \times V$ 以及任意数 λ 定义:

$$(u_1, v_1) + (u_2, v_2) = (u_1 + u_2, v_1 + v_2), \qquad \lambda(u, v) = (\lambda u, \lambda v).$$

则容易验证, $U \times V$ 在如上的加法和数乘下构成线性空间, 称为 U 与 V 的直和, 记为 $U \oplus V$, U 与 V 分别称为**直和项**.

关于直和有下述基本定理 (证明见习题47):

定理 2.5.1 (直和空间的基) 设 $\alpha_1, \alpha_2, \cdots, \alpha_n$(称为 α- 基) 与 $\beta_1, \beta_2, \cdots, \beta_m$(称为 β-基) 分别是 U 与 V 的一组基, 则向量组

$$(\alpha_i, 0), (0, \beta_j), 1 \le i \le n, 1 \le j \le m \tag{2.5.1}$$

构成 $U \oplus V$ 的一组基, 称为 $\alpha \oplus \beta$- 基. 特别地,

$$\dim (U \oplus V) = \dim U + \dim V. \tag{2.5.2}$$

(如果 U 与 V 还是内积空间, 则还可以按分量定义 $U \oplus V$ 的内积, 即定义

$$((u_1, v_1), (u_2, v_2)) = (u_1, u_2) + (v_1, v_2). \tag{2.5.3}$$

见习题 48.)

一般将子空间的直和称为“内直和”, 而将上面定义的直和称为“外直和”. 但这两种直和本质上是一样的. 这是因为“外直和” $U \oplus V$ 中的两个直和项 U 与 V 可以分别看成是线性空间 $W = U \oplus V$ 的子空间 $\{(u, v) \in W \mid u \in U, v = 0\}$ 与 $\{(u, v) \in W \mid v \in V, u = 0\}$. 例如实平面 $\mathbb{R}^2$ 既可以看成是由 x 轴和 y 轴的内直和产生的, 也可以看成是由两个数轴的外直和产生的, 因为这两个数轴可以分别看成是 x 轴和 y 轴.

常将外直和 $U \oplus V$ 中的向量 (u, v) 改记为 $u + v$, 这样外直和与内直和就完全统一了. 特别, $\alpha \oplus \beta$- 基就变成了 $\alpha \cup \beta$- 基. 以下我们将不区分外直和与内直和.

现考虑直和空间的线性变换. 设 $\sigma_i \in \mathrm{Hom}(U_i, V_i)$, $i = 1, 2$. 则可以定义 $U_1 \oplus U_2$ 到 $V_1 \oplus V_2$ 的映射 σ 如下 (其中 $u_i \in U_i$):

$$\sigma(u_1 + u_2) = \sigma_1(u_1) + \sigma_2(u_2) \tag{2.5.4}$$

容易验证 (习题 49), σ 是一个线性变换, 称为 σ_1 与 σ_2 的直和, 记为 $\sigma_1 \oplus \sigma_2$.

关于线性变换的直和有下面的定理:

定理 2.5.2 设线性变换 σ_1 与 σ_2 关于 α-α'- 基与 β-β'- 基的矩阵分别为 A 与 B, 则 $\sigma_1 \oplus \sigma_2$ 关于 $\alpha \oplus \beta$-$\alpha' \oplus \beta'$- 基的矩阵为 $A \oplus B$.

证 由于 $\sigma_1 \oplus \sigma_2$ 在 $\alpha \oplus \beta$- 基上的作用是按照分量进行的, 故

$$\sigma_1 \oplus \sigma_2(\alpha_i) = \sigma_1(\alpha_i), \quad \sigma_1 \oplus \sigma_2(\beta_j) = \sigma_2(\beta_j).$$

因此

$$\sigma_1 \oplus \sigma_2(\alpha \oplus \beta) = (\sigma_1(\alpha_1), \cdots, \sigma_1(\alpha_n), \sigma_2(\beta_1), \cdots, \sigma_2(\beta_m)) = (\alpha \oplus \beta)(A \oplus B).$$

□

例 2.5.1 设 $\sigma, \tau \in (\mathbb{R}^2)^*$ 分别为函数 $\sigma(x, y) = ax + by$ 与 $\tau(x, y) = cx + dy$. 则 $\sigma \oplus \tau \in \mathrm{Hom}(\mathbb{R}^2 \oplus \mathbb{R}^2, \mathbb{R} \oplus \mathbb{R})$ 为:

$$\sigma \oplus \tau(x, y, u, v) = (ax + by, cu + dv).$$

进一步, 直接计算可知 $\sigma \oplus \tau$ 关于标准基 - 标准基的矩阵为

$$C = \begin{pmatrix} a & b & 0 & 0 \\ 0 & 0 & c & d \end{pmatrix}$$

注意 σ 与 τ 关于标准基的矩阵分别为 $A = (a \quad b)$ 与 $B = (c \quad d)$, 所以确有 $C = A \oplus B$.

定理 2.5.2 可以推广到任意有限个线性空间的直和, 即有下面的定理 (证明见习题 50):

定理 2.5.3 设 $\sigma_i \in \mathrm{Hom}(U_i, V_i)$, $1 \le i \le n$. 再设 σ 关于 α_i-α_i'- 基的矩阵为 A_i. 则 $\sum_{i=1}^{n} \oplus \sigma_i \in \mathrm{Hom}(\sum_{i=1}^{n} \oplus U_i, \sum_{i=1}^{n} \oplus V_i)$ 关于 $\sum_{i=1}^{n} \oplus \alpha_i$-$\sum_{i=1}^{n} \oplus \alpha_i'$- 基的矩阵为 $\sum_{i=1}^{n} \oplus A_i$.

上面的过程实际上是利用较小的空间的线性变换构造较大的空间上的线性变换. 这个过程可以反过来考虑, 即讨论如何将较大空间的线性变换化为较小的空间的线性变换的直和. 设 $V = U \oplus W$, $\sigma \in \mathrm{End}V$. 如果 σ 分解为 $\sigma_1 \in \mathrm{End}U$ 与 $\sigma_2 \in \mathrm{End}W$ 的直和, 即 $\sigma = \sigma_1 \oplus \sigma_2$, 则 $\sigma(u+w) = \sigma_1(u) + \sigma_2(w)$. 特别, 对任意 $u \in U$, $w \in W$, 有

$$\sigma(u) = \sigma_1(u) \in U, \quad \sigma(w) = \sigma_2(w) \in W.$$

换句话说, $\sigma(U) \subseteq U$, $\sigma(W) \subseteq W$. 这样的子空间 U(与 W) 称为 σ 的**不变子空间**. 平凡子空间是任何线性变换的不变子空间. 另外, 核空间与像空间也是不变子空间.

定理 2.5.4 设 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$, $\sigma \in \mathrm{End}V$. 如果诸 V_i 均为 σ 的不变子空间, 则存在 $\sigma_i \in \mathrm{End}V_i$, $1 \le i \le s$, 使得

$$\sigma = \sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s.$$

证 由于诸 V_i 均为 σ 的不变子空间, 故可定义诸 V_i 的映射 σ_i 如下:

$$\sigma_i(x) = \sigma(x), \forall x \in V_i.$$

σ_i 显然是线性变换, 且对任意 $v \in V$, 设

$$v = v_1 + v_2 + \cdots + v_s, \quad v_i \in V_i,$$

有

$$\begin{aligned}
&\sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s(v) \\
&= \sigma_1(v_1) + \sigma_2(v_2) + \cdots + \sigma_s(v_s) \\
&= \sigma(v_1) + \sigma(v_2) + \cdots + \sigma(v_s) \\
&= \sigma(v_1 + v_2 + \cdots + v_s) = \sigma(v),
\end{aligned}$$

即 $\sigma = \sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s$. □

一般将上述诸 σ_i 称为线性变换 σ 在子空间 V_i 上的限制, 记为 $\sigma|_{V_i}$.

推论 2.5.1 设 $V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$, $\sigma \in \mathrm{End}V$. 如果诸 V_i 均为 σ 的不变子空间, 则存在 V 的基, 使得 σ 在该基下的矩阵是分块对角矩阵.

