

3A 线性映射的向量空间

线性映射的定义和实例

3.1 线性映射（线性变换）

函数 $T: V \rightarrow W$ 是从 V 到 W 的线性映射若满足

- 可加性 $\forall u, v \in V$, 均有 $T(u + v) = Tu + Tv$
 - 齐次性 $\forall \lambda \in \mathbf{F}, \forall v \in V, T(\lambda v) = \lambda(Tv)$
- 对线性映射常用记号 Tv , 也会用一般的函数记号 $T(v)$

3.2 $\mathcal{L}(V, W), \mathcal{L}(V)$

从 V 到 W 的全体线性映射构成的集合记作 $\mathcal{L}(V, W)$
 $\mathcal{L}(V) := \mathcal{L}(V, V)$

3.3 线性映射

1. 零映射

$0 \in \mathcal{L}(V, W)$ 定义为

$$0v = 0$$

LHS的 0 是从 V 到 W 的函数, RHS的 0 是 W 中的加法单位元
根据上下文区分符号 0 的多种用法

2. 恒等算子

恒等算子记作 I , 是某个向量空间上的一个这样的线性映射: 它将该空间中每个元素都对应到该元素本身. 具体而言, $I \in \mathcal{L}(V)$ 定义为

$$Iv = v$$

3. 微分映射

定义 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 为

$$Dp = p'$$

断言该线性映射等价于微积分中的: 任可微函数 f, g 和常数 λ , $(f + g)' = f' + g'$ 和 $(\lambda f)' = \lambda f'$ 总成立

4. Riemann积分映射

定义 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}), \mathbf{R})$ 为

$$Tp = \int_0^1 p$$

断言该线性映射用另一种方式表达了: 两个函数之和的积分等于这两个函数的积分之和, 一个函数乘以常数再作积分等于该函数作积分后乘以同一常数

5. 线性映射 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$:

$$\forall x \in \mathbf{R}, (Tp)(x) = x^2 p(x)$$

6. 移位映射

线性映射 $T \in \mathcal{L}(\mathbf{F}^\infty)$:

$$T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$$

7. 线性映射 $T \in \mathcal{L}(\mathbf{R}^3, \mathbf{R}^2)$

$$T(x, y, z) = (2x - y + 3z, 7x + 5y - 6z)$$

8. m 和 n 为正整数, $A_{j,k} \in \mathbf{F}$ ($j = 1, \dots, m$ 且 $k = 1, \dots, n$), 线性映射 $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \dots + A_{1,n}x_n, \dots, A_{m,1}x_1 + \dots + A_{m,n}x_n)$$

实际上每个从 \mathbf{F}^n 到 \mathbf{F}^m 的线性映射都是这种形式（后面讲表示矩阵就知道了）

9. $q \in \mathcal{P}(\mathbf{R})$

线性映射 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$

$$(Tp)(x) = p(q(x))$$

下个结论的

存在性说明总能构造一线性映射以将一基中各向量对应到任想要的值上

唯一性说明一线性映射完全决定于它将一个基映到的值

3.4 线性映射引理

假定 ν_1, \dots, ν_n 是 V 的基且 $w_1, \dots, w_n \in W$. 那么存在唯一的线性映射 $T: V \rightarrow W$ 使得对每个 $k = 1, \dots, n$ 都有

$$T\nu_k = w_k$$

证

【 T 的存在性】

定义 $T: V \rightarrow W$ 为

$$T(c_1\nu_1 + \dots + c_n\nu_n) = c_1w_1 + \dots + c_nw_n$$

其中 c_1, \dots, c_n 是 \mathbf{F} 中任意元素。组 ν_1, \dots, ν_n 是 V 的基，由此，上述方程的确定了一个从 V 到 W 的函数 T （因为 V 中每个元素都能被唯一地写成 $c_1\nu_1 + \dots + c_n\nu_n$ 的形式）

对于每个 k ，取上式中 $c_k = 1$ 且其他各 c 都等于 0 就证明了 $T\nu_k = w_k$

如果 $u, \nu \in V$ ，且有 $u = a_1\nu_1 + \dots + a_n\nu_n$ 和 $\nu = c_1\nu_1 + \dots + c_n\nu_n$ ，则

$$\begin{aligned} T(u + \nu) &= T((a_1 + c_1)\nu_1 + \dots + (a_n + c_n)\nu_n) \\ &= (a_1 + c_1)w_1 + \dots + (a_n + c_n)w_n \\ &= (a_1w_1 + \dots + a_nw_n) + (c_1w_1 + \dots + c_nw_n) \\ &= Tu + T\nu \end{aligned}$$

类似地，对 $\lambda \in \mathbf{F}$ ， $\nu = c_1\nu_1 + \dots + c_n\nu_n$ 有 $T(\lambda\nu) = \lambda T(\nu)$

故 T 就是从 V 到 W 的线性映射

【 T 的唯一性】

设 $T \in \mathcal{L}(V, W)$ 且对每个 $k = 1, \dots, n$ 都有 $T\nu_k = w_k$

令 $c_1, \dots, c_n \in \mathbf{F}$

那么 T 的齐次性表明对每个 $k = 1, \dots, n$ ，都有 $T(c_k\nu_k) = c_kw_k$

结合 T 的可加性可得

$$T(c_1\nu_1 + \dots + c_n\nu_n) = c_1w_1 + \dots + c_nw_n$$

于是 T 在 $\text{span}(\nu_1, \dots, \nu_n)$ 上由上式唯一确定。因为 ν_1, \dots, ν_n 是 V 的基，所以这表明 T 唯一

$\mathcal{L}(V, W)$ 上的代数运算



3.5 $\mathcal{L}(V, W)$ 上的加法和标量乘法

$S, T \in \mathcal{L}(V, W)$ ， $\lambda \in \mathbf{F}$

和 $S + T$ 与积 λT 都是从 V 到 W 的线性映射：

$$\forall \nu \in V, (S + T)(\nu) = S\nu + T\nu, (\lambda T)(\nu) = \lambda(T\nu)$$

读者自证如上的 $S + T$ 和 λT 确实是线性映射

换言之若 $S, T \in \mathcal{L}(V, W)$ 且 $\lambda \in \mathbf{F}$ ，那么 $S + T \in \mathcal{L}(V, W)$ ， $\lambda T \in \mathcal{L}(V, W)$

3.6 $\mathcal{L}(V, W)$ 是向量空间

$\mathcal{L}(V, W)$ 带有如上定义的加法和标量乘法是向量空间

读者自证。注意 $\mathcal{L}(V, W)$ 的加法恒等元是例3.3中定义为零线性映射

通常将一个向量空间中的两个元素相乘没有意义，但有时能构造出有用的乘积

3.7 线性映射的乘积

如果 $T \in \mathcal{L}(U, V)$ 且 $S \in \mathcal{L}(V, W)$ ，那么乘积 $ST \in \mathcal{L}(U, W)$ ：

$$\forall u \in U, (ST)(u) = S(Tu)$$

由此可见， ST 就是一般的两函数复合 $S \circ T$ ，不过当两个函数都是线性函数时，我们通常使用 ST 这个记号而不用 $S \circ T$ 。用 ST 这样的乘积表示法，有助于使下个结果中的分配性质看起来更自然些

注意只有当 T 映射到 S 的定义空间中时， ST 才有定义。读者验证对于任意 $T \in \mathcal{L}(U, V)$ 和 $S \in \mathcal{L}(V, W)$ ， ST 的确是从 U 到 W 的线性映射

3.8 线性映射乘积的代数性质

可结合性

对于任意使乘积有意义的线性映射 T_1, T_2, T_3 （意即 T_3 映射到 T_2 的定义空间中， T_2 映射到 T_1 的定义空间中），有 $(T_1 T_2) T_3 = T_1 (T_2 T_3)$

恒等元

对于任意 $T \in \mathcal{L}(V, W)$ ，有 $TI = IT = T$ 。这里第一个 I 是 V 上的恒等算子，而第二个 I 是 W 上的恒等算子

分配性

对于任意 $T, T_1, T_2 \in \mathcal{L}(U, V)$ 和 $S, S_1, S_2 \in \mathcal{L}(V, W)$ ，有 $(S_1 + S_2)T = S_1 T + S_2 T$ 且 $S(T_1 + T_2) = ST_1 + ST_2$

读者自证

线性映射的乘法通常不满足交换律，即便式子两侧的乘积都是有意义的

3.9 不可交换的两个从 $\mathcal{P}(\mathbf{R})$ 到 $\mathcal{P}(\mathbf{R})$ 的线性映射

设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是例 3.3 中定义的分微映射， $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是例 3.3 中定义的“与 x^2 相乘”映射。那么

$$((TD)p)(x) = x^2 p'(x)$$

$$((DT)p)(x) = x^2 p'(x) + 2xp(x)$$

于是 $TD \neq DT$ ——先作微分再乘以 x^2 ，和先乘以 x^2 再作微分是不一样的

3.10 线性映射将 0 映射为 0

T 是由 V 到 W 的线性映射，则 $T(0) = 0$

证

由可加性

$$T(0) = T(0 + 0) = T(0) + T(0)$$

等式两侧同加 $-(T(0))$ 得证

设 $m, b \in \mathbf{R}$. 当且仅当 $b = 0$ 时, 由

$$f(x) = mx + b$$

定义的函数 $f: \mathbf{R} \rightarrow \mathbf{R}$ 是线性映射 (由3.10)

3B 零空间和值域

零空间和单射性

3.11 零空间

$T \in \mathcal{L}(V, W)$, T 的零空间

$$\mathbf{n}T = \{\nu \in V : T\nu = 0\}$$

3.12 零空间

如果 T 是从 V 到 W 的零映射, 意即对于每个 $\nu \in V$ 都有 $T\nu = 0$, 那么 $\mathbf{null} T = V$

设 $T \in \mathcal{L}(\mathbf{C}^3, \mathbf{C})$, 定义为 $(z_1, z_2, z_3) \mapsto z_1 + 2z_2 + 3z_3$. 那么 $\mathbf{null} T$ 等于 $\{(z_1, z_2, z_3) \in \mathbf{C}^3 : z_1 + 2z_2 + 3z_3 = 0\}$, 它是 \mathbf{C}^3 的定义空间的子空间. 我们很快就会看到, 每个线性映射的零空间都是其定义空间的子空间

设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是微分映射, 即其定义为 $Dp = p'$. 导函数为 0 的函数仅有常值函数于是, D 的零空间等于常值函数构成的集合

设 $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是与 x^2 相乘映射, 即其定义为 $Tp(x) = x^2 p(x)$. 唯一使得 $x^2 p(x) = 0$ 对所有 $x \in \mathbf{R}$ 都成立的多项式 p 是多项式 0. 于是 $\mathbf{n}T = \{0\}$

设 $T \in \mathcal{L}(\mathbf{F}^\infty)$ 是后向移位映射, 即其定义为 $T(x_1, x_2, x_3, \dots) = (x_2, x_3, \dots)$. 那么 $T(x_1, x_2, x_3, \dots)$ 等于 0 当且仅当数 x_2, x_3, \dots 全为 0. 于是 $\mathbf{n}T = \{(a, 0, 0, \dots) : a \in \mathbf{F}\}$