证 由 定理2.5.4, 可设 $\sigma = \sigma_1 \oplus \sigma_2 \oplus \cdots \oplus \sigma_s$. 设诸 σ_i 在 V_i 的 α_i- 基下的矩阵是 A_i, 则 σ 在 $\sum\limits_{i=1}^{s} \oplus \alpha_i$- 基下的矩阵为

$$\sum_{i=1}^{s} \oplus A_i.$$

□

例 2.5.2 设 $\sigma \in \mathrm{End}V$. 对任意非零向量 v, 投影变换 $x \mapsto P_v x$ 是幂等变换. 对任何幂等变换 σ 有

$$V = \mathrm{Ker}(\sigma) \oplus \mathrm{Im}(\sigma) \tag{2.5.5}$$

这是因为 $v = (v - \sigma(v)) + \sigma(v)$, 其中 $v - \sigma(v) \in \mathrm{Ker}(\sigma)$, $\sigma(v) \in \mathrm{Im}(\sigma)$ 且 $\mathrm{Ker}(\sigma) \cap \mathrm{Im}(\sigma) = 0$. 现选取 $\mathrm{Ker}(\sigma)$ 与 $\mathrm{Im}(\sigma)$ 的各一组基, 它们的并构成 V 的一组基, 在这组基下, σ 的矩阵为分块矩阵

$$0_{n-r} \oplus I_r,$$

其中 r 是 σ 的秩, n 是 V 的维数. 换句话说, σ 在 $\mathrm{Ker}(\sigma)$ 上的限制是 0 变换, 在 $\mathrm{Im}(\sigma)$ 上的限制是恒等变换.

由矩阵与线性变换的对应关系可知, 方阵 A 是幂等矩阵 $\iff$ 它所对应的线性变换 $x \mapsto Ax$ 是投影变换, 因此幂等矩阵也称为**投影矩阵**.

由于两个线性空间的直和中的运算是按分量进行的, 故两个直和项没有相互“融入”, 因此这个“新”线性空间本质上没有带来任何真正新的信息. 可以设想将 U 与 V 的基相互融合, 即以

$$(\alpha_i, \beta_j), 1 \le i \le n, 1 \le j \le m \tag{2.5.6}$$

为基构造新的线性空间, 称为 U 与 V 的**张量积**[22], 记为 $U \otimes V$. 一般, 将 $U \otimes V$ 的基 (2.5.6) 写成

$$\alpha_i \otimes \beta_j, 1 \le i \le n, 1 \le j \le m \tag{2.5.7}$$

从而 $U \otimes V$ 中的向量形如

$$\sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} \alpha_i \otimes \beta_j. \tag{2.5.8}$$

由上述定义立即可得下述的

定理 2.5.5 设 U 与 V 均是有限维线性空间, 则 $\dim (U \otimes V) = (\dim U)(\dim V)$.

所以“张量积”可以看成是线性空间的“乘法”. 显然, 如上定义的张量积对无限维线性空间依然有意义 (只需在式 (2.5.8) 中取任意有限和即可).

例 2.5.3 考察行向量空间 $U = \mathbb{F}^{1\times n}$ 与列向量空间 $V = \mathbb{F}^{m\times 1}$ 的张量积. 选取 U 与 V 的标准基 $e_1^T, \cdots, e_n^T$ 与 $e_1^{(m)}, \cdots, e_m^{(m)}$ (其中 $e_i^{(m)}$ 表示第 i 个 m 维标准向量). 则 $\mathbb{F}^{1\times n} \otimes \mathbb{F}^{m\times 1}$ 有一组标准基如下:

$$e_i^T \otimes e_j^{(m)},\ 1 \le i \le n,\ 1 \le j \le m. \tag{2.5.9}$$

如果将基向量 $e_i^T \otimes e_j^{(m)}$ 对应到基本矩阵 E_{ji}, 则可得

$$\mathbb{F}^{1\times n} \otimes \mathbb{F}^{m\times 1} \cong \mathbb{F}^{m\times n}. \tag{2.5.10}$$

例 2.5.4 (二元多项式的产生) 证明 $\mathbb{F}[x] \otimes \mathbb{F}[y] \cong \mathbb{F}[x, y]$.

证 分别取 $\mathbb{F}[x]$ 与 $\mathbb{F}[y]$ 的标准基

$$1, x, x^2, \cdots, x^n, \cdots$$

以及

$$1, y, y^2, \cdots, y^m, \cdots$$

[22]张量 (运算) 由意大利数学家 Gregorio Ricci-Curbastro (1853-1925) 提出的, 后经其学生 Tullio Levi-Civita (1873 − 1941) 发展而成为众多学科的重要理论和工具 (包括相对论).

则可得 $\mathbb{F}[x]\otimes\mathbb{F}[y]$ 的一组基如下

$$x^n\otimes y^m,\ m\,,n\geq 0.$$

将 $x^n\otimes y^m$ 对应到 x^ny^m 即给出需要的一个同构变换.

请注意, 本例并未涉及多项式的乘法结构, 有兴趣的读者可自行查阅相关材料. □

以下我们讨论张量积空间的线性变换. 设 $\sigma_i\in \mathrm{Hom}(U_i,\,V_i)$, $i=1,2$. 类似于直和的情形, 可以定义 $U_1\otimes U_2$ 到 $V_1\otimes V_2$ 的张量积映射 $\sigma_1\otimes\sigma_2$ 如下 (其中 $u_i\in U_i$):

$$\sigma_1\otimes\sigma_2(u_1\otimes u_2)\;=\;\sigma_1(u_1)\otimes\sigma_2(u_2) \tag{2.5.11}$$

容易验证 (习题51), $\sigma_1\otimes\sigma_2$ 是一个线性变换, 称为 σ_1 与 σ_2 的张量积. 需要注意, 张量积空间 $U_1\otimes U_2$ 中的元素的一般形式如式 (2.5.8), 因此公式 (2.5.11) 只定义了 $U_1\otimes U_2$ 中的一部分向量的像, 其余向量的像可由线性扩张得到, 见 定理2.2.1.

为了得到线性变换的张量积的矩阵表示, 我们先考察一个例子.

例 2.5.5 设 $\sigma\in(\mathbb{F}^2)^*$, $\tau\in\mathrm{Hom}(\mathbb{F},\mathbb{F}^2)$. 设 σ 与 τ 关于标准基 - 标准基的矩阵分别为 $A=(a\ \ b)$ 与 $B=(x\ \ y)^T$. 则 $\sigma\otimes\tau$ 将 $\mathbb{F}^2\otimes\mathbb{F}$ 的基向量分别映到

$$\begin{array}{l}\sigma\otimes\tau(e_1\otimes 1)=\sigma(e_1)\otimes\tau(1)=a\otimes(x\,,y)^T=ax(1\otimes e_1)+ay(1\otimes e_2)\\ \sigma\otimes\tau(e_2\otimes 1)=\sigma(e_2)\otimes\tau(1)=b\otimes(x\,,y)^T=bx(1\otimes e_1)+by(1\otimes e_2)\end{array}.$$

因此 $\sigma\otimes\tau$ 在 $\{e_1\otimes 1,\,e_2\otimes 1\}$-$\{1\otimes e_1,\,1\otimes e_2\}$- 基下的矩阵为 $\begin{pmatrix}ax & bx\\ ay & by\end{pmatrix}$. 该矩阵正是所谓矩阵 $A=(a\ \ b)$ 与 $B=(x\ \ y)^T$ 的张量积. 一般地, 有下述

定义 2.5.1 (矩阵的张量积) 设 $A=(a_{ij})$ 与 $B=(b_{st})$ 分别是 $m\times n$ 与 $p\times q$ 矩阵. A 与 B 的 (左) 张量积 (矩阵) 或Kronecker积, 记作 $A\otimes B$, 是指 $mp\times nq$ 矩阵

$$\begin{pmatrix}a_{11}B & a_{12}B & \cdots & a_{1n}B\\ a_{21}B & a_{22}B & \cdots & a_{2n}B\\ \cdots & \cdots & \cdots & \cdots\\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B\end{pmatrix}.$$

例 2.5.6 (1) $e_i^{(n)}\otimes(e_j^{(m)})^T=E_{ij}$;

(2) $I_n\otimes I_m=I_{mn}$;

(3) $\begin{pmatrix}a\\ b\end{pmatrix}\otimes\begin{pmatrix}x & y\end{pmatrix}=\begin{pmatrix}ax & ay\\ bx & by\end{pmatrix}=\begin{pmatrix}a\\ b\end{pmatrix}\begin{pmatrix}x & y\end{pmatrix}$.

(4) $\begin{pmatrix}a & b\end{pmatrix}\otimes\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}ax & bx\\ ay & by\end{pmatrix}\neq\begin{pmatrix}a & b\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=ax+by$.

例 2.5.7 设 A, B 均为2阶矩阵. 设 σ 是 $\mathbb{F}^{2\times 2}$ 上的一个线性变换, 其中

$$\sigma:\ X\mapsto\sigma(X)=AXB^T.$$

则直接计算可知, σ 关于 (按行顺序) 标准基 E_{11}, E_{12}, E_{21}, E_{22} 的矩阵是 $A\otimes B$, 关于 (按列顺序) 标准基 E_{11}, E_{21}, E_{12}, E_{22} 的矩阵是 $B\otimes A$.

此例提示我们矩阵的张量积与线性矩阵方程$AXB = C$ 的求解有密切关系.