接下来的结果说明, 每个线性映射的零空间都是其定义空间的子空间. 特别地, 每个线性映射的零空间都包含 0

3.13 零空间是子空间

设 $T \in \mathcal{L}(V, W)$. 那么 $\mathbf{null} T$ 是 V 的子空间

证

因为 T 是线性映射, 故 $T(0) = 0$ (由3.10). 于是 $0 \in \mathbf{n}T$

假定 $u, \nu \in \mathbf{n}T$. 那么

$$T(u + \nu) = Tu + T\nu = 0 + 0 = 0$$

所以 $u + \nu \in \mathbf{n}T$. 于是 $\mathbf{null} T$ 对于加法封闭“null”的意思是“零”. 于是, “null space”这个术语应该能让你想起它和 0 的联系. 有些数学家使用术语“核” (kernel) 而不是零空间

假定 $u \in \mathbf{n}T$ 且 $\lambda \in \mathbf{F}$. 那么

$$T(\lambda u) = \lambda Tu = \lambda 0 = 0$$

所以 $\lambda u \in \mathbf{n}T$. 于是 $\mathbf{null} T$ 对于标量乘法封闭

我们证明了 $\mathbf{n}T$ 含有 0, 并且对于加法和标量乘法运算都封闭. 因此 $\mathbf{null} T$ 是 V 的子空间 (由 1.34)

我们很快就会看到, 对于线性映射来说, 接下来这条定义和零空间紧密相关

3.14 单射 (一对一)

对于函数 $T: V \rightarrow W$ ，若 $Tu = T\nu$ 蕴涵 $u = \nu$ （如果 $u \neq \nu$ 蕴涵 $Tu \neq T\nu$ ）（当且仅当 T 将不同的输入映射到不同的输出时），则称该函数是单射

接下来的结果说，可通过检验0是否为唯一被映射为0的向量来检验线性映射是否为单射

3.12 中我们计算了一些线性映射的零空间，下面结论可看出这些线性映射中仅有“与 x^2 相乘”映射是单射（特殊情况 $V = \{0\}$ 下，零映射也是单射）

3.15 单射性 零空间等于 0

$T \in \mathcal{L}(V, W)$ 是单射 $\iff \text{null } T = \{0\}$

证

\Rightarrow)

设 T 是单射，想要证明 $\text{null } T = \{0\}$

已知 $\{0\} \subseteq \text{null } T$ （由3.10）

为证这个包含反过来也成立，设 $\nu \in \text{null } T$ ，则

$$T(\nu) = 0 = T(0)$$

因为 T 是单射，所以上述等式表明 $\nu = 0$ 。于是我们能得出 $\text{null } T = \{0\}$ ，此方向得证 \Leftarrow)

设 $\text{null } T = \{0\}$

想证 T 是单射

设 $u, \nu \in V$ 且 $Tu = T\nu$ ，则

$$0 = Tu - T\nu = T(u - \nu).$$

于是 $u - \nu$ 在 $\text{null } T$ 中，而 $\text{null } T$ 等于 $\{0\}$

故 $u - \nu = 0$ ，则 $u = \nu$

故 T 是单射

值域和满射性

现在我们给线性映射的输出所构成的集合起个名字

3.16 值域

$T \in \mathcal{L}(V, W)$ ， T 的值域

$$\text{range } T = \{T\nu : \nu \in V\}$$

3.17 值域

若 T 是从 V 到 W 的零映射，意即对于所有的 $\nu \in V$ 都有 $T\nu = 0$ ，那么 $\text{range } T = \{0\}$ 假设 $T \in \mathcal{L}(\mathbf{R}^2, \mathbf{R}^3)$ 定义为 $T(x, y) = (2x, 5y, x + y)$

注意到 $\text{range } T$ 是 \mathbf{R}^3 的子空间。我们很快将看到， $\mathcal{L}(V, W)$ 中每个线性映射的值域都是 W 的子空间

设 $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ 是由 $Dp = p'$ 定义的微分映射。因为对于每个多项式 $q \in \mathcal{P}(\mathbf{R})$ 都存在一个多项式 $p \in \mathcal{P}(\mathbf{R})$ 使得 $p' = q$ ，所以 D 的值域是 $\mathcal{P}(\mathbf{R})$

接下来的结果表明，每个线性映射的值域都是它所映射到 W 的向量空间的子空间

3.18 值域是子空间

$T \in \mathcal{L}(V, W)$, 则 $\text{range } T$ 是 W 的子空间

证

设 $T \in \mathcal{L}(V, W)$. 那么 $T(0) = 0$ (根据3.10), 意味着 $0 \in \text{range } T$

如果 $w_1, w_2 \in \text{range } T$, 那么存在 $\nu_1, \nu_2 \in V$ 使 $T\nu_1 = w_1$ 和 $T\nu_2 = w_2$ 成立. 于是

$$T(\nu_1 + \nu_2) = T\nu_1 + T\nu_2 = w_1 + w_2$$

所以 $w_1 + w_2 \in \text{range } T$, 于是 $\text{range } T$ 对于加法封闭

如果 $w \in \text{range } T$ 且 $\lambda \in \mathbf{F}$, 那么存在 $\nu \in V$ 使 $T\nu = w$ 成立. 于是

$$T(\lambda\nu) = \lambda T\nu = \lambda w$$

所以 $\lambda w \in \text{range } T$, 于是 $\text{range } T$ 对于标量乘法封闭

我们证明了 $\text{range } T$ 含有 0 以及对于加法和标量乘法运算都封闭. 因此 $\text{range } T$ 是 W 的子空间 (由1.34)



3.19 满射 (映成)

若函数 $T: V \rightarrow W$ 的值域等于 W , 则称该函数为满射

3.17 中计算过值域的那些映射, 其中只有微分映射是满射 (特殊情况 $W = \{0\}$, 零映射也是满射)
一个线性映射是否为满射, 取决于我们认为它映射到哪个向量空间



3.20 是否为满射取决于目标空间的选取

定义为 $\frac{d}{dx}$ 的微分映射 $D \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ 不是满射, 因为多项式 x^5 不在 D 的值域中然而, 定义为 $\frac{d}{dx}$ 的微分映射 $S \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}), \mathcal{P}_4(\mathbf{R}))$ 是满射, 因为它的值域等于 $\mathcal{P}_4(\mathbf{R})$, 恰是 S 映射到的向量空间

线性映射基本定理

下个结果十分重要, 所以它有个引人注目的名字



3.21 线性映射基本定理

V 是有限维的, $T \in \mathcal{L}(V, W)$, 则 $\text{range } T$ 是有限维的, 且

$$\dim V = \dim \text{null } T + \dim \text{range } T$$

证

令 u_1, \dots, u_m 是 $\text{null } T$ 的基, 于是 $\dim \text{null } T = m$. 由 2.32, 线性无关组 u_1, \dots, u_m 可被扩充成 V 的基

$$u_1, \dots, u_m, \nu_1, \dots, \nu_n$$

于是 $\dim V = m + n$

需证 $\text{range } T$ 是有限维的且 $\dim \text{range } T = n$. 为此, 我们证明 $T\nu_1, \dots, T\nu_n$ 是 $\text{range } T$ 的基

令 $\nu \in V$, 因为 $u_1, \dots, u_m, \nu_1, \dots, \nu_n$ 张成 V , 所以我们可以写出

$$\nu = a_1 u_1 + \dots + a_m u_m + b_1 \nu_1 + \dots + b_n \nu_n$$

其中各 a 和 b 都属于 \mathbf{F} . 将 T 作用于等式两侧, 可得

$$T\nu = b_1 T\nu_1 + \dots + b_n T\nu_n$$

其中形如 Tu_k 的项消失了, 因为各 u_k 均属于 $\text{null } T$. 上方程表明组 $T\nu_1, \dots, T\nu_n$ 张成 $\text{range } T$ 特别地, $\text{range } T$ 是有限维的

为证 $T\nu_1, \dots, T\nu_n$ 是线性无关的, 设 $c_1, \dots, c_n \in \mathbf{F}$ 且 $c_1T\nu_1 + \dots + c_nT\nu_n = 0$

那么 $T(c_1\nu_1 + \dots + c_n\nu_n) = 0$

因此 $c_1\nu_1 + \dots + c_n\nu_n \in \mathbf{n}T$

因 u_1, \dots, u_m 张成 $\mathbf{n}T$, 所以可以写出 $c_1\nu_1 + \dots + c_n\nu_n = d_1u_1 + \dots + d_mu_m$

其中各 d 都属于 \mathbf{F} . 这个等式表明, 所有各 c (以及各 d) 都是 0 (因为 $u_1, \dots, u_m, \nu_1, \dots, \nu_n$ 是线性无关的). 于是 $T\nu_1, \dots, T\nu_n$ 是线性无关的, 进而它是 $\text{range } T$ 的基. 原命题得证

现在可证将一个有限维向量空间映射到比它“更小”的向量空间的线性映射都不是单射, 这里的“更小”是用维数来度量的

3.22 映到更低维空间上的线性映射不是单射

设 V 和 W 是有限维向量空间且满足 $\dim V > \dim W$, 则从 V 到 W 的线性映射一定不是单射

Example

令 $T \in \mathcal{L}(V, W)$, 则

$$\begin{aligned} \dim V - \dim W \\ > 0 \end{aligned}$$

上述第一行来自线性映射基本定理 (3.21), 第二行来自 2.37. 上述不等式表明, $\dim \text{null } T > 0$ 这就意味着 $\text{null } T$ 包含除了 0 以外的其他向量. 于是 T 不是单射 (由 3.15)

3.23 从 \mathbf{F}^4 到 \mathbf{F}^3 的线性映射不是单射

按下式定义一个线性映射 $T: \mathbf{F}^4 \rightarrow \mathbf{F}^3$:

$$T(z_1, z_2, z_3, z_4) = (\sqrt{7}z_1 + z_2 + z_4, 97z_1 + 3z_2 + 2z_3, z_2 + 6z_3 + 7z_4)$$

因为 $\dim \mathbf{F}^4 > \dim \mathbf{F}^3$, 由 3.22 T 不是单射

接下来的结果表明, 从一个有限维向量空间映射到比它“更大”的向量空间的线性映射都不是满射, 这里的“更大”是从维数上来度量的

3.24 映到更高维空间上的线性映射不是满射

V 和 W 是有限维向量空间且满足 $\dim V < \dim W$. 那么从 V 到 W 的线性映射一定不是满射

证

令 $T \in \mathcal{L}(V, W)$, 则

$$\begin{aligned} &\leq \dim V \\ &< \dim W, \end{aligned}$$

上述式子中的等号源于线性映射基本定理 (3.21). 该不等式表明, $\dim \text{range } T < \dim W$. 这就意味着 $\text{range } T$ 不会等于 W . 于是 T 不是满射

我们马上就会看到, 3.22 和 3.24 在线性方程理论中有一些重要推论, 其中的想法是将有关线性方程组的问题用线性映射的语言来表达. 首先, 我们将齐次线性方程组是否有非零解这个问题用线性映射的语言来改写

固定正整数 m 和 n , 并令 $A_{j,k} \in \mathbf{F}$ ($j = 1, \dots, m$ 且 $k = 1, \dots, n$). 考虑齐次线性方程组 (每个方程右端的常数项都是 0)

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

$$\sum_{k=1}^n A_{m,k} x_k = 0.$$

显然, $x_1 = \cdots = x_n = 0$ 是上述方程组的一个解. 此处的问题是, 是否存在其他的解

定义 $T: \mathbf{F}^n \rightarrow \mathbf{F}^m$ 为

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^n A_{1,k} x_k, \dots, \sum_{k=1}^n A_{m,k} x_k \right)$$

方程 $T(x_1, \dots, x_n) = 0$ (此处的 0 是 \mathbf{F}^m 中的加法恒等元, 也就是长度为 m 的全 0 构成的组) 与上面的齐次线性方程组是一样的. 于是我们就想知道 $\text{null } T$ 是否严格大于 $\{0\}$, 这也就等价于 T 不是单射 (由 3.15). 接下来的结果给出了保证 T 不是单射的条件

3.26 齐次线性方程组

未知数个数多于方程个数的齐次线性方程组具有非零解

证

沿用前面讨论中所用的记号和结论. 于是 T 是从 \mathbf{F}^n 到 \mathbf{F}^m 的线性映射, 且我们有共含 m 个方程、 n 个未知数 x_1, \dots, x_n 的齐次线性方程组. 由 3.22 即可知若 $n > m$ 则 T 就不是单射.

上述结果的一个实例: 具有四个方程和五个未知数的齐次线性方程组具有非零解

一个线性方程组是否在常数项取某些值时会是无解的? 为将这个问题用线性映射的语言来描述, 固定正整数 m 和 n , 并令 $A_{j,k} \in \mathbf{F}$ ($j = 1, \dots, m$ 且 $k = 1, \dots, n$), 对于 $c_1, \dots, c_m \in \mathbf{F}$, 考虑线性方程组

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

$$\sum_{k=1}^n A_{m,k} x_k = c_m.$$

用这样的记号, 我们可将问题表述成: 是否存在常数项 $c_1, \dots, c_m \in \mathbf{F}$ 的某些取值, 使得上述方程组无解

如式 (3.25) 一样定义 $T: \mathbf{F}^n \rightarrow \mathbf{F}^m$. 方程组 $T(x_1, \dots, x_n) = (c_1, \dots, c_m)$ 与上面式 (3.27) 的个数作比较. 它们也可利用高斯消元法证得方程组是一样的. 于是我们想知道是明, 但此处所用的抽象方法似乎提供了更简否有 $\text{range } T \neq \mathbf{F}^m$. 所以我们可以将问题的证明($c_1, \dots, c_m \in \mathbf{F}$ 的某些取值使得方程组无解)改写成这样: 什么条件能使 T 一定不是满射? 接下来的结论就给出了一个这样的条件

3.28 方程个数多于未知数个数的线性方程组

方程个数多于未知数个数的线性方程组当常数项取某些值时无解

证

沿用前面讨论中所用的记号和结论. 于是 T 是从 \mathbf{F}^n 到 \mathbf{F}^m 的线性映射, 且我们有共含 m 个方程、 n 个未知数 x_1, \dots, x_n 的线性方程组【见式 (3.27)】. 若 $n < m$, 由 3.24 即可知 T 不是满射. 按照上面的讨论, 这表明, 如果一个线性方程组中, 方程个数多于未知数个数, 那么当常数项取某些值时该方程组无解

上述结果的一个实例: 具有五个方程和四个未知数的线性方程组在常数项取某些值时会是无解的

3C 矩阵

用矩阵表示线性映射

我们知道, 如果 ν_1, \dots, ν_n 是 V 的基且 $T: V \rightarrow W$ 是线性的, 那么 $T\nu_1, \dots, T\nu_n$ 的取值决定了 T 将 V 中任意向量所映射到的值——见线性映射引理 (3.4). 利用 W 的基, 矩阵可以高效地记录各 $T\nu_k$ 的值

3.29 矩阵

设 m 和 n 是非负整数. $m \times n$ 矩阵 A 是由 \mathbf{F} 中元素构成的 m 行 n 列的矩形阵列:

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix}.$$

记号 $A_{j,k}$ 表示 A 的第 j 行第 k 列中的元素

3.30 矩阵

$A_{j,k}$ 等于 A 的第 j 行第 k 列中的元素假设 $A = \begin{pmatrix} 8 & 4 & 5 - 3i \\ 1 & 9 & 7 \end{pmatrix}$. 于是 $A_{2,3}$ 表示 A 的第二行第三列中的元素, 意即 $A_{2,3} = 7$

讨论矩阵时, 第一个下标代表第几行, 第二个坐标代表第几列

现在我们给出本节中的关键定义

3.31 线性映射的矩阵

假设 $T \in \mathcal{L}(V, W)$, ν_1, \dots, ν_n 是 V 的基, w_1, \dots, w_m 是 W 的基, T 关于这些基的矩阵是 $m \times n$ 矩阵 (T) , 其中各元素 $A_{j,k}$ 由下式定义:

$$T\nu_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m$$

如果从上下文无法明确得知基 ν_1, \dots, ν_n 和 w_1, \dots, w_m 取什么, 那么就用 $(T, (\nu_1, \dots, \nu_n), (w_1, \dots, w_m))$ 这个记号

线性映射 $T \in \mathcal{L}(V, W)$ 的矩阵 (T) 决定于 V 的基 ν_1, \dots, ν_n , W 的基 w_1, \dots, w_m , 以及 T . 不过, 从上下文中往往能明确得知这些基, 因此常常把它们从记号里省去

为记住 (T) 是如何由 T 构造出来的, 你可以在矩阵的上方横着标上定义空间的基向量 ν_1, \dots, ν_n , 并在矩阵的左侧竖着列出 T 映射到的向量空间的基向量 w_1, \dots, w_m , 就像下面这样:

$$(T) = \begin{matrix} & \nu_1 & \cdots & \nu_k & \cdots & \nu_n \\ \begin{matrix} w_1 \\ \vdots \\ w_m \end{matrix} & A_{1,k} \\ & A_{m,k} \end{matrix}.$$

我们只展示了上述矩阵的第 k 列, 于是此处矩阵各元素的第二个下标都是 k . 上面这种描述还会让你想出由 (T) 计算 $T\nu_k$ 的方法: 先将第 k 列中的各元素分别与它们左侧对应的 w_j 相乘, 再将所得各向量都加起来

如果 T 是从 \mathbf{F}^n 到 \mathbf{F}^m 的线性映射, 那么除非另有说明, 总假定所讨论的基都是标准基 (T) 的第 k 列包含了将 $T\nu_k$ 写成 w_1, \dots, w_m 的线性组合所需的各标量:

$$T\nu_k = \sum_{j=1}^m A_{j,k}w_j.$$

如果 T 是从 n 维向量空间到 m 维向量空间的线性映射, 那么 (T) 是 $m \times n$ 矩阵

基 (其中第 k 个向量的第 k 个坐标是 1 而其他各坐标都是 0). 如果你将 \mathbf{F}^m 中的元素看成由 m 个数组成的列, 那么你可以将 (T) 中的第 k 列看成将 T 作用于第 k 个标准基向量所得

3.32 从 \mathbf{F}^2 到 \mathbf{F}^3 的线性映射的矩阵

$T \in \mathcal{L}(\mathbf{F}^2, \mathbf{F}^3)$ 定义为

$$T(x, y) = (x + 3y, 2x + 5y, 7x + 9y).$$

因为 $T(1, 0) = (1, 2, 7)$ 且 $T(0, 1) = (3, 5, 9)$, 所以 T 关于标准基的矩阵就是如下所示的 3×2 矩阵:

$$\begin{pmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{pmatrix}$$

在研究 $\mathcal{P}_m(\mathbf{F})$ 时, 除非上下文有另作说明, 都取标准基 $1, x, x^2, \dots, x^m$

3.33 从 $\mathcal{P}_3(\mathbf{R})$ 到 $\mathcal{P}_2(\mathbf{R})$ 的微分映射的矩阵

设 $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ 是由 $p = p'$ 所定义微分映射. 因为 $(x^n)' = nx^{n-1}$, 所以 D 关于标准基的矩阵就是如下所示的 3×4 矩阵:

$$(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

矩阵的加法和标量乘法

在本节的剩余部分, 假定 U, V, W 是有限维的且对于其中每个向量空间都已选定了个基于是对于每个从 V 到 W 的线性映射, 我们都能够讨论它们 (关于所选择的基) 的矩阵

两个线性映射之和的矩阵, 是否等于这两个映射的矩阵之和呢? 这个问题目前还没有意义, 因为尽管我们已定义两个线性映射的和, 但我们还没定义两个矩阵的和. 幸运的是, 按一种很自然的方式来作出两矩阵之和的定义, 它便具备我们期望的性质. 具体而言, 我们给出以下定义

3.34 矩阵加法

两个相同大小的矩阵之和, 是将两矩阵对应位置上的元素相加所得的矩阵:

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} + \begin{pmatrix} B_{1,1} & \cdots & B_{1,n} \\ \vdots & \ddots & \vdots \\ B_{m,1} & \cdots & B_{m,n} \end{pmatrix} = \begin{pmatrix} A_{1,1} + B_{1,1} & \cdots & A_{1,n} + B_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} + B_{m,1} & \cdots & A_{m,n} + B_{m,n} \end{pmatrix}$$

在接下来这条结果中, 假设对于 $S + T, S$ 和 T 这三个线性映射都选取相同的基

3.35 线性映射之和的矩阵

设 $S, T \in \mathcal{L}(V, W)$ 那么 $(S + T) = (S) + (T)$

由定义即可验证上面这条结论, 验证过程留给读者完成

3.36 矩阵的标量乘法

$$\lambda \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} = \begin{pmatrix} \lambda A_{1,1} & \cdots & \lambda A_{1,n} \\ \vdots & \ddots & \vdots \\ \lambda A_{m,1} & \cdots & \lambda A_{m,n} \end{pmatrix}$$

3.37 矩阵的加法和标量乘法

$$2 \begin{pmatrix} 6 & 1 \\ -1 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ -2 & 10 \end{pmatrix} + \begin{pmatrix} 4 & 2 \\ 1 & 6 \end{pmatrix} = \begin{pmatrix} 10 & 4 \\ -1 & 16 \end{pmatrix}$$

在接下来的结果中, 假设对于 λT 和 T 这两个线性映射都选取相同的基

3.38 标量与线性映射之积的矩阵

设 $\lambda \in \mathbf{F}$ 且 $T \in \mathcal{L}(V, W)$. 那么 $(\lambda T) = \lambda(T)$

上述结论的验证同样留给读者完成

因为现已定义了矩阵的加法和标量乘法, 由此产生一个向量空间也就不足为奇了. 我们先引入个记号, 给这个新的向量空间取个名字, 然后得出这个新的向量空间的维数

3.39 记号

对于正整数 m 和 n , 各元素均属于 \mathbf{F} 的所有 $m \times n$ 矩阵构成的集合记作 $\mathbf{F}^{m,n}$

3.40 $\dim \mathbf{F}^{m,n} = mn$

m 和 n 为正整数. 按上面定义的加法和标量乘法, $\mathbf{F}^{m,n}$ 是维数为 mn 的向量空间

读者自证。

注意 $\mathbf{F}^{m,n}$ 的加法恒等元是所有元素都等于 0 的 $m \times n$ 矩阵

读者也应自行验证, 所有 $m \times n$ 矩阵中, 那些除某元素等于 1 外其余各元素都等于 0 的互异矩阵, 构成 $\mathbf{F}^{m,n}$ 的一个基. 这样的矩阵共有 mn 个, 因此 $\mathbf{F}^{m,n}$ 的维数等于 mn

矩阵乘法

与之前一样, 假设 ν_1, \dots, ν_n 是 V 的基, w_1, \dots, w_m 是 W 的基. 另设 u_1, \dots, u_p 是 U 的基