以下列出矩阵张量积的一些性质 (证明见习题 52).

命题 2.5.1 (1)(结合律) $(A\otimes B)\otimes C = A\otimes(B\otimes C)$;
(2)(分配律 1) $(A\otimes B)(C\otimes D) = AC\otimes BD$;
(3)(分配律 2) $A\otimes(B\pm C) = A\otimes B\pm A\otimes C$;
(4)(保转置) $(A\otimes B)^T = A^T\otimes B^T$;
(5)(保可逆) $(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}$ (故 $A\otimes B$ 可逆 $\Longleftrightarrow A$ 与 B 均可逆);
(6) 设 A 与 B 分别为 m 阶与 n 阶矩阵, 则 $|A\otimes B| = |A|^m|B|^n$;
(7)(保秩与迹) $r(A\otimes B) = r(A)r(B)$; $\operatorname{tr}(A\otimes B) = \operatorname{tr}(A)\operatorname{tr}(B)$;
(8)(保特征值) 设 λ 与 μ 分别是 m 阶矩阵 A 与 n 阶矩阵 B 的特征值, x 与 y 是相应的特征向量, 则 $\lambda\mu$ 是 $A\otimes B$ 的特征值, $\lambda+\mu$ 是 $A\otimes I_m + I_n\otimes B$ 的特征值, $x\otimes y$ 是相应的特征向量.

实际上 例2.5.5 的结论具有一般性, 即关于线性变换的张量积与矩阵的张量积有下述关系 (证明见习题 53):

定理 2.5.6 设 $\sigma_i\in \operatorname{Hom}(U_i, V_i)$, $i = 1, 2$. 设 σ_i 关于 α_i-α_i' 基的矩阵分别为 A_i , $i = 1, 2$. 则 $\sigma_1\otimes\sigma_2$ 关于 $\alpha_1\otimes\alpha_2$-$\alpha_1'\otimes\alpha_2'$ 基的矩阵为 $A_1\otimes A_2$.

由公式 (2.5.7) 定义的线性空间的张量积依赖于两个因子空间的基的选择, 故常用另一种办法来定义张量积, 这就是**商空间**.

设 U 是线性空间 V 的子空间. 对任意 $\alpha\in V$, 将 V 的子集 $\{\beta\in V\,|\,\beta = \alpha+u, u\in U\}$ 称为向量 α 关于 U 的**陪集**, 记为 $\alpha+U$. 注意对两个不同的向量 $\alpha_1\neq\alpha_2$, 它们的陪集可能相同, 即可以有 $\alpha_1+U = \alpha_2+U$. 将 V 的关于 U 的所有不同的陪集构成的集合记为 V/U. 则该集合有一个自然的线性空间结构 (见习题 54), 其中的加法和数乘为

$$(\alpha_1+U)+(\alpha_2+U) = (\alpha_1+\alpha_2)+U,\ \ a(\alpha+U) = a\alpha+U. \tag{2.5.12}$$

线性空间 V/U 称为 V 关于 U 的商空间. 一般, 常将商空间 V/U 中的向量 $\alpha+U$ 记为 $\bar{\alpha}$.

例 2.5.8 考虑 $V=\mathbb{R}^2$ 的子空间 $U = x$- 轴. 对任意向量 $\alpha\in\mathbb{R}^2$, 陪集 $\alpha+U$ 就是 x- 轴平移 α 所得到的平行于 x- 轴的一条直线, 因此所有关于 x- 轴的陪集就是全体平行于 x- 轴的直线, 它们就是 V/U 的全体元素 (向量). 所以商空间 V/U 的几何意义就是将平行于 x- 轴的直线看成一个点, 或者说将它们压缩成 y- 轴上的点, 于是商空间 V/U 就是 y- 轴, 即 $V/U\cong y$-轴. 请读者自行建立该同构映射.

类似于 例2.5.8, 线性空间 V 的子空间 U 的陪集 $\alpha+U$ 实际上是将 U 沿向量 α “平移” 所得, 故几何学中常将点, 直线, 平面等称为**线性流形**. 因此, 线性空间的子空间实际上是过原点的线性流形.

商空间的本质是“分类”, 即将所谓“等价的元素”视为相同的. 比如, 在整数环 $\mathbb{Z}$ 中, 记全体偶数的集合为 $2\mathbb{Z}$. 如果 $a-b\in 2\mathbb{Z}$, 则称 a 与 b 等价. 因此所有偶数等价, 每个偶数 n 的陪集 $n+2\mathbb{Z} = 0+2\mathbb{Z} = \bar{0}$; 所有奇数等价, 每个奇数 m 的陪集 $m+2\mathbb{Z} = 1+2\mathbb{Z} = \bar{1}$. 因此

“商” $\mathbb{Z}/2\mathbb{Z}$ 实际上是一个 2 元集合 $\{\bar{0},\bar{1}\}$. 可以自然地定义该集合上的加法运算“+”与乘法运算“•”(即由 $\mathbb{Z}$ 的运算诱导) 如下:

$$
\begin{array}{c}
\bar{0}+\bar{0}=\bar{0},\quad \bar{0}+\bar{1}=\bar{1}+\bar{0}=\bar{1}\\
\bar{0}\bullet\bar{0}=\bar{0}\bullet\bar{1}=\bar{1}\bullet\bar{0}=\bar{0},\quad \bar{1}\bullet\bar{1}=\bar{1}
\end{array}
\tag{2.5.13}
$$

容易看出, 由公式 (2.5.13) 定义的 $\mathbb{Z}/2\mathbb{Z}$ 上的加法与乘法运算就是熟知的“偶数 + 偶数 = 偶数, 奇数 • 奇数 = 奇数”等的简洁表达. 在这样的加法与乘法下, 商 $\mathbb{Z}/2\mathbb{Z}$ 构成了应用中最为重要的一个有限域 – 二元域 $\mathbb{Z}_2$ 或 $\mathbb{F}_2$. 一般, 将 $\bar{0}$ 与 $\bar{1}$ 分别简记为 0 与 1, 即有 $\mathbb{F}_2=\{0,1\}$.

上述做法可以推广到任意素数 p, 即将 $2\mathbb{Z}$ 换成 $p\mathbb{Z}$, 即可得到有限域 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$.

为了更好地理解商空间的结构, 我们先列出以下简单性质, 证明见习题 55.

命题 2.5.2 (1) 对任意 $u\in U$, 有 $u+U=0+U=U$, 即 U 是商空间 V/U 的 0 向量;
(2) $v+U=\bar{0} \iff v\in U$;
(3) $\alpha+U=\beta+U \iff \alpha-\beta\in U$.

定理 2.5.7 设 $V=U\oplus W$, 则 $V/U\cong W$. 特别地, $\dim V/U=\dim V-\dim U$.

证 设 $\alpha_1,\cdots,\alpha_s$ 是 U 的一组基, $\alpha_{s+1},\cdots,\alpha_n$ 是 W 的一组基. 则

$$\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$$

是 V 的一组基. 断言

$$\alpha_{s+1}+U,\cdots,\alpha_n+U \tag{2.5.14}$$

是 V/U 的一组基. 一方面, 该组向量是线性无关的: 设

$$k_1(\alpha_{s+1}+U)+\cdots+k_{n-s}(\alpha_n+U)=0,$$

则

$$(k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n)+U=0,$$

即

$$k_1\alpha_{s+1}+\cdots+k_{n-s}\alpha_n\in U,$$

因此诸系数均为 0.

另一方面, 对任意 $\alpha+U\in V/U$, 设 $\alpha=x_1\alpha_1+\cdots+x_s\alpha_s+x_{s+1}\alpha_{s+1}+\cdots+x_n\alpha_n$, 则

$$
\begin{aligned}
\alpha+U &= (x_1\alpha_1+\cdots+x_s\alpha_s+x_{s+1}\alpha_{s+1}+\cdots+x_n\alpha_n)+U\\
&= [(x_1\alpha_1+\cdots+x_s\alpha_s)+U]+[(x_{s+1}\alpha_{s+1}+\cdots+x_n\alpha_n)+U]\\
&= (x_{s+1}\alpha_{s+1}+\cdots+x_n\alpha_n)+U\\
&= x_{s+1}(\alpha_{s+1}+U)+\cdots+x_n(\alpha_n+U),
\end{aligned}
$$

即 (2.5.14) 是 V/U 的一组基. □

由上面的证明可以知道商空间的基的结构, 即有下面的

推论 2.5.2 设 $\alpha_1,\cdots,\alpha_s$ 是 U 的一组基, $\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n$ 是 V 的一组基. 则

$$\alpha_{s+1}+U,\cdots,\alpha_n+U$$

是商空间 V/U 的一组基.