考虑线性映射 $T: U \rightarrow V$ 和 $S: V \rightarrow W$. 复合映射 ST 是从 U 到 W 的映射. 那么 $\mathcal{M}()$ 等于 $(S)(T)$ 吗? 这个问题目前还没有意义, 因为我们还没有定义两矩阵的乘积. 我们将选定一种矩阵乘法的定义, 来让上面的问题必有肯定的回答. 现在来看该怎么做

设 $(S) = A$ 且 $(T) = B$. 对于 $1 \leq k \leq p$,

$$\begin{aligned}(ST)u_k &= S\left(\sum_{j=1}^n B_{j,k}\nu_j\right) \\ &= \sum_{j=1}^n B_{j,k}S\nu_j \\ &= \sum_{j=1}^n B_{j,k} \sum_{i=1}^m A_{i,j}w_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n A_{i,j}B_{j,k}\right)w_i.\end{aligned}$$

于是 $()$ 是 $m \times p$ 矩阵, 其中第 j 行第 k 列的元素等于

$$\sum_{j=1}^n A_{j,k}B_{j,k}.$$

这样我们就明白了如何定义矩阵乘法才能使我们期望的等式 $(ST) = (S)(T)$ 成立

3.41 矩阵乘法

设 A 是 $m \times n$ 矩阵且 B 是 $n \times p$ 矩阵. 那么 AB 定义为一个 $m \times p$ 矩阵, 其中第 j 行第 k 列的元素由下式给出:

$$(AB)_{j,k} = \sum_{i=1}^n A_{j,i}B_{i,k}.$$

于是, 取 A 的第 j 行和 B 的第 k 列, 将它们对应位置上的元素分别相乘再相加, 就得到了 AB 第 j 行第 k 列的元素

注意只有当第一个矩阵的列数等于第二个矩阵的行数时, 我们才能定义这两个矩阵的乘积

你可能在之前的课程中已经学过矩阵乘法的定义, 尽管你可能并没看出这里所说的定义它的动机

3.42 矩阵乘积

将一个 3×2 矩阵和一个 2×4 矩阵相乘, 得到一个 3×4 矩阵:

$$\begin{pmatrix} 1 & 2 & 6 & 5 & 4 & 3 \\ 3 & 4 & 2 & 1 & 0 & -1 \\ 5 & 6 & 2 & 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 10 & 7 & 4 & 1 \\ 26 & 19 & 12 & 5 \\ 42 & 31 & 20 & 9 \end{pmatrix}$$

矩阵乘法不满足交换律—— AB 并不一定等于 BA ，即使两个乘积都有定义（见习题 10）矩阵乘法满足分配律和结合律（见习题11、12）

接下来的结果中假设考虑 $T \in \mathcal{L}(U, V)$ 和 $S \in \mathcal{L}(V, W)$ 时取 V 中的同一个基，考虑 $S \in \mathcal{L}(V, W)$ 和 $ST \in \mathcal{L}(U, W)$ 时取 W 中的同一个基，考虑 $T \in \mathcal{L}(U, V)$ 和 $ST \in \mathcal{L}(U, W)$ 时取 U 中的同一个基

3.43 线性映射之积的矩阵

若 $T \in \mathcal{L}(U, V)$ 且 $S \in \mathcal{L}(V, W)$ ，那么 $(ST) = (S)(T)$

上述结果的证明，就是在定义矩阵乘法之前，我们说明其动机时所做的计算
在下面的记号中，注意，如往常一样，第一个下标代表行，第二个下标代表列；垂直居中的点“.”用于占位

3.44 记号

A 是 $m \times n$ 矩阵

若 $1 \leq j \leq m$ ，那么 $A_{j, \cdot}$ 表示由 A 的第 j 行构成的 $1 \times n$ 矩阵如果 $1 \leq k \leq n$ ，那么 $A_{\cdot, k}$ 表示由 A 的第 k 列构成的 $m \times 1$ 矩阵

3.45 $A_{j, \cdot}$ 等于 A 的第 j 行且 $A_{\cdot, k}$ 等于 A 的第 k 列

记号 $A_{2, \cdot}$ 表示 A 的第 2 行， $A_{\cdot, 2}$ 表示 A 的第 2 列。于是如果 $A = \begin{pmatrix} 8 & 4 & 5 \\ 1 & 9 & 7 \end{pmatrix}$ 那么 $A_{2, \cdot} = (1 \quad 9 \quad 7)$, $A_{\cdot, 2} = (9)$

$1 \times n$ 矩阵和 $n \times 1$ 矩阵的乘积是 1×1 矩阵。然而，我们常常会把 1×1 矩阵与其元素等同起来看。例如，

$$(34)(6) = (26)$$

因为 $3 \cdot 6 + 4 \cdot 2 = 26$ 。然而，我们可以把(26)与26等同起来看，也就是将上式写成 $(34)\begin{pmatrix} 6 \\ 2 \end{pmatrix} = 26$ 。

接下来的结论就采用了上段里讨论的这个约定，从而给出了另一种思考矩阵乘法的方式举一实例：将接下来的结论结合上段所作的计算，就能解释为什么例 3.42 中所得乘积的第 2 行第 1 列中的元素等于 26

3.46 矩阵之积的元素等于行乘以列

设 A 是 $m \times n$ 矩阵且 B 是 $n \times p$ 矩阵。那么如果 $1 \leq j \leq m$ 且 $1 \leq k \leq p$ ，则

$$(AB)_{j,k} = A_{j, \cdot} B_{\cdot, k}.$$

换言之， AB 中第 j 行第 k 列的元素等于：(A 的第 j 行) 乘以 (B 的第 k 列)

证明

设 $1 \leq j \leq m$ 且 $1 \leq k \leq p$ 。矩阵乘法的定义表明

$$(AB)_{j,k} = A_{j,1}B_{1,k} + \cdots + A_{j,n}B_{n,k}.$$

矩阵乘法的定义还表明， $1 \times n$ 矩阵 $A_{j, \cdot}$ 和 $n \times 1$ 矩阵 $B_{\cdot, k}$ 的乘积是一个 1×1 矩阵，且该矩阵的元素就是上述等式的右端。于是 AB 中第 j 行第 k 列的元素就等于：(A 的第 j 行) 乘以 (B 的第 k 列)

接下来的结论又给出了另一种思考矩阵乘法的方式. 在下面的结论中, $(AB)_{\cdot,k}$ 是 $m \times p$ 矩阵 AB 的第 k 列, 于是 $(AB)_{\cdot,k}$ 是 $m \times 1$ 矩阵. 同时, 因为 $AB_{\cdot,k}$ 是 $m \times n$ 矩阵和 $n \times 1$ 矩阵的乘积, 所以它也是 $m \times 1$ 矩阵. 于是下述结论中等式两侧的矩阵大小相同, 猜测它们相等就是合理的了

3.48 矩阵之积的列等于矩阵与列之积

假设 A 是 $m \times n$ 矩阵且 B 是 $n \times p$ 矩阵. 那么如果 $1 \leq k \leq p$, 则

$$(AB)_k = AB_k$$

(AB 的第 k 列等于 A 乘以 B 的第 k 列)

证

上面已讨论过, $(AB)_{\cdot,k}$ 和 $AB_{\cdot,k}$ 都是 $m \times 1$ 矩阵. 如果 $1 \leq j \leq m$, 那么 $(AB)_{\cdot,k}$ 中第 j 行的元素是式(3.47)的左端项, $AB_{\cdot,k}$ 中第 j 行的元素是式(3.47)的右端项. 得证

接下来的结论由下面例子所启发, 给出了另一种思考 $m \times n$ 矩阵和 $n \times 1$ 矩阵之积的方式

3.49 例: 3×2 矩阵和 2×1 矩阵的乘积

利用定义和基础的算术运算, 我们可验证下式:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 7 \\ 1 \end{pmatrix} = 19 = 53 + 14.$$

于是在本例中, 3×2 矩阵和 2×1 矩阵的乘积是 3×1 矩阵各列的线性组合, 而与各列相乘的标量 (5 和 1) 则来自 2×1 矩阵

下面的结论推广了上述例子

3.50 列的线性组合

设 A 是 $m \times n$ 矩阵且 $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ 是 $n \times 1$ 矩阵. 那么 $Ab = b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$

换言之, Ab 是 A 中各列的线性组合, 而与这些列相乘的标量则来自 b

证

若 $k \in \{1, \dots, m\}$, 那么由矩阵乘法的定义知, 所得 $m \times 1$ 矩阵 Ab 的第 k 行中的元素是

$$A_{k,1}b_1 + \cdots + A_{k,n}b_n$$

而 $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$ 的第 k 行中的元素也等于该数

因为对于每个 $k \in \{1, \dots, m\}$, Ab 和 $b_1 A_{\cdot,1} + \cdots + b_n A_{\cdot,n}$ 的第 k 行元素都相等, 得证

前述两个结论重点关注矩阵的列. 类似的结论对于矩阵的行也是成立的, 详见习题第8、9题. 只要适当修改3.48和3.50的证明过程, 即可证明接下来的结论, 是在下个小节中证明行列分解 (3.56) 及矩阵的列秩等于行秩 (3.57) 时, 所用的主要工具. 为了和叙述行列分解时常用的记号保持一致, 下面结论乃至下个小节中的矩阵记作 c 和 r 而不是 A 和 B

3.51 将矩阵乘法视为列或行的线性组合

设 c 是 $m \times c$ 矩阵且 c 是 $c \times n$ 矩阵

(a) 如果 $k \in \{1, \dots, n\}$, 那么 c 的第 k 列是 c 的各列的线性组合, 其中各系数来自 c 的第 k 列

(b) 如果 $j \in \{1, \dots, m\}$, 那么 c 的第 j 行是 c 的各行的线性组合, 其中各系数来自 c 的第 j 行

证

设 $k \in \{1, \dots, n\}$. 那么 c 的第 k 列等于 $c_{:,k}$ (由 3.48). 而 $c_{:,k}$ 还等于 c 中各列的线性组合, 其中系数取自 $c_{:,k}$ (由 3.50). 于是 (a) 成立

证明 (b), 仿照 (a) 的证明即可, 只需用行代替列, 并用习题第 8、9 题的结论代替 3.48 和 3.50

行列分解和矩阵的秩

我们从定义与每个矩阵都相关的两个非负整数开始

3.52 列秩 行秩

A 是 $m \times n$ 矩阵, 其各元素 $\in \mathbf{F}$

A 的列秩是 A 的各列在 $\mathbf{F}^{m,1}$ 中的张成空间的维数

A 的行秩是 A 的各列在 $\mathbf{F}^{1,n}$ 中的张成空间的维数

命题

A 是 $m \times n$ 矩阵, 那么 A 的列秩不超过 n (因为 A 有 n 列) 同时不超过 m (因为 $\dim \mathbf{F}^{m,1} = m$). 类似地, A 的行秩也不超过 $\min\{m, n\}$

3.53 列秩和行秩

假设

$$A = \begin{pmatrix} 4 & 7 & 1 & 8 \\ 7 & 5 & 2 & 9 \end{pmatrix}$$

A 的列秩是 $\mathbf{F}^{2,1}$ 中的

$$\text{span} \left(\begin{pmatrix} 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 8 \\ 9 \end{pmatrix} \right)$$

的维数. 上面列出的 $\mathbf{F}^{2,1}$ 中各向量的前两个不成标量倍数关系. 于是这个长度为 4 的向量组的张成空间的维数至少是 2. 而该向量组在 $\mathbf{F}^{2,1}$ 中, 则其张成空间的维数不超过 2, 因为 $\dim \mathbf{F}^{2,1} = 2$. 于是这个组的张成空间的维数就是 2, 意味着 A 的列秩是 2