定义 2.5.2 设 V 是线性空间, $\sigma \in \mathrm{End}V$. 设 U 是 σ 的不变子空间. 定义 V/U 到自身的映射 $\bar{\sigma}$ 如下:

$$\bar{\sigma}:\ \alpha + U \mapsto \sigma(\alpha) + U. \tag{2.5.15}$$

则 $\bar{\sigma}$ 是线性变换 (习题 56), 称为 σ 的**诱导变换** 或**商变换**.

例 2.5.9 设 σ 是 $\mathbb{R}^3$ 的关于 xoy- 平面的投影变换. 则 σ 关于基 e_3, e_1, e_2 的矩阵为 $(0) \oplus I_2$. 设 U 是 z- 轴 (看成是 $\mathbb{R}^3$ 的子空间), 则 σ 诱导的商空间 V/U 的线性变换为 $\bar{\sigma}((x,y,z)^T + U) = (x,y,0)^T + U = (x,y,z)^T + U$, $\bar{\sigma}$ 关于基 $\bar{e}_1, \bar{e}_2$ 的矩阵为 I_2. σ 关于基 e_3, e_1+e_3, e_2+e_3 的矩阵为 $\begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. $\bar{\sigma}$ 关于基 $\overline{e_1+e_3}, \overline{e_2+e_3}$ 的矩阵仍然为 I_2(实际上 $\bar{\sigma}$ 是恒等变换).

以下考察诱导变换的矩阵表示.

定理 2.5.8 设 $A \in M_s(\mathbb{F}), B, C \in M_{n-s}(\mathbb{F})$. 设 V 是有限维线性空间, $\sigma \in \mathrm{End}V$. 设 σ 在 V 的两组基 $\alpha_1, \cdots, \alpha_s, \alpha_{s+1}, \cdots, \alpha_n$ 与 $\alpha_1, \cdots, \alpha_s, \beta_{s+1}, \cdots, \beta_n$ 下的矩阵分别为 $A \oplus B$ 与 $A \oplus C$. 则 $U = \mathrm{Span}\{\alpha_1, \cdots, \alpha_s\}$ 是 σ 的不变子空间, 且 σ 的诱导映射 $\bar{\sigma}$ 关于商空间 V/U 的两组基 $\bar{\alpha}_{s+1}, \cdots, \bar{\alpha}_n$ 与 $\bar{\beta}_{s+1}, \cdots, \bar{\beta}_n$ 下的矩阵分别为 B 与 C.

证 由 σ 在此两组基下的矩阵为分块对角矩阵可知, 不仅 $U = \mathrm{Span}\{\alpha_1, \cdots, \alpha_s\}$ 是 σ 的不变子空间, $\mathrm{Span}\{\alpha_{s+1}, \cdots, \alpha_n\}$ 与 $\mathrm{Span}\{\beta_{s+1}, \cdots, \beta_n\}$ 也均是 σ 的不变子空间. 由 推论 2.5.2 可知, $\bar{\alpha}_{s+1}, \cdots, \bar{\alpha}_n$ 与 $\bar{\beta}_{s+1}, \cdots, \bar{\beta}_n$ 都是 V/U 的基. 由诱导映射 $\bar{\sigma}$ 的定义可得

$$(\bar{\sigma}(\bar{\alpha}_{s+1}), \cdots, \bar{\sigma}(\bar{\alpha}_n)) = (\bar{\alpha}_1, \cdots, \bar{\alpha}_s, \bar{\alpha}_{s+1}, \cdots, \bar{\alpha}_n) \begin{pmatrix} 0 \\ B \end{pmatrix}.$$

由于 $\bar{\alpha}_1 = 0, \cdots, \bar{\alpha}_s = 0$, 故有

$$(\bar{\sigma}(\bar{\alpha}_{s+1}), \cdots, \bar{\sigma}(\bar{\alpha}_n)) = (\bar{\alpha}_{s+1}, \cdots, \bar{\alpha}_n)B,$$

即 $\bar{\sigma}$ 在基 $\bar{\alpha}_{s+1}, \cdots, \bar{\alpha}_n$ 下的矩阵为 B. 同理知 $\bar{\sigma}$ 在基 $\bar{\beta}_{s+1}, \cdots, \bar{\beta}_n$ 下的矩阵 C. $\square$

上面的证明实际上并不需要子空间 $\mathrm{Span}\{\alpha_{s+1}, \cdots, \alpha_n\}$ 与 $\mathrm{Span}\{\beta_{s+1}, \cdots, \beta_n\}$ 是 σ 的不变子空间, 其中的细节见习题 57.

由于线性变换在不同基下的矩阵彼此相似, 故由 定理 2.5.8 可得下述

推论 2.5.3 设 A, B, C 均是方阵, 则 $A \oplus B$ 与 $A \oplus C$ 相似 $\Longleftrightarrow$ B 与 C 相似.

现在我们可以利用商空间来重新定义线性空间的张量积了.

设 U 与 V 是两个线性空间, 将以集合 $U \times V$ 为基的无限维线性空间记为 $F(U \times V)$, 称为 $U \times V$ 上的**自由线性空间**, 即

$$F(U \times V) = \{\sum_{i=0}^{n} x_i(u_i, v_i) \mid u \in U,\ v \in V, x_i \in \mathbb{F}, n \in \mathbb{N}\} \tag{2.5.16}$$

记线性空间 $F(U\times V)$ 的由下列三种向量生成的子空间为 N:

$$(u_1+u_2,v)-(u_1,v)-(u_2,v)$$

$$(u,v_1+v_2)-(u,v_1)-(u,v_2)$$

$$x(u,v)-(xu,v),\quad x(u,v)-(u,xv).$$

定义 2.5.3 称商空间 $F(U\times V)/N$ 为 U 与 V 的张量积 (空间).

利用商空间定义的张量积空间 $F(U\times V)/N$ 不依赖于 U 与 V 的特定的基, 但其困难也是显而易见的, 即涉及到无限维空间 $F(U\times V)$, 而且子空间 N 往往也不易确定.

例 2.5.10 考察最简单的情形, 即取 $U=V=\mathbb{F}$. 固定 U 与 V 的各一组基 α 与 β. 则 定义2.5.3 中的子空间 N 由以下三种元素生成:

$$(x_1\alpha+x_2\alpha,y\beta)-(x_1\alpha,y\beta)-(x_2\alpha,y\beta)$$

$$(x\alpha,y_1\beta+y_2\beta)-(x\alpha,y_1\beta)-(x\alpha,y_2\beta)$$

$$z(x\alpha,y\beta)-(zx\alpha,y\beta),\quad z(x\alpha,y\beta)-(x\alpha,zy\beta)$$

因此在商空间 $F(U\otimes V)/N$ 中, 任意向量 $(x\alpha,y\beta)+N=xy(\alpha,\beta)+N$, 因此向量 $(\alpha,\beta)+N$ 是商空间 $F(U\otimes V)/N$ 的一组基. 从而商空间 $F(U\otimes V)/N$ 是1维线性空间. 实际上, 对应 $(\alpha,\beta)\mapsto\alpha\otimes\beta$ 给出商空间 $F(U\otimes V)/N$ 和张量积 $U\otimes V$ 之间的一个同构映射, 见习题61.

更一般地, 有下述 (证明见习题62)

定理 2.5.9 设 U 与 V 是两个线性空间, 则商空间 $F(U\otimes V)/N$ 与张量积 $U\otimes V$ 同构.

我们将在下一节介绍矩阵的张量积在解线性矩阵方程中的应用. 它的另一个应用是矩阵函数乘积的导数公式, 我们将在第五章中讨论该问题.

思考题

1. 设 $U_i,V_j,1\le i\le n,1\le j\le m$ 是线性空间. 描述 $\mathrm{Hom}(\sum\limits_{i=1}^{n}\oplus U_i,\ \sum\limits_{j=1}^{m}\oplus V_i)$ 中的元素的结构, 并以此给出分块矩阵的一个几何解释.

2. $\mathbb{Q}(\sqrt{2})$ 是有理数域上的2维线性空间. 它与 $\mathbb{Q}$ 及自身的张量积 (空间) 分别是什么?

3. 复数域 $\mathbb{C}$ 是实数域 $\mathbb{R}$ 上的2维线性空间. 商空间 $\mathbb{C}/\mathbb{R}$ 是什么?

4. 设 p 是素数, 有限域 $\mathbb{F}_p=\mathbb{Z}/p\mathbb{Z}$ 的加法与乘法结构如何? 能否建立 $\mathbb{F}_2$(或 $\mathbb{F}_p$) 上的线性空间 (线性变换) 理论?

5. 实多项式空间 $\mathbb{R}[x]$ 与复数域 $\mathbb{C}$ 均是 $\mathbb{R}$ 上的线性空间, 它们的张量积是什么?

6. 设 V 是线性空间, $\sigma\in\mathrm{End}V$ 是幂零 (幂等, 同构, 等) 变换. 设 U 是 σ 的不变子空间, 设 $\bar{\sigma}$ 是由 σ 诱导的 V/U 上的商变换, 问 $\bar{\sigma}$ 是否也是幂零 (幂等, 同构, 等) 变换?

第六节 应用: 拟合曲线, 移动通信, 滤波, 线性矩阵方程

本章简要介绍子空间, 线性变换, 张量积等理论与方法的一些应用.

一. 最小二乘拟合曲线

当观测的数据点明显不接近任何直线或平面时, 仍然用一条直线或平面等去拟合所有数据会产生较大的误差, 这时合适的方法是尝试用更高次的拟合曲线, 比如, 若数据点 $(x_1, y_1), \cdots, (x_n, y_n)$ 明显位于某条抛物线之上时, 就应该选择方程

$$y = a_0 + a_1 x + a_2 x^2$$

来逼近数据. 更一般地, 记上面的方程导出的方程组为

$$Y = X\alpha$$

其中 X 称为**设计矩阵**, α 称为参数 (系数) 向量, y 称为观测向量. 为了使误差向量(也称为余差向量)$Y - X\alpha$ 的长度最小, 相当于求出 $Y = X\alpha$ 的最小二乘解, 即方程组 $X^T Y = X^T X\alpha$ 的解.

例 2.6.1 试分别求观测数据

$$(-2, 12), (-1, 5), (0, 3), (1, 2), (2, 4)$$

的最小二乘二次拟合曲线与三次拟合曲线.(见习题 63)

二. 正交性与移动通信

例 2.6.2 设 $\mathbb{S} = \{\{x_k\} = (\cdots, x_{-2}, x_{-1}, x_0, x_1, x_2, \cdots) \,|\, x_i \in \mathbb{R}\}$ 表示双向无穷实数列的集合, 按分量自然定义 $\mathbb{S}$ 中的向量加法和数乘, 则 V 构成一个无限维实线性空间. 工程上常将 $\mathbb{S}$ 称为 (离散时间) 信号空间.

实践中常用的是信号空间 $\mathbb{S}$ 的一个子空间, 即平方可和子空间

$$l_2 = \{\{x_k\} \in \mathbb{S} \,|\, \sum_{k=-\infty}^{+\infty} x_k^2 < +\infty\}.$$

信号空间 $\mathbb{S}$ 的另几个常用的子空间是:
(1) 绝对收敛子空间

$$l_1 = \{\{x_k\} \in \mathbb{S} \,|\, x_k = 0, 若\ k \le 0\ 且\ \sum_{k=1}^{+\infty} |x_k| < +\infty\};$$

(2) 有限非零序列子空间

$$\{\{x_k\} \in \mathbb{S} \,|\, x_k\ 几乎处处为\ 0\,\}\ (即仅有有限个\ x_k\ 非\ 0\,)$$

(3) 有界序列子空间

$$l_\infty = \{\{x_k\} \in \mathbb{S} \,|\, x_k = 0, 若\ k \le 0,\ 且存在整数 C, |x_k| < C, \forall k\}.$$

平方可和子空间 l_2 是一个无限维内积空间, 其中的内积是自然定义的 (见习题 64), 即

$$(\{x_k\}, \{y_k\}\,) = \sum_{k=-\infty}^{+\infty} x_k y_k. \tag{2.6.1}$$

于是两个离散信号 $\{x_k\}$ 与 $\{y_k\}$ 正交 $\iff \sum\limits_{k=-\infty}^{+\infty} x_k y_k = 0.$

在实际应用中, 所有用户共享整个频率信道, 但不同用户被分配至不同的时区, 并且不同用户的通信信号在时域上没有任何重叠, 即若以 $\{x_k\}$ 与 $\{y_k\}$ 分别表示两个不同用户的通信信号, 则 $x_k y_k = 0, \forall k$ (此时由向量 $\{x_k\}$ 与 $\{y_k\}$ 生成的子空间不但相交为 0, 而且还是正交的), 因此 $\{x_k\}$ 与 $\{y_k\}$ 正交. 这就是时分多址通信的工作原理. 移动通信中的**频分多址**与**码分多址**的工作原理完全类似. 所以正交性对移动通信起着至为关键的作用.

三. 差分方程与线性滤波器

在数字信号处理中, 线性滤波器可以用如下的 n 阶线性差分方程来表示 (a_i 为常数且 $a_0 a_n \neq 0$)

$$y_k = a_n x_k + a_{n-1} x_{k+1} + \cdots + a_1 x_{k+n-1} + a_0 x_{k+n} \tag{2.6.2}$$

其中 $\{x_k\}$ 是输入信号而 $\{y_k\}$ 为输出信号. 如果 $\{y_k\}$ 是零序列, 则称该差分方程为齐次的, 此时的解对应于被过滤掉的信号 (并被换为零信号).

例 2.6.3 对 2 阶滤波器 $y_k = \frac{1}{2\sqrt{2}} x_k + \frac{1}{2} x_{k+1} + \frac{1}{2\sqrt{2}} x_{k+2}$ 输入信号 (来自于连续信号 $x = \cos\frac{\pi t}{4}$ 的整数抽样)

$$\{x_k\} = \{\cdots, x_0 = 1, \frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 1, -\frac{1}{\sqrt{2}}, 0, \cdots\}$$

可得输出信号 (具体计算见习题 65)

$$\{y_k\} = \{\cdots, y_0 = \frac{1}{\sqrt{2}}, y_1 = 1, \frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, -1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 1, -\frac{1}{\sqrt{2}}, 0, \cdots\}.$$

如果再输入一个新的更高的频率信号 $\{u_k\}$(来自于连续信号 $x = \cos\frac{3\pi t}{4}$ 的整数抽样), 则输入信号为 (具体计算见习题 65)

$$\{u_k\} = \{\cdots, u_0 = 1, -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, -1, \frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, 1, -\frac{1}{\sqrt{2}}, 0, \cdots\}.$$

此时滤波器产生的输出序列是零序列, 因此该滤波器将高频信号过滤掉了 (故称为低通滤波器).

命题 2.6.1 n 阶齐次线性差分方程的解集是一个 n 维线性空间. 因此, 被线性滤波器过滤掉的所有信号构成一个线性空间.

研究 n 阶齐次线性差分方程的一种方法是降阶法, 即化为等价的一阶差分方程, 而每个一阶差分方程可写成如下形式:

$$y_{k+1} = A y_k, \forall k \tag{2.6.3}$$

其中 $y_k \in \mathbb{R}^n$, $A \in M_n(\mathbb{R})$. 比如差分方程 (注意 $a_0 = 1$)

$$a_n x_k + a_{n-1} x_{k+1} + \cdots + a_1 x_{k+n-1} + x_{k+n} = 0, \forall k$$

可以化为 $y_{k+1} = A y_k\ (\forall k)$, 其中 $y_k = (x_k, x_{k+1}, \cdots, x_{k+n-1})^T$, 而

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{pmatrix}.$$

注意, 矩阵 A 正是多项式 $f(x)=a_n+a_{n-1}x+\cdots+a_1x^{n-1}+x^n$ 的友矩阵.

如果设 $x_k=0,k<0$, 则由迭代法可以将 (2.6.3) 化为

$$y_k=A^ky_0 \tag{2.6.4}$$

其中 $y_0=(x_0,x_1,\cdots,x_{n-1})^T$. 于是矩阵 A 的高次幂的计算成为解决问题的关键. 我们将在下一章继续讨论该问题.

四. 线性变换与度量

在线性代数课程中, 我们已经知道2阶与3阶行列式的绝对值的几何意义分别是 (平面或立体) 图形的面积与体积. 由于行列式可以看成是矩阵的一个数字特征, 所以研究图形在线性变换下的度量变化是有意义的. 确切地说, 设 S 是欧氏空间 V(此处为 $\mathbb{R}^2$ 或 $\mathbb{R}^3$, 当然可以更一般地研究高维欧氏空间中的类似度量) 中的图形, σ 是 V 的线性变换, 问 $\sigma(S)=\{y\in V\mid y=\sigma(x),\exists x\in S\}$ 与 S 的面积或体积的关系. 下面的定理回答了此问题 (证明见习题67).

定理 2.6.1 记 V 中图形 S 的面积或体积为 $vol(S)$. 则 $vol(\sigma(S))=|\sigma|vol(S)$.

例 2.6.4 不使用微积分, 求椭圆 $S=\{(x,y)\mid \frac{x^2}{a^2}+\frac{y^2}{b^2}\le 1\}$ 的面积.