A 的行秩是 $\mathbf{F}^{1,4}$ 中的

$$\text{span} \left((4 \ 7 \ 1 \ 8), (7 \ 5 \ 2 \ 9) \right)$$

的维数. 上面列出的 $\mathbf{F}^{1,4}$ 中的两个向量不成标量倍数关系. 于是这个长度为 2 的向量组的张成空间的维数是 2, 这意味着 A 的行秩是 2

3.54 转置

A 是 $m \times n$ 矩阵, 那么 A^T 是 $n \times m$ 矩阵

$$(A^T)_{k,j} = A_{j,k}$$

3.55 转置

如果 $A = \begin{pmatrix} 5 & -7 \\ 3 & 8 \\ -4 & 2 \end{pmatrix}$, , 那么 $A^T = \begin{pmatrix} 5 & 3 & -4 \\ -7 & 8 & 2 \end{pmatrix}$ 注意此处 A 是 3×2 矩阵而 A^T 是 2×3 矩阵

矩阵的转置有着很好的代数性质: 对于所有 $m \times n$ 矩阵 A, B 、所有 $\lambda \in \mathbf{F}$ 和所有 $n \times p$ 矩阵 C , 都有 $(A+B)^T = A^T + B^T$, $(\lambda A)^T = \lambda A^T$ 以及 $(A^T)^T = A$ (见习题第 14、15 题)

接下来的结果将是用于证明列秩等于行秩的主要工具 (见 3.57)

3.56 行列分解

A 是 $m \times n$ 矩阵, 其中各元素均在 \mathbf{F} 中且列秩 c . 那么存在各元素均属于 \mathbf{F} 的 $m \times c$ 矩阵 C 和 $c \times n$ 矩阵 R , 使得 $A = CR$ 成立

证

A 的各列都是 $m \times 1$ 矩阵. 由 A 的各列构成的组 $A_{\cdot 1}, \dots, A_{\cdot c}$ 可以被削减成 A 的各列的张成空间的基 (由 2.30). 由列秩的定义, 这个基的长度是 c . 将该基中的 Ψ_c 个列向量合在一起就形成了 $m \times c$ 矩阵 C

如果 $k \in \{1, \dots, n\}$, 那么 A 的第 k 列是 C 的各列的线性组合. 令该线性组合中的系数组成一个 $c \times n$ 矩阵的第 k 列, 并记该矩阵为 R . 那么由 3.51 (a) 得, $A = CR$

例 3.53 中的矩阵满足列秩等于行秩. 接下来的结论表明, 这一点对于每个矩阵都成立

3.57 列秩等于行秩

$A \in \mathbf{F}^{m,n}$, 那么 A 的列秩等于 A 的行秩

证

令 c 表示 A 的列秩. 令 $A = CR$ 是由 3.56 给出的 A 的行列分解式, 其中 C 是 $m \times c$ 矩阵且 R 是 $c \times n$ 矩阵. 那么 3.51 (b) 告诉我们, A 的每一行都是 R 的各行的线性组合. 因为 R 有 Ψ_c 行, 所以这表明 A 的行秩小于或等于 Ψ_c (也就是 A 的列秩)

为证明上面这个不等关系反过来也成立, 将上一段中的结论应用于 A^T , 可得
于是 A 的列秩就等于 A 的行秩

由于列秩等于行秩, 所以我们无需使用“列秩”和“行秩”这两个术语, 而用更简单的术语“秩”代替它们

3.58 秩

矩阵 $A \in \mathbf{F}^{m,n}$ 的秩是 A 的列秩

列秩等于行秩的其他证明见于 3.133 和 7A 节的习题 8

3D 可逆性和同构

可逆线性映射

3.59 可逆的、逆

对于线性映射 $T \in \mathcal{L}(V, W)$, 如果存在线性映射 $S \in \mathcal{L}(W, V)$, 使得 ST 等于 V 上的恒等算子且 TS 等于 W 上的恒等算子, 则称 T 是可逆的一个满足 $ST = I$ 及 $TS = I$ 的线性映射 $S \in \mathcal{L}(W, V)$ 被称为 T 的一个逆. (注意, 第一个 I 是 V 上的恒等算子, 第二个 I 是 W 上的恒等算子)

上面的定义中用的说法是线性映射的“一个逆”. 然而, 接下来的结果表明, 我们可以将这个说法换成线性映射的“逆”

3.60 逆唯一

可逆的线性映射具有唯一的逆

证

设 $T \in \mathcal{L}(V, W)$ 是可逆的, 且 S_1 和 S_2 是 T 的逆. 那么

$$S_1 = S_1 I = S_1 (TS_2) = (S_1 T) S_2 = I S_2 = S_2$$

于是 $S_1 = S_2$

3.61 记号

如果 T 是可逆的, 那么它的逆记作 T^{-1} . 换言之, 如果 $T \in \mathcal{L}(V, W)$ 是可逆的, 那么 T^{-1} 是 $\mathcal{L}(W, V)$ 中唯一使得 $T^{-1}T = I$ 和 $TT^{-1} = I$ 成立的元素

3.62 一个从 \mathbb{R}^3 到 \mathbb{R}^3 的线性映射的逆

$T \in \mathcal{L}(\mathbb{R}^3)$ 定义为 $T(x, y, z) = (-y, x, 4z)$. 于是 T 就将向量在 xy 平面上逆时针旋转 90° 并将其沿 z 轴的分量拉伸至原先长度的4倍. 因此, 逆映射 $T^{-1} \in \mathcal{L}(\mathbb{R}^3)$ 就将向量在 xy 平面上顺时针旋转 90° 并将其沿 z 轴的分量压缩至原先长度的 $\frac{1}{4}$ 倍:

$$T^{-1}(x, y, z) = (y, -x, \frac{1}{4}z)$$

3.63 可逆性与单射性和满射性

一个线性映射是可逆的, 当且仅当它既是单射又是满射

证

设 $T \in \mathcal{L}(V, W)$. 我们需要证明, T 是可逆的当且仅当它既是单射又是满射

设 T 是可逆的. 为了证明 T 是单射, 假设 $u, v \in V$ 且 $Tu = Tv$. 那么

$$u = T^{-1}(Tu) = T^{-1}(Tv) = v$$

所以 $u = v$. 所以 T 是单射

设 T 是可逆的, 现在我们欲证 T 是满射. 为此, 令 $w \in W$. 那么 $w = T(T^{-1}w)$, 这就表明 w 在 T 的值域中. 于是 $\text{range } T = W$. 所以 T 是满射, 这就完成了此方向的证明

设 T 既是单射又是满射. 我们欲证 T 是可逆的. 对于每个 $w \in W$, 定义 $S(w)$ 是 V 中唯一使得 $T(S(w)) = w$ 成立的元素 (这样一个元素的存在性和唯一性分别来自 T 的满射性和单射性). 根据 s 的定义可得, $T \circ S$ 就等于 W 上的恒等算子. 为了证明 $S \circ T$ 等于 V 上的恒等算子, 令 $v \in V$. 那么

$$T((S \circ T)v) = (T \circ S)(Tv) = I(Tv) = Tv$$

此式表明 $(S \circ T)v = v$ (因为 T 是单射). 从而 $S \circ T$ 等于 V 上的恒等算子

为完成证明, 我们还需要证明 s 是线性的. 为此, 假设 $w_1, w_2 \in W$. 那么

$$T(S(w_1) + S(w_2)) = T(S(w_1)) + T(S(w_2)) = w_1 + w_2$$

于是 $S(w_1) + S(w_2)$ 是 V 中唯一被 T 映射到 $w_1 + w_2$ 的元素. 根据 s 的定义, 这意味着 $S(w_1 + w_2) = S(w_1) + S(w_2)$. 所以 s 满足可加性, 这是证明其线性所必需的

如果 $w \in W$ 且 $\lambda \in \mathbf{F}$, 那么

$$T(\lambda S(w)) = \lambda T(S(w)) = \lambda w$$

于是 $\lambda S(w)$ 是 V 中唯一被 T 映射到 λw 的元素. 由 s 的定义, 这意味着 $S(\lambda w) = \lambda S(w)$. 所以 s 是线性的, 则原命题得证

你可能会好奇, 对于从某向量空间映射到该空间本身的线性映射, 是否仅凭单射性或满射性就可以推出可逆性. 在无限维向量空间中, 只凭两个条件之一不能推出可逆性. 接下来这个例子, 就用例3.3 中我们所熟知的两个线性映射, 说明了这一点

3.64 仅凭单射性或满射性不能推出可逆性

从 $\mathcal{P}(\mathbf{R})$ 到 $\mathcal{P}(\mathbf{R})$ 的与 x^2 相乘线性映射 (见例3.3) 是单射, 但它不可逆, 因为它不是满射 (多项式1 不在其值域内)

从 \mathbf{F}^∞ 到 \mathbf{F}^∞ 的后向移位线性映射 (见例 3.3) 是满射, 但它不可逆, 因为它不是单射【向量 $(1, 0, 0, 0, \dots)$ 在其零空间中】

鉴于上面例子, 下个结果就非同寻常——它指出, 对于从某有限维向量空间映射到与之维数相同的向量空间的线性映射, 单射性和满射性可以互推. 注意, 对于 V 是有限维的且 $W = V$ 这一特殊情形, 下面的前提条件 $\dim V = \dim W$ 自动满足

3.65 命题

假设 V 和 W 都是有限维向量空间, $\dim V = \dim W$, 且 $T \in \mathcal{L}(V, W)$. 那么 T 可逆 $\iff T$ 是单射 $\iff T$ 是满射

证

线性映射基本定理 (3.21) 表明

$$\dim V = \dim \text{n}T + \dim \text{range}T$$

如果 T 是单射 (由 3.15, 这就等价于 $\dim \text{n}T = 0$), 那么由上式可得

$$\dim \text{range}T = \dim V - \dim \text{n}T = \dim V = \dim W$$

这表明 T 是满射 (根据 2.39)

反之, 如果 T 是满射, 那么式(3.66) 就表明

$\dim \text{n}T = \dim V - \dim \text{range}T = \dim V - \dim W = 0$ 这就表明 T 是单射. 于是, 我们证明了 T 是单射当且仅当 T 是满射. 于是只要 T 满足单射性或满射性

下面的例子说明了上述结论有多么强大. 尽管不用线性代数也可以证明下例中的结论, 但是利用线性代数能让证明更为简洁容易

3.67 存在多项式 p 使得 $((x^2 + 5x + 7)p)'' = q$

从 $\mathcal{P}(\mathbf{R})$ 到该空间本身的线性映射

$$p \mapsto ((x^2 + 5x + 7)p)''$$

是单射, 你可自行证明. 于是我们想用3.65 来证明这个映射是满射. 然而, 例3.64 表明, 不能在无限维向量空间 $\mathcal{P}(\mathbf{R})$ 上发挥 3.65 之妙处. 我们通过将考虑的范围限制在有限维向量空间 $\mathcal{P}_m(\mathbf{R})$ 来绕开这个问题

假设 $q \in \mathcal{P}(\mathbf{R})$. 存在非负整数 m 使得 $q \in \mathcal{P}_m(\mathbf{R})$. 定义 $T: \mathcal{P}_m(\mathbf{R}) \rightarrow \mathcal{P}_m(\mathbf{R})$ 为

$$Tp = ((x^2 + 5x + 7)p)'.$$

将一个非零的多项式乘以 $(x^2 + 5x + 7)$ 会将其次数增加2, 再求两次微分又会将其次数降低2于是 T 的确是从 $\mathcal{P}_m(\mathbf{R})$ 到该空间本身的线性映射