这里的本质是任何椭圆一定是单位圆在线性变换下的像. 实际上本题中的椭圆是单位圆盘在 $\alpha\mapsto A\alpha$ 下的像, 其中 $A=\begin{pmatrix}a&0\\0&b\end{pmatrix}$. 因此由 定理2.6.1 知 S 的面积为 $|A|\pi=\pi ab$.

例 2.6.5 ($|AB|=|A||B|$ 的几何意义) 设 $A,B\in M_n(\mathbb{F})$, 有界图形 $G\subset\mathbb{F}^n$ 的体积为 $V\neq 0$. 考察 $\mathbb{F}^n$ 到自身的两个线性变换 $\sigma: x\mapsto Ax$ 与 $\tau: x\mapsto Bx$ 及其复合 $\sigma\tau: x\mapsto (AB)x$. 由 定理2.6.1 知, $\tau(G)$ 的体积为 $|B|V$, $\sigma(\tau(G))$ 的体积为 $|A||B|V$, 而 $\sigma\tau(G)$ 的体积为 $|AB|V$. 由于 $\sigma(\tau(G))=\sigma\tau(G)$, 故 $|AB|V=|A||B|V$. 所以

$$|AB|=|A||B|.$$

五. 张量积与线性矩阵方程

利用矩阵的张量积可以求线性矩阵方程

$$A_1XB_1+A_2XB_2+\cdots+A_sXB_s=C \tag{2.6.5}$$

的解, 其中 $A_i\in M_m(\mathbb{C}),B_i\in M_n(\mathbb{C}),C\in\mathbb{C}^{m\times n}$ 是已知矩阵, $X\in\mathbb{C}^{m\times n}$ 为未知矩阵.

首先考察方程 (2.6.5) 的左端只有一项的情形. 此时, 矩阵方程 $AXB^T=C$ 可以看成是求向量 C 在 σ 下的原像, 其中 $\sigma: X\mapsto AXB^T$ 是矩阵空间 $\mathbb{C}^{m\times n}$ 的线性变换. 我们在 例2.5.7 的讨论中已经知道, σ 在 (按列顺序) 标准基 E_{ij} 下的矩阵恰好是 $B\otimes A$, 因此若记 X 在 (按列顺序) 标准基下的坐标为 $\mathbf{x}$, 则 σ 可以改写为

$$\sigma:(E_{11},\cdots,E_{mn})\mathbf{x}\mapsto(E_{11},\cdots,E_{mn})(B\otimes A)\mathbf{x} \tag{2.6.6}$$

于是矩阵方程 $AXB^T=C$ 不过是下面的线性方程组而已

$$(B\otimes A)\mathbf{x}=\mathbf{c} \tag{2.6.7}$$

其中 $\mathbf{c}$ 是矩阵 C 在 (按列顺序) 标准基下的坐标.

容易看出, 一个 $m\times n$ 矩阵在 (按列顺序) 标准基下的坐标恰好是将其元素按列排成的 mn 维列向量, 因此有下面的

定义 2.6.1 设 $A=(a_{ij})_{m\times n}$, 将 A 的各列依次竖排得到 mn 维列向量, 称为矩阵 A 的**列展开**, 记为 $\mathrm{vec}(A)$, 即

$$\mathrm{vec}(A)=(a_{11},a_{21},...,a_{m1},a_{12},a_{22},...,a_{m2},......,a_{1n},a_{2n},...,a_{mn})^T.$$

类似地, 矩阵 A 的**行展开**为:

$$\mathrm{rvec}(A)=(a_{11},a_{12},...,a_{1n},a_{21},a_{22},...,a_{2n},......,a_{m1},a_{m2},...,a_{mn}).$$

显然, vec 可以看成是线性空间 $\mathbb{F}^{m\times n}$ 到线性空间 $\mathbb{F}^{mn}$ 的同构线性变换, 称为 (矩阵的) **向量化变换**. 由上述定义, 得 $\mathrm{rvec}(A^T)=(\mathrm{vec}(A))^T$, $\mathrm{vec}(A^T)=\mathrm{rvec}(A)^T$.

下面的定理列出了矩阵的行 (列) 展开的一些性质, 证明见习题 70.

定理 2.6.2 (1) $\mathrm{rvec}(ABC)=\mathrm{rvec}(B)(A^T\otimes C)$, $\mathrm{vec}(ABC)=(C^T\otimes A)\mathrm{vec}(B)$. 特别地,
$\mathrm{vec}(A_{m\times m}B_{m\times n})=(I_n\otimes A)\mathrm{vec}(B),\quad \mathrm{vec}(B_{m\times n}C_{n\times n})=(C^T\otimes I_m)\mathrm{vec}(B)$;
(2) $\mathrm{vec}(xA+yB)=x\mathrm{vec}(A)+y\mathrm{vec}(B)$, $x,y\in\mathbb{C}$;
(3) $\mathrm{vec}(AC+CB)=(I_n\otimes A+B^T\otimes I_m)\mathrm{vec}(C)$.

现在可以给出线性矩阵方程 (2.6.5) 的解了.

定理 2.6.3 矩阵 $X\in\mathbb{C}^{m\times n}$ 是方程 (2.6.5) 的解 $\iff$ $\mathrm{vec}(X)$ 是方程 $G\mathbf{x}=\mathbf{c}$ 的解, 其中 $\mathbf{x}=\mathrm{vec}(X)$, $\mathbf{c}=\mathrm{vec}(C)$, $G=\sum_{i=1}^s(B_i^T\otimes A_i)$.

证 只需将方程 (2.6.5) 两端按列展开, 即可得
$\mathbf{c}=\mathrm{vec}(C)=\sum_{i=1}^s\mathrm{vec}(A_iXB_i)=\sum_{i=1}^s(B_i^T\otimes A_i)\mathrm{vec}(X)=G\,\mathrm{vec}(X)=G\mathbf{x}.$ □

推论 2.6.1 矩阵方程 (2.6.5) 有解 $\iff$ $r([G,\mathbf{c}])=r(G)$; 有唯一解 $\iff$ G 是可逆矩阵.

例 2.6.6 求 **Lyapunov**[23] **方程** $AX+XB=C$ 的解, 其中

$$A=\begin{pmatrix}1&1\\0&1\end{pmatrix},\quad B=\begin{pmatrix}2&1\\0&2\end{pmatrix},\quad C=\begin{pmatrix}-2&3\\3&4\end{pmatrix}.$$

解

$$G=I_2\otimes A+B^T\otimes I_2=\begin{pmatrix}3&1&0&0\\0&3&0&0\\1&0&3&1\\0&1&0&3\end{pmatrix},$$

$\mathbf{c}=\mathrm{vec}(C)=(-2,3,3,4)^T$. 解方程 $G\mathbf{x}=\mathbf{c}$ 得

$$\mathbf{x}=(-1,1,1,1)^T.$$

[23] Aleksandr Mikhailovich Lyapunov(1857-1918), 著名俄罗斯数学家, 力学家, 物理学家, 稳定性理论与动力系统的奠基人. 于其妻病逝当天开枪自杀.

因此原矩阵方程的解为

$$X = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}.$$

习 题 二

1. 证明 定理2.1.1.

2. 证明 命题2.1.1.

3. (1) 证明维数定理 (定理2.1.2);
(2) 设 V 是有限维线性空间. 证明并解释下面的维数公式:

$$\dim V = \max\{m \mid 0 = V_0 \subset V_1 \subset \cdots \subset V_{m-1} \subset V_m = V,\ V_i \text{ 是 } V_{i+1} \text{ 的真子空间}\}$$

4. 证明 定理2.1.4(多子空间直和的判定).

5. 设

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{pmatrix},$$

求 A 的四个相关子空间.

6. 设 V 是线性空间, $W_1, W_2, \cdots, W_s$ 是 V 的真子空间. 证明 $W_1 \cup W_2 \cup \cdots \cup W_s \neq V$.(提示: 利用 Vandermonde 行列式或归纳法.)

7. 设 V 是所有 n 阶实数矩阵按矩阵的加法和数乘作成的实线性空间, U 是 V 中所有迹为零的矩阵的集合. 证明 U 是 V 的子空间, 并求 U 的维数和一个补空间.

8. 设 V 是所有次数小于 n 的实系数多项式组成的实线性空间, $U = \{f(x) \in V \mid f(1) = 0\}$. 证明 U 是 V 的子空间, 并求 V 的一个补空间.

9. 设 $U = [(1,2,3,6)^T, (4,-1,3,6)^T, (5,1,6,12)^T]$, $W = [(1,-1,1,1)^T, (2,-1,4,5)^T]$ 是 $\mathbb{R}^4$ 的两个子空间,
(1) 求 $U \cap W$ 的基;
(2) 扩充 $U \cap W$ 的基, 使其成为 U 的基;
(3) 扩充 $U \cap W$ 的基, 使其成为 W 的基;
(4) 求 $U + W$ 的基.