二阶导函数等于0的每个多项式都形如 $ax + b$ ，其中 $a, b \in \mathbf{R}$ 。于是 $\ker T = \{0\}$ ，所以 T 是单射，进而 T 是满射（由3.65），这意味着存在一个多项式 $p \in \mathcal{P}_m(\mathbf{R})$ 使得 $((x^2 + 5x + 7)p)'' = q$ ，正如本例标题所言

6A 节的习题35 给出了类似但是更精彩的3.65 应用实例

下面结论中的前提 $\dim V = \dim W$ ，对于 V 是有限维的且 $W = V$ 这一特殊情形，是自动满足的。那么，在这种情况下，等式 $ST = I$ 蕴涵 $ST = TS$ ，即便乘法交换律并不对于从 V 到 V 的任意线性映射都成立

3.68 $ST = I \implies TS = I$

设 V 和 W 是维数相同的有限维向量空间， $S \in \mathcal{L}(W, V)$ 且 $T \in \mathcal{L}(V, W)$ 。那么 $ST = I$ 当且仅当 $TS = I$

证

设 $ST = I$ 若 $v \in V$ 且 $Tv = 0$ ，那么

$$v = Iv = (ST)v = S(Tv) = S(0) = 0$$

故 T 是单射（由3.15）。因为 V 和 W 维数相同，故可得 T 是可逆的（由 3.65）

现在将 $ST = I$ 的两侧都右乘 T^{-1} ，得 $S = T^{-1}$ 于是 $TS = TT^{-1} = I$

为了证明另一方向的蕴涵关系，只要将上面的证明中 S 和 T 以及 V 和 W 的角色互换，即可证明如果 $TS = I$ 那么 $ST = I$

线性空间的同构

接下来的定义有助于描述除了元素名称不同外本质上相同的两个向量空间

3.69 同构、同构的

同构在线性空间中等价于是可逆线性映射

对于两个向量空间，若存在将其中一个向量空间映成另一个向量空间的同构，则称它们是同构的

可将同构 $T: V \rightarrow W$ 理解为把 $v \in V$ 改写成 $Tv \in W$ 。这个观点能解释为何两个同构的向量空间有相同的向量空间性质。“同构”和“可逆线性映射”这两个术语同义，当你想强调两个空间本质上相同时，就用“同构”这个词

对于两个数学结构（例如群或拓扑空间），要判定它们（除了基础集合 $\mathbf{9}$ 中元素的名称不同外）在本质上相同，可能很困难。然而，接下来的结论说明了，要判定两个向量空间是否同构，我们只需要关注一个数——维数——即可

3.70 维数表明了向量空间是否同构

对于 F 上的两个有限维向量空间，当且仅当它们的维数相同时，它们才是同构的

证

设 V 和 W 是同构的有限维向量空间。于是，存在将 V 映成 W 的同构 T 。因为 T 是可逆的，所以我们有 $\ker T = \{0\}$ 且 $\text{range } T = W$ 。于是

$$\dim V = \dim \ker T + \dim \text{range } T$$

就可以化成 $\dim V = \dim W$ 。这就完成了一个方向的证明

另一方面，设 V 和 W 是维数相同的有限维向量空间。令 v_1, \dots, v_n 是 V 的基， w_1, \dots, w_n 是 W 的基。令 $T \in \mathcal{L}(V, W)$

$$T(c_1 v_1 + \dots + c_n v_n) = c_1 w_1 + \dots + c_n w_n$$

那么由 v_1, \dots, v_n 是 V 的基，所以 T 就是一个定义完善的线性映射。而且，由于 w_1, \dots, w_n 张成 W ，所以 T 是满射；由于 w_1, \dots, w_n 线性无关，可得 $\ker T = \{0\}$ ，所以 T 是单射。因为 T 既是单射又是满射，所以它就是同构（见3.63）。所以 V 和 W 是同构的

上结论表明, 每个有限维向量空间 V 都与 \mathbf{F}^n 同构 (其中 $n = \dim V$). 例如, 如果 m 是非负整数, 那么 $\mathcal{P}_m(\mathbf{F})$ 与 \mathbf{F}^{m+1} 同构. 回顾一下, 记号 $\mathbf{F}^{m,n}$ 代表各元素都属于 \mathbf{F} 的 $m \times n$ 矩阵所构成的向量空间. 如果 ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基, 那么每个 $T \in \mathcal{L}(V, W)$ 都有一对应的矩阵 $(T) \in \mathbf{F}^{m,n}$. 于是, 一旦为 V 和 W 取定了基, T 就成为了一个从 $\mathcal{L}(V, W)$ 到 $\mathbf{F}^{m,n}$ 的函数. 又注意到 3.35 和 3.38 表明 $T \mapsto (T)$ 是一个同构. 既然每个有限维向量空间都与某个 \mathbf{F}^n 同构, 那么为何不只研究 \mathbf{F}^n , 而还要研究更一般的向量空间呢? 要回答这个问题, 就要看到对 \mathbf{F}^n 的研究势必引入其他向量空间. 例如, 我们会碰上线性映射的零空间和值域. 尽管这些向量空间都分别和某个 \mathbf{F}^m 同构, 但这样考虑问题往往更复杂且并不带来新的见解. 线性映射. 现在我们证明, 这个线性映射其实是一个同构.

3.71 定理

设 ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基. 那么 $T \mapsto (T)$ 是 $\mathcal{L}(V, W)$ 与 $\mathbf{F}^{m,n}$ 间的同构.

证

注意到 $T \mapsto (T)$ 是线性的, 只需证 $T \mapsto (T)$ 既是单射又是满射.

我们从证明单射性开始. 如果 $T \in \mathcal{L}(V, W)$ 且 $(T) = 0$, 那么对每个 $k = 1, \dots, n$ 均有 $T\nu_k = 0$. 因为 ν_1, \dots, ν_n 是 V 的基, 故可得 $T = 0$. 于是 $T \mapsto (T)$ 是单射 (由 3.15).

为证 $T \mapsto (T)$ 是满射, 假设 $A \in \mathbf{F}^{m,n}$. 由线性映射引理 (3.4), 存在 $T \in \mathcal{L}(V, W)$ 使得对每个 $k = 1, \dots, n$ 均有

$$T\nu_k = \sum_{j=1}^m A_{j,k} w_j$$

因为 (T) 等于 A , 所以 $T \mapsto (T)$ 的值域等于 $\mathbf{F}^{m,n}$, 满射性得证.

现在我们可以确定两个有限维向量空间之间的线性映射所构成的向量空间的维数了.

3.72 $\dim \mathcal{L}(V, W) = \dim V \dim W$

V 和 W 是有限维的. 那么 $\mathcal{L}(V, W)$ 是有限维的, 且

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$$

证

由 3.71、3.70 和 3.40 可证.

将线性映射视为矩阵乘法

之前我们定义了线性映射的矩阵, 现在我们定义向量的矩阵.

3.73 向量的矩阵

假设 $\nu \in V$ 且 ν_1, \dots, ν_n 是 V 的基. ν 关于该基的矩阵是 $n \times 1$ 矩阵

$$(\nu) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

其中 b_1, \dots, b_n 是使得下式成立的标量:

$$\nu = b_1 \nu_1 + \cdots + b_n \nu_n.$$

向量 $\nu \in V$ 的矩阵 (ν) 取决于 V 的基 ν_1, \dots, ν_n , 也取决于 ν . 然而, 由上下文应该可明确基取什么, 因此它没被包含在记号里.

3.74 向量的矩阵

多项式 $2 - 7x + 5x^3 + x^4$ 关于 $\mathcal{P}_4(\mathbf{R})$ 的标准基的矩阵是

$$\begin{pmatrix} 2 \\ -7 \\ 0 \\ 5 \\ 1 \end{pmatrix}.$$

一向量 $x \in \mathbf{F}^n$ 关于标准基的矩阵可通过将 x 的各坐标写成 $n \times 1$ 矩阵的各元素来获得. 换言之, 如果 $x = (x_1, \dots, x_n) \in \mathbf{F}^n$, 那么

$$(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

有时我们想把 V 中的元素改写成 $n \times 1$ 矩阵. 一旦取定一个基 ν_1, \dots, ν_n , 将 $\nu \in V$ 对应到 (ν) 的函数 就是 (用以实现这种改写的) 将 V 映成 $\mathbf{F}^{n,1}$ 的同构

回忆一下, 如果 A 是一个 $m \times n$ 矩阵, 那么 $A_{\cdot,k}$ 代表 A 的第 k 列, 我们可将它看成一个 $m \times 1$ 矩阵. 下面结论中, $(T\nu_k)$ 是关于 W 的基 w_1, \dots, w_m 来计算的

3.75 定理

设 $T \in \mathcal{L}(V, W)$, ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基. 令 $1 \leq k \leq n$. 那么 (T) 的第 k 列, 记作 $(T)_{\cdot,k}$, 就等于 $(T\nu_k)$

证

由 (T) 和 (ν_k) 的定义可立刻推得上述结论

下面结果展示了线性映射的矩阵、向量的矩阵以及矩阵乘法的概念是怎样融合在一起的

3.76 线性映射的作用就像矩阵乘法

$T \in \mathcal{L}(V, W)$ 且 $\nu \in V$. 假设 ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基. 那么

$$(\nu) = ()(\nu)$$

证

设 $\nu = b_1\nu_1 + \dots + b_n\nu_n$, 其中 $b_1, \dots, b_n \in \mathbf{F}$. 于是

$$T\nu = b_1T\nu_1 + \dots + b_nT\nu_n$$

因此

$$\begin{aligned} (T\nu) &= b_1(T\nu_1) + \dots + b_n(T\nu_n) \\ &= b_1(T)_{\cdot,1} + \dots + b_n(T)_{\cdot,n} \\ &= (T)(\nu) \end{aligned}$$

其中第一个等号来自式(3.77) 和 的线性, 第二个等号来自3.75, 最后一个等号来自3.50

每个 $m \times n$ 矩阵 A 都能诱导出一个从 $\mathbf{F}^{n,1}$ 到 $\mathbf{F}^{m,1}$ 的线性映射, 也就是将 $x \in \mathbf{F}^{n,1}$ 对应至 $Ax \in \mathbf{F}^{m,1}$ 的矩阵乘法函数. 利用上述结论, 通过同构 作适当的改写后, 可以将任何 (从一个有限维向量空间到另一个有限维向量空间的) 线性映射看成矩阵乘法映射. 具体而言, 如果 $T \in \mathcal{L}(V, W)$ 且我们把 $\nu \in V$

与 $(\) \in \mathbf{F}^{n,1}$ 等同起来看, 那么上述结论就告诉我们, 可以把 $T\nu$ 和 $(T)(\nu)$ 等同起来看

因为上述结论使我们得以 (通过同构) 将每个线性映射都看成在 $\mathbf{F}^{n,1}$ 上与某矩阵 A 相乘, 所以要谨记这个矩阵 A 不仅依赖于该线性映射, 还依赖于基的选取. 后续章节中, 许多至关重要的结论的主旨之一, 就是选取使矩阵 A 尽可能简单的基

本书中专注于线性映射而不是矩阵, 然而有时将线性映射看成矩阵 (或将矩阵看成线性映射) 能得出许多重要且有用的见解

下面结论的表述中没出现“基”. (T) 依赖于 V 和 W 的基的选取, 但是该结论表明, 对于所有的选取方式来说, (T) 的列秩总是相同的 (因为 $\text{range } T$ 并不依赖于基的选取)