10. 设 $U = \{(x,y,z,w) \mid x+y+z+w=0\}$, $W = \{(x,y,z,w) \mid x-y+z-w=0\}$. 求 $U \cap W$, $U + W$ 的维数与基.

11. 设 A, B 分别是 $n\times m$, $m\times p$ 矩阵, V 是齐次线性方程组 $xAB = 0$ 的解空间. 证明 $U = \{y = xA \mid x \in V\}$ 是 $\mathbb{F}^n$ 的子空间, 并求 U 的维数.

12. 设 A 是 n 阶方阵. 证明
(1) A 可以唯一地表示成一个对称矩阵和一个反对称矩阵的和. 试用子空间的直和分解理论解释这一结果;
(2) A 可以唯一地表示成一个 Hermite 矩阵和一个反 Hermite 矩阵的和. 试用子空间的直和分解理论解释这一结果;
(3) 解释定义域为 $\mathbb{R}$ 的任意实函数可以唯一地表示成一个偶函数与一个奇函数的和;
(4) 请举一个类似于上面 (1)-(3) 的例子并解释之.

13. 证明数域 $\mathbb{F}$ 上的一元多项式的欧几里德带余除法: 设 $f(x), g(x)$ 是任意两个多项式, 其中 $g(x) \neq 0$, 则存在唯一一对多项式 $q(x)$ 与 $r(x)$ 使得

$$f(x) = g(x)q(x) + r(x)$$

其中 $r(x) = 0$ 或 $\partial r(x) < \partial g(x)$. 试用线性空间的理论解释这一结果.

14. (1) 设 f 是定义在实数域上的加性函数. 证明: 如果 f 是连续的, 则它一定是齐次的, 从而是线性变换;
(2) 试将 (1) 中的结论推广到一般情形.

15. 若 $\sigma(\alpha_1), \sigma(\alpha_2), \cdots, \sigma(\alpha_s)$ 线性相关, 证明或否定 $\alpha_1, \alpha_2, \cdots, \alpha_s$ 也线性相关.

16.(1) 证明 定理 2.2.2;

(2) 设 $\sigma \in \mathrm{Hom}(U, V)$ 是可逆线性变换, 证明其逆唯一, 且若 $\tau = \sigma^{-1}$, 则 $\sigma = \tau^{-1}$;
(3) 计算 2 维实线性空间 $\mathbb{C}$ 的所有自同构.

17. 设 $\sigma \in \mathrm{Hom}(U, V)$.
(1) 证明 σ 的核 $\mathrm{Ker}\sigma$ 与像 $\mathrm{Im}\sigma$ 分别是 U 与 V 的子空间;
(2) 证明商空间 $U/\mathrm{Ker}\sigma$ 与 $\mathrm{Im}\sigma$ 同构并建立相应的一个同构映射.

18. 设 $\sigma \in \mathrm{End}V$ 在 V 的某组基下的矩阵为 A. 证明或否定: 存在 $\tau \in \mathrm{End}V, \tau \neq \sigma$, 使得 τ 在另一组基下的矩阵也是 A.

19. 证明线性空间的同构定理即 定理 2.2.5.

20. 第一章习题 22 与 23 中的线性空间各是几维的? 试分别建立它们与某 $\mathbb{R}^n$ 之间的同构变换.

21. 设 U 与 V 均是有限维线性空间, 证明 $\dim_{\mathbb{F}} \mathrm{Hom}(U, V) = (\dim_{\mathbb{F}} U)(\dim_{\mathbb{F}} V)$.

22. 证明 定理 2.2.6.

23. 证明线性变换基本定理 (定理 2.2.7).

24. 证明 例 2.2.28 中的映射 ϕ 是 V 与 V^{**} 的同构变换并求其逆变换.

25. 分别求导数运算 $\partial: f(x) \mapsto f'(x)$ 在标准基 $1, x, x^2, \cdots, x^{n-1}$ 与基 $1, (x-a), (x-a)^2, \cdots, (x-a)^{n-1}$ 下的矩阵. 问 ∂ 的行列式与迹是多少? 解释之.

26. 设 V 是数域 $\mathbb{F}$ 上的 n 阶矩阵全体, σ 是将 V 中任意元素的严格下三角部分变为 0 的映射. 判断 σ 是否为 V 的线性变换. 若是, 求其核与像; 并任选 V 的一组基, 求 σ 在该组基下的矩阵.

27. (1) 求 例 2.2.22 中的幂零变换 τ 的幂零指数及其在标准基下的矩阵;
(2) 设 $\sigma, \tau \in \mathrm{End}V$ 分别是线性空间 V 的同构变换和幂零变换, 证明 $\sigma + \tau$ 是 V 的同构变换;
(3) 设 A, D 是可逆矩阵, B, C 是幂零矩阵, 证明分块矩阵 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 可逆.

28. 设 V 是数域 $\mathbb{F}$ 上的 n 阶矩阵全体, A 是 V 中一个固定元素, P 是 V 中一个固定的可逆矩阵, σ 是左乘 A 的映射, τ 是左乘 P 逆右乘 P 的映射. 判断 σ 与 τ 是否为 V 的线性变换. 若是, 求其核与像. 并任选 V 的一组基, 计算 σ 与 τ 在该组基下的矩阵.

29. 设 $V = \mathbb{R}^3$, $\sigma(x, y, z) = (x + 2y - z, y + z, x + y - 2z)$. 求
(1) σ 的核与像空间的基与维数;
(2) σ 的行列式与迹.

30. 设 V 是 n 维内积空间, U 是 V 的子空间. 令 $W = \{\alpha \in V \mid (\alpha, \beta) = 0, \forall \beta \in U\}$. 证明 W 是 V 的子空间且 $V = U \oplus W$.

31. 证明 定理 2.3.2(2).

32. 设 $V = \mathbb{R}[x]_n$, 其上的内积为

$$(f(x), g(x)) = \int_0^1 f(x)g(x)\,\mathrm{d}\,x.$$

设 $U = \{f(x) \in V \mid f(0) = 0\}$.
(1) 证明 U 是 V 的一个 $n-1$ 维子空间, 并求 U 的一组基;
(2) 当 $n = 3$ 时, 求 U 的正交补 $U^{\perp}$.

33. 在欧氏空间 $\mathbb{R}^n$ 中求一个超平面 W, 使得向量 $e_1 + e_2$ 在 W 中的最佳近似向量为 e_2.

34. 证明: 函数 $f(x)$ 的 Fourier 级数中的系数 $a_n, b_n\,(n > 0)$ 恰好是 $f(x)$ 与诸基向量 $\cos nx, \sin nx$ 的内积.

35. 证明 命题2.3.1.

36. 试任意构造维数大于5的一个线性空间 V 以及 V 的一个线性映射 σ, 使得 σ 的核的维数等于5. 进一步, 试将 V 改造成内积空间, 求 $\mathrm{Im}\sigma$ 的正交补空间. 再构造一个线性变换 τ, 使得 $\mathrm{Ker}\tau = \mathrm{Im}\sigma$, $\mathrm{Im}\tau = \mathrm{Ker}\sigma$.

37. 设 α_0 是欧氏空间 V 中的单位向量, $\sigma(\alpha) = \alpha - 2(\alpha, \alpha_0)\alpha_0$, $\alpha \in V$. 证明
(1) σ 是线性变换;
(2) σ 是正交变换.

38. 证明: 欧氏空间 V 的线性变换 σ 是反对称变换 (即 $(\sigma(\alpha), \beta) = -(\alpha, \sigma(\beta))$) $\Longleftrightarrow$ σ 在 V 的标准正交基下的矩阵是反对称矩阵.

39. 设 σ 是实平面 $\mathbb{R}^2$ 上的线性变换, 其关于标准基的矩阵为

$$P = \begin{pmatrix} c & s \\ s & -c \end{pmatrix}$$

其中 $c^2 + s^2 = 1$. 证明 σ 是反射变换, 并计算其对称轴.

40. 证明 例2.4.3 关于三维正交矩阵的结论. (提示: 利用三维正交矩阵有1个或3个实特征值并结合二维正交矩阵的分类.)

41. 设 $\sigma \in \mathrm{End}\mathbb{R}^2$. 记单位正方形 $S = \{(x, y) : 0 \leq x, y \leq 1\}$ 在 σ 下的图形为 $G = \sigma(S) = \{\sigma(x, y) : (x, y) \in S\}$. 回答下列问题:
(1) 列出 G 所有可能的形状;
(2) 如果 G 仍为正方形, σ 应满足什么条件?
(3) 如果 σ 可逆, 则 G 是什么形状?