3.78 $\text{range } T$ 的维数等于 T 的列秩

V 和 W 是有限维的, $T \in \mathcal{L}(V, W)$. 那么 $\dim \text{range } T$ 等于 (T) 的列秩

证

假设 ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基. 将 $w \in W$ 对应到 $(\)$ 的线性映射是一个同构, 它将 W 映成由 $m \times 1$ 向量构成的空间 $\mathbf{F}^{m,1}$. 将该同构限制于 $\text{range } T$ 【由3B节的习题10, 它等于 $\text{span}(T\nu_1, \dots, T\nu_n)$ 】, 该同构就成为将 $\text{range } T$ 映成 $\text{span}((T\nu_1), \dots, (T\nu_n))$ 的同构. 对于每个 $k \in \{1, \dots, n\}$, $m \times 1$ 矩阵 (ν_k) 就等于 (T) 的第 k 列. 原命题得证

换基

3C节中, 我们定义了由 V 到可能与之不同的向量空间 W 的线性映射 T 的矩阵

$$(T, (\nu_1, \dots, \nu_n), (w_1, \dots, w_m))$$

其中 ν_1, \dots, ν_n 是 V 的基且 w_1, \dots, w_m 是 W 的基. 对于从一个向量空间到其本身的线性映射, 我们通常为其定义空间和目标空间选择相同的基. 这时, 我们通常只将所选的基写出一次. 换言之, 如果 $T \in \mathcal{L}(V)$ 且 ν_1, \dots, ν_n 是 V 的基, 那么记号 $(T, (\nu_1, \dots, \nu_n))$ 定义为

$$(T, (\nu_1, \dots, \nu_n)) = (T, (\nu_1, \dots, \nu_n), (\nu_1, \dots, \nu_n))$$

如果从上下文能确知基 ν_1, \dots, ν_n 取什么, 那么我们只写 (T) 即可

3.79 恒等矩阵

设 n 为正整数. 仅对角线上 (即那些行号和列号相等的位置) 的元素为1 而其他各元素均为0 的 $n \times n$ 矩阵

$$\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}$$

称恒等矩阵, 记作 I

上述定义中, 矩阵左下角的0 表示对角线下方的所有元素都是0, 右上角的0 则表示对角线上方的元素都是0

恒等算子 $I \in \mathcal{L}(V)$ 关于 V 的每个基的矩阵都是恒等矩阵 I . 注意, 符号 I 既被用于表示恒等算子, 又被用于表示恒等矩阵. 根据上下文即可知 I 指的是什么意思. 例如, 等式 $\mathcal{M}(\) =$ 中, 左侧的 I 表示恒等算子, 右侧的 I 则表示恒等矩阵

如果 A 是与 I 大小相同的方阵 (行数等于列数的矩阵), 那么 $AI = IA = A$ (你应自行验证)

3.80 可逆的、逆

称方阵 A 是可逆的, 如果存在与之大小相等的方阵 B 使得 $AB = BA = I$. 我们称 B 是 A 的逆且将其记为 A^{-1}

逆唯一

如果 A 是可逆的方阵, 则存在唯一的矩阵 B 使得 $AB = BA = I$ 成立 (由此亦得 $B = A^{-1}$ 这个记号是合理的)

证

用与证明 3.60 相同的方法，可以证明

奇异与非奇异

有些数学家使用术语“非奇异”和“奇异”，它们分别与“可逆”和“不可逆”同义

逆矩阵的性质

如果 A 是可逆矩阵，那么 $(A^{-1})^{-1} = A$ ，因为

$$A^{-1}A = AA^{-1} = I$$

如果 A 和 A^{-1} 是相同大小的可逆方阵，那么 A^{-1} 也可逆，且 $(A^{-1})^{-1} = A$ ，因为

$$\begin{aligned}(A^{-1})(A^{-1})^{-1} &= A^{-1}A \\ &= IA^{-1} \\ &= A^{-1}A \\ &= I\end{aligned}$$

类似地，可得 $A(A^{-1})^{-1} = I$

下面结果成立，是因为我们定义矩阵乘法就是为了使其成立——见 3.43 及其前面的内容现在我们只是把与之有关的基写得更明确些

3.81 线性映射之积的矩阵

$T \in \mathcal{L}(U, V)$ 且 $S \in \mathcal{L}(V, W)$ 。如果 u_1, \dots, u_m 是 U 的基， v_1, \dots, v_n 是 V 的基且 w_1, \dots, w_p 是 W 的基，那么

$$\begin{aligned}(ST, (u_1, \dots, u_m), (w_1, \dots, w_p)) \\ = (S, (v_1, \dots, v_n), (w_1, \dots, w_p))(T, (u_1, \dots, u_m), (v_1, \dots, v_n)).\end{aligned}$$

接下来的结论讨论的是恒等算子 I 关于两个不同的基的矩阵。注意到

$$b(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$$

的第 k 列是由将 u_k 写成基 v_1, \dots, v_n 的线性组合所需的标量构成的
其中 I 表示由 V 到 V 的恒等算子。在证明中， I 也表示 $n \times n$ 恒等矩阵

3.82 恒等算子关于两个基的矩阵

假设 u_1, \dots, u_n 和 v_1, \dots, v_n 是 V 的两个基。那么矩阵

$$(I, (u_1, \dots, u_n), (v_1, \dots, v_n)) \quad \text{和} \quad (I, (v_1, \dots, v_n), (u_1, \dots, u_n))$$

都是可逆的，且互为对方的逆

证

3.81 中，将各 w_k 替换成 u_k 并将 s 和 T 换成 I ，可得

$$I = (I, (\nu_1, \dots, \nu_n), (u_1, \dots, u_n)) (I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n))$$

现在将 u 和 ν 的角色互换，可得

$$I = (I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)) (I, (\nu_1, \dots, \nu_n), (u_1, \dots, u_n))$$

上面两个等式就给出了我们想要的结果

3.83 \mathbf{F}^2 上的恒等算子关于两个基的矩阵

考虑 \mathbf{F}^2 的基 $(4, 2)$, $(5, 3)$ 和 $(1, 0)$, $(0, 1)$ 。因为 $I(4, 2) = 4(1, 0) + 2(0, 1)$ 且 $I(5, 3) = 5(1, 0) + 3(0, 1)$ ，所以我们有

$$(I, ((4, 2), (5, 3)), ((1, 0), (0, 1))) = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$$

你应自行验证，上述矩阵的逆是

$$\begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}$$

于是 3.82 就表明

$$(I, ((1, 0), (0, 1)), ((4, 2), (5, 3))) = \begin{pmatrix} \frac{3}{2} & -\frac{5}{2} \\ -1 & 2 \end{pmatrix}$$

下面的结果展示了当我们改变所选的基时， T 的矩阵是如何改变的。在下面结果中，我们取 V 的两个不同的基，其中每个基都被同时用作 T 的定义空间和目标空间的基。回忆一下，这种情况下，使用下面的简短记号就只需将基写出一次：

$$(T, (u_1, \dots, u_n)) = (T, (u_1, \dots, u_n), (u_1, \dots, u_n))$$

3.84 换基公式

$T \in \mathcal{L}(V)$ ， u_1, \dots, u_n 和 ν_1, \dots, ν_n 都是 V 的基

$$A = (T, (u_1, \dots, u_n)) \quad B = (T, (\nu_1, \dots, \nu_n)) = (I, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)), \text{ 则}$$

$$A = C^{-1} B C$$

证

3.81 中，将各 w_k 替换成 u_k 并将 s 替换为 I ，可得

$$A = {}^{-1}(T, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n))$$

其中用到了 3.82

再利用 3.81，这次将各 w_k 替换成 ν_k ，并将 T 替换为 I ， S 替换为 T ，可得

$$(T, (u_1, \dots, u_n), (\nu_1, \dots, \nu_n)) = B$$

代入 3.85，得 $A = {}^{-1}B$

下面结论的证明留作习题

3.86 逆的矩阵等于矩阵的逆

设 ν_1, \dots, ν_n 是 V 的基且 $T \in \mathcal{L}(V)$ 是可逆的. 那么 $(T^{-1}) = (T)^{-1}$, 式中两个矩阵均是关于基 ν_1, \dots, ν_n 的

3E 向量空间的积和商

向量空间的积

通常与多个向量空间打交道时, 这些向量空间都应在相同的域上

3.87 向量空间的积

V_1, \dots, V_m 都是 \mathbf{F} 上的向量空间, 乘积

$$V_1 \times \cdots \times V_m = \{(\nu_1, \dots, \nu_m) : \nu_1 \in V_1, \dots, \nu_m \in V_m\}$$

$V_1 \times \cdots \times V_m$ 上的加法定义为

$$(u_1, \dots, u_m) + (\nu_1, \dots, \nu_m) = (u_1 + \nu_1, \dots, u_m + \nu_m)$$

$V_1 \times \cdots \times V_m$ 上的标量乘法定义为

$$\lambda(\nu_1, \dots, \nu_m) = (\lambda\nu_1, \dots, \lambda\nu_m)$$

3.88 向量空间 $\mathcal{P}_5(\mathbf{R})$ 和 \mathbf{R}^3 的乘积

$\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ 的元素是长度为2的组, 组中第一项是 $\mathcal{P}_5(\mathbf{R})$ 的元素, 第二项是 \mathbf{R}^3 的元素例如, $(5 - 6x + 4x^2, (3, 8, 7))$ 和 $(x + 9x^5, (2, 2, 2))$ 都是 $\mathcal{P}_5(\mathbf{R}) \times \mathbf{R}^3$ 的元素. 它们的和定义为

$$\begin{aligned} (5 - 6x + 4x^2, (3, 8, 7)) + (x + 9x^5, (2, 2, 2)) \\ = (5 - 5x + 4x^2 + 9x^5, (5, 10, 9)) \end{aligned}$$

此外, $2(5 - 6x + 4x^2, (3, 8, 7)) = (10 - 12x + 8x^2, (6, 16, 14))$

在按3.87 定义的加法和标量乘法运算下, 下面结论成立

3.89 向量空间的积是向量空间

V_1, \dots, V_m 都是 \mathbf{F} 上的向量空间. 那么 $V_1 \times \cdots \times V_m$ 是 \mathbf{F} 上的向量空间

证

证明留给读者

3.90 $\mathbf{R}^2 \times \mathbf{R}^3 \neq \mathbf{R}^5$, 但 $\mathbf{R}^2 \times \mathbf{R}^3$ 与 \mathbf{R}^5 是同构的

注意 $V_1 \times \cdots \times V_m$ 的加法恒等元是 $(0, \dots, 0)$, 其中第 k 个坐标中的0 是 V_k 的加法恒等元. $(\nu_1, \dots, \nu_m) \in V_1 \times \cdots \times V_m$ 的加法逆元是 $(-\nu_1, \dots, -\nu_m)$

向量空间 $\mathbf{R}^2 \times \mathbf{R}^3$ 的元素是这样的组:

$$((x_1, x_2), (x_3, x_4, x_5)),$$

其中 $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$. \mathbf{R}^5 的元素是这样的组:

$$(x_1, x_2, x_3, x_4, x_5),$$

其中 $x_1, x_2, x_3, x_4, x_5 \in \mathbf{R}$

尽管 $\mathbf{R}^2 \times \mathbf{R}^3$ 的元素和 \mathbf{R}^5 的元素看起来很像，它们却不是同类对象： $\mathbf{R}^2 \times \mathbf{R}^3$ 的元素是长度为 2 的组（其中第一项本身就是长度为 2 的组，而第二项是长度为 3 的组），但 \mathbf{R}^5 的元素是长度为 5 的组。因此， $\mathbf{R}^2 \times \mathbf{R}^3$ 不等于 \mathbf{R}^5 。