42. (1) 设 $\sigma \in \mathrm{End}\mathbb{R}^2$. 设 C 是一个二次曲线 (即抛物线, 椭圆或双曲线). 计算 $\sigma(C)$ 所有可能的形状 (可设 C 的方程均为标准方程);
(2) 设 P 是一个平面 n 次代数曲线 (即 C 的方程是 n 次多项式), 计算 $\sigma(Q)$ 所有可能的形状;
(3) 分别对指数函数, 对数函数, 三角函数研究其曲线在 σ 下的图像;
(4) 设 $\sigma \in \mathrm{End}\mathbb{R}^3$. 设 Q 是一个二次曲面. 计算 $\sigma(Q)$ 所有可能的形状 (可设 Q 的方程均为标准方程).

43. 证明Householder变换 H_v 是关于超平面 $v^{\perp}$ 的反射, 从而是正交变换. 试画出三维Householder变换的示意图.

44. 证明Givens旋转矩阵 G 是正交矩阵. 对任意向量 $x = (x_1, x_2, \cdots, x_n)^T$, 计算 Gx 的各个分量. 设 x 是单位向量, 讨论如何重复使用若干Givens旋转矩阵将 x 变为标准向量 e_1.

45. (1) 证明任何线性变换的伴随变换是唯一存在的;
(2) 验证 例2.4.9 的结论, 并证明 定理2.4.5;
(3) 证明正交投影变换是自伴变换.

46. (1) 如何在酉空间中定义Hermite矩阵对应的Hermite变换? 导出Hermite变换的一个判断准则;
(2) 在酉空间中定义伴随变换与自伴变换, 并导出伴随变换的基本性质 (参考 定理2.4.5).

47. 证明 定理2.5.1.

48. 证明由公式 (2.5.3) 定义的函数是 $U \oplus V$ 上的内积.

49. 证明公式 (2.5.4) 中的映射是线性变换.

50. 证明 定理2.5.3.

51. 证明公式 (2.5.11) 定义了一个线性变换.

52. 证明 命题2.5.1.

53. 证明 定理2.5.6.

54. 证明集合 V/U 按公式 (2.5.12) 定义的加法和数乘作成线性空间.

55. 证明 命题2.5.2.

56. 验证公式 (2.5.15) 定义的诱导映射 $\bar{\sigma}$ 是线性变换. 条件"U 是 σ 的不变子空间"可以去掉吗?

57. 设 $A\in M_s(\mathbb{F}), B, C\in M_{n-s}(\mathbb{F})$. 设 V 是有限维线性空间, $\sigma\in \mathrm{End}V$. 设 σ 在 V 的两组基

$$\alpha_1,\cdots,\alpha_s,\alpha_{s+1},\cdots,\alpha_n \text{ 与 } \alpha_1,\cdots,\alpha_s,\beta_{s+1},\cdots,\beta_n$$

下的矩阵分别为 $\begin{pmatrix} A & X\\ 0 & B\end{pmatrix}$ 与 $\begin{pmatrix} A & Y\\ 0 & C\end{pmatrix}$. 证明 $U=\mathrm{Span}\{\alpha_1,\cdots,\alpha_s\}$ 是 σ 的不变子空间, 且 σ 的诱导映射 $\bar{\sigma}$ 关于商空间 V/U 的两组基 $\bar{\alpha}_{s+1},\cdots,\bar{\alpha}_n$ 与 $\bar{\beta}_{s+1},\cdots,\bar{\beta}_n$ 下的矩阵分别为 B 与 C. 并据此推广 推论2.5.3.

58. 推广 例2.5.7, 即设 A, B 均为 n 阶矩阵. 设 σ 是 $\mathbb{F}^{n\times n}$ 上的一个线性变换, 其中

$$\sigma:\ X\mapsto \sigma(X)=AXB^T.$$

证明 σ 关于标准基 $E_{11},\cdots,E_{1n},E_{21},\cdots,E_{2n},E_{n1},\cdots,E_{nn}$ 的矩阵是 $B\otimes A$.

59. (复数,位似与旋转矩阵) 设 σ 是 $\mathbb{C}$ 到自身的线性变换, 其定义为

$$\sigma:\begin{pmatrix}x\\y\end{pmatrix}\mapsto A\begin{pmatrix}x\\y\end{pmatrix},$$

其中

$$A=\begin{pmatrix} a & b\\ -b & a\end{pmatrix}.$$

将 $(x,y)^T$ 记为普通复数 $x+y\mathrm{i}$, 证明 $\sigma((x,y)^T)=(a+b\mathrm{i})(x+y\mathrm{i})$. 请解释之.

60. 设 $V=M_2(\mathbb{C}), U=M_2(\mathbb{R})$. 记 $\mathrm{i}U=\{A\in V\,|\,A=\mathrm{i}X,\exists X\in U\}$. 研究 (实) 线性空间 $U\oplus(\mathrm{i}U)$, V/U 以及 $(V/U)\otimes\mathbb{C}$ 的结构.

61. 验证 例2.5.10.

62. 证明 定理2.5.9.

63. 建立 例2.6.1 的线性方程组, 并利用适当软件计算之.

64. 验证公式 (2.6.1) 是绝对平方收敛子空间 C 的一个内积. 它是信号空间 $\mathbb{S}$ 的内积吗?

65. 验证 例2.6.3 的滤波器确实将高频信号 $\{u_k\}$ 过滤掉. 研究被该滤波器过滤掉的所有信号具有什么特征?

66. 求差分方程 $y_{k+3}-2y_{k+2}-5y_{k+1}+6y_k=0$(对 $\forall k$ 成立) 的解空间的一组基, 并解释之.

67. 证明 定理2.6.1.

68. 设 σ 同29题, $S=\{(x,y,z)^T\in\mathbb{R}^3\,|\,0\le x,y,z\le 1\}$ 为标准单位立方体. 计算 $\sigma(S)$ 的体积.

69. 利用 定理2.6.1 求椭球的体积公式.

70. 证明 定理2.6.2.

71. 求方程 $AX+XB=C$ 的解, 其中

$$A=\begin{pmatrix} 1 & -1 & 3\\ -2 & 2 & 2\\ -1 & -1 & 5\end{pmatrix}, B=\begin{pmatrix} 1 & -1 & 3 & 0\\ -2 & 2 & 2 & 0\\ -1 & -1 & 5 & 0\\ 0 & 1 & -3 & 1\end{pmatrix}, C=\begin{pmatrix} 2 & 3 & 1 & 0\\ 2 & 0 & 1 & 1\\ 0 & -1 & -2 & 3\end{pmatrix}.$$

72. 设 $\sigma: X\mapsto AX+XB$ 是 $M_n(\mathbb{C})$ 的线性变换. 证明 σ 是同构 $\iff$ A 和 $-B$ 没有相同的特征值.

73. 设 U, W 是线性空间 V 的两个子空间.
(1) 证明 $(U+W)/W$ 与 $U/(U\bigcap W)$ 是同构的线性空间;
(2) 试建立 (1) 中的不依赖于任何基的一个同构映射.

74. (双线性函数与张量积) 设 $A\in\mathbb{F}^{m\times n}$. 对任意 $x\in\mathbb{F}^m, y\in\mathbb{F}^m$, 定义

$$f(x,y)=x^TAy.$$

则 $f(x,y)$ 称为 $\mathbb{F}$ 上的一个 $m\times n$ 维的双线性函数, A 称为该双线性函数的矩阵. $\mathbb{F}$ 上的 $m\times n$ 维双线性函数的全体记为 $\mathcal{B}(m,n)$.

(1) 证明 $\mathcal{B}(m,n)$ 按照普通加法与数乘运算构成 $\mathbb{F}$ 上的线性空间;(对照第一章第五节思考题 6)

(2) 计算 $\mathcal{B}(m,n)$ 的维数与一组基;

(3) 建立 $\mathcal{B}(m,n)$ 与 $\mathrm{Hom}_{\mathbb{F}}(\mathbb{F}^m\otimes\mathbb{F}^n,\mathbb{F})=(\mathbb{F}^m\otimes\mathbb{F}^n)^*$ 之间的一个同构映射. 解释之.

75. 设 U 与 W 是线性空间 V 的两个子空间, $\alpha,\beta\in V$. 证明:

(1) $\alpha+U=P_{U^\perp}(\alpha)+U$;

(2) $(\alpha+U)+(\beta+W)=(\alpha+\beta)+(U+W)$;

(3) $(\alpha+U)\cap(\beta+W)\neq\emptyset \iff \alpha-\beta\in U+W$.

76. 证明迹函数 $\mathrm{tr}: X\mapsto \mathrm{tr}(X)$ 是线性空间 $M_n(\mathbb{F})$ 到 $\mathbb{F}$ 的满足性质 $\sigma(XY)=\sigma(YX)$ 以及 $\sigma(I)=n$ 的唯一线性变换.