线性映射

$$\left((x_1, x_2), (x_3, x_4, x_5) \right) \mapsto (x_1, x_2, x_3, x_4, x_5)$$

是将向量空间 $\mathbf{R}^2 \times \mathbf{R}^3$ 映成向量空间 \mathbf{R}^5 的这个同构非常自然，我们把它看成改写即可。有些人通俗地说“ $\mathbf{R}^2 \times \mathbf{R}^3$ 等于 \mathbf{R}^5 ”，尽管严格来讲这说法并不正确，但它体现了通过改写将两个向量空间等同起来看的思想。于是，这两个向量空间是同构的，尽管它们并不相等。下面例子形象地展示了我们将在 3.92 的证明中使用的思想。



3.91 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 的一个基

考虑 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 中这个长度为 5 的向量组：

$$(1, (0, 0)), (x, (0, 0)), (x^2, (0, 0)), (0, (1, 0)), (0, (0, 1))$$

上述组是线性无关的，且张成 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 。所以它是 $\mathcal{P}_2(\mathbf{R}) \times \mathbf{R}^2$ 的一个基。



3.92 向量空间之积的维数是各向量空间维数之和

V_1, \dots, V_m 都是有限维向量空间。那么 $V_1 \times \dots \times V_m$ 是有限维的，且

$$\dim(V_1 \times \dots \times V_m) = \dim V_1 + \dots + \dim V_m$$

证

为每个 V_k 选定一个基。对于每个 V_k 的每个基向量，考虑 $V_1 \times \dots \times V_m$ 中这样一个元素：它的第 k 个坐标等于这个基向量，其他坐标都等于 0。所有这样的元素组成的向量组，是线性无关的且张成 $V_1 \times \dots \times V_m$ ，从而该组是 $V_1 \times \dots \times V_m$ 的一个基。这个基的长度是 $\dim V_1 + \dots + \dim V_m$ ，原命题得证。

在下面结果中，由 $V_1 + \dots + V_m$ 的定义可得映射 是满射。于是下述结论的最后一个词可以由“单射”替换成“可逆的”。



3.93 积与直和

设 V_1, \dots, V_m 都是 V 的子空间。由下式定义线性映射 $\phi: V_1 \times \dots \times V_m \rightarrow V_1 + \dots + V_m$ ：

$$(\nu_1, \dots, \nu_m) \mapsto \nu_1 + \dots + \nu_m$$

那么 $V_1 + \dots + V_m$ 是直和，当且仅当 ϕ 是单射。

证

由 3.15 可得， ϕ 是单射，当且仅当以 $\nu_1 + \dots + \nu_m$ （式中各向量 ν_k 属于 V_k ）表达 0 的唯一方式是将其中的每个向量都取 0。于是 1.45 表明， ϕ 是单射，当且仅当 $V_1 + \dots + V_m$ 是直和，此即待证命题。



3.94 向量空间的和是直和，当且仅当该和的维数等于各求和项维数之和

设 V 是有限维的， V_1, \dots, V_m 都是 V 的子空间。那么 $V_1 + \dots + V_m$ 是直和，当且仅当

$$\dim(V_1 + \dots + V_m) = \dim V_1 + \dots + \dim V_m.$$

证

3.93 中的映射 φ 是满射. 于是根据线性映射基本定理 (3.21), φ 是单射, 当且仅当

$$\dim(V_1 + \cdots + V_m) = \dim(V_1 \times \cdots \times V_m).$$

将 3.93 和 3.92 相结合, 就证明了 $V_1 + \cdots + V_m$ 是直和, 当且仅当

$$\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m,$$

原命题得证

在 $m = 2$ 这个特殊情形下, “ $V_1 + V_2$ 是直和当且仅当 $\dim(V_1 + V_2) = \dim V_1 + \dim V_2$ ”这个命题还可以通过将 1.46 与 2.43 相结合来证明

商空间

我们从定义向量与子集之和开启迈向商空间的第一步

3.95 记号

设 $v \in V$ 且 $U \subseteq V$

$$v + U = \{v + u : u \in U\}$$

3.96 一向量与 \mathbf{R}^2 的一个一维子空间之和

设

$$U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$$

所以 U 是 \mathbf{R}^2 中过原点且斜率为 2 的直线, 故

$$(17, 20) + U$$

找不到“images/fa7756e0a927bc23b8311158f5d51aab7a14881fc9be99edec0840093b57d770.jpg”。

是 \mathbf{R}^2 中过点 $(17, 20)$ 且斜率为 2 的直线
因为

$$(10, 20) \in U \quad (17, 20) \in (17, 20) + U,$$

我们可知 $(17, 20) + U$ 是由 U 向右平移 7 个单位所得到
 $(17, 20) + U$ 平行于子空间 U

3.97 平移

对于 $v \in V$ 和 V 的一个子集 U , 称集合 $v + U$ 是 U 的一个平移

3.98 平移

如果 U 是 \mathbf{R}^2 中定义为 $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$ 的直线, 那么 \mathbf{R}^2 中所有斜率为 2 的直线都是 U 的平移. U 和其一个平移的图形如上面的例 3.96 所示

更一般地说, 如果 U 是 \mathbf{R}^2 中的直线, 那么 U 的所有平移构成的集合就是 \mathbf{R}^2 中所有与 U 平行的直线构成的集合

如果 $U = \{(x, y, 0) \in \mathbf{R}^3 : x, y \in \mathbf{R}\}$ ，那么 U 的平移是 \mathbf{R}^3 中所有与 xy 平面平行的平面
更一般地说，如果 U 是 \mathbf{R}^3 中的平面，那么 U 的所有平移构成的集合就是 \mathbf{R}^3 中所有与 U 平行的平面构成的集合（实例见本节习题7）

3.99 商空间

设 U 是 V 的子空间. 那么商空间 V/U 是由 U 的所有平移构成的集合

$$V/U = \{\nu + U : \nu \in V\}.$$

3.100 商空间

如果 $U = \{(x, 2x) \in \mathbf{R}^2 : x \in \mathbf{R}\}$ ，那么 \mathbf{R}^2/U 是 \mathbf{R}^2 中所有斜率为2的直线所构成的集合如果 U 是 \mathbf{R}^3 中过原点的直线，那么 \mathbf{R}^3/U 是 \mathbf{R}^3 中所有与 U 平行的直线所构成的集合如果 U 是 \mathbf{R}^3 中过原点的平面，那么 \mathbf{R}^3/U 是 \mathbf{R}^3 中所有与 U 平行的平面所构成的集合

我们接下来的目标，是让 V/U 成为向量空间. 为此，我们需要下面这个结果

3.101 子空间的两个平移要么相等要么不相交

设 U 是 V 的一个子空间且 $\nu, w \in V$. 那么

$$\nu - w \in U \iff \nu + U = w + U \iff (\nu + U) \cap (w + U) \neq \emptyset$$

证

先设 $\nu - w \in U$. 如果 $u \in U$ ，那么

$$\nu + u = w + ((\nu - w) + u) \in w + U.$$

从而 $\nu + U \subseteq w + U$. 类似地， $w + U \subseteq \nu + U$. 于是 $\nu + U = w + U$ ，也就证明了 $\nu - w \in U$ 蕴涵 $\nu + U = w + U$

等式 $\nu + U = w + U$ 蕴涵 $(\nu + U) \cap (w + U) \neq \emptyset$

现在设 $(\nu + U) \cap (w + U) \neq \emptyset$. 于是存在 $u_1, u_2 \in U$ 使得

$$\nu + u_1 = w + u_2$$

于是 $\nu - w = u_2 - u_1$. 所以 $\nu - w \in U$ ，这表明 $(\nu + U) \cap (w + U) \neq \emptyset$ 蕴涵 $\nu - w \in U$ ，证毕

现在我们可以定义 V/U 上的加法和标量乘法了

3.102 V/U 上的加法和标量乘法

U 是 V 的一个子空间. 那么 V/U 上的加法和标量乘法分别由下面两式定义：对所有 $\nu, w \in V$ 和所有 $\lambda \in \mathbf{F}$ ，

$$(\nu + U) + (w + U) = (\nu + w) + U$$

$$\lambda(\nu + U) = (\lambda\nu) + U$$

以上定义是合理的——对这一点的证明，是对接下来这条结果的证明中的一部分

3.103 商空间是向量空间

U 是 V 的一个子空间. 那么带有定义如上的加法和标量乘法的 V/U 就是一个向量空间

证

上述 V/U 上加法和标量乘法的定义可能存在一个问题，就是 U 的同一个平移存在不唯一的表达法。更具体地说，为了证明上面给出的 V/U 上的加法定义是合理的，假设有向量 $\nu_1, \nu_2, w_1, w_2 \in V$ 满足

我们必须证明 $(\nu_1 + w_1) + U = (\nu_2 + w_2) + U$ 。

由 3.101

因为 U 是 V 的子空间，从而对加法封闭，所以此式表明 $(\nu_1 - \nu_2) + (w_1 - w_2) \in U$ 。于是 $(\nu_1 + w_1) - (\nu_2 + w_2) \in U$ 。再次利用 3.101 可见

$$(\nu_1 + w_1) + U = (\nu_2 + w_2) + U$$

即我们欲证的结论成立。于是 V/U 上加法的定义是合理的

类似地，假设 $\lambda \in \mathbf{F}$ 。仍假设 $\nu_1 + U = \nu_2 + U$ 。因为 U 是 V 的子空间，从而对标量乘法封闭，所以我们有 $\lambda(\nu_1 - \nu_2) \in U$ 。于是 $\lambda\nu_1 - \lambda\nu_2 \in U$ 。所以 3.101 表明

$$(\lambda\nu_1) + U = (\lambda\nu_2) + U.$$

于是 V/U 上加法和标量乘法的定义都是合理的

既然 V/U 上的加法和标量乘法运算已定义清楚，那么验证这些运算使 V/U 成为向量空间就很简单了，请读者自行完成。注意， V/U 的加法恒等元是 $0 + U$ （也就等于 U ）， $\nu + U$ 的加法逆元是 $(-\nu) + U$

接下来的概念将引出 V/U 的维数计算方法

3.104 商映射

U 是 V 的一个子空间。商映射 $\pi: V \rightarrow V/U$ 是由下式定义的线性映射：对每个 $\nu \in V$ ，

$$\pi(\nu) = \nu + U$$

读者应自证 π 的确是一个线性映射。尽管 π 既依赖于 U 又依赖于 V ，记号中却省略了这些向量空间，因为根据上下文就能明确它们是什么

3.105 商空间的维数

V 是有限维的， U 是 V 的子空间。则

$$\dim V/U = \dim V - \dim U$$

证

令 π 表示从 V 到 V/U 的线性映射。如果 $\nu \in V$ ，那么 $\nu + U = 0 + U$ 当且仅当 $\nu \in U$ （由 3.101），这表明 $\pi^{-1}(0) = U$ 。的定义表明 $\text{range } \pi = V/U$ 。现在，线性映射基本定理（3.21）表明 $\dim V = \dim U + \dim V/U$ ，这正是我们想要的结果

V 上的每个线性映射 T 都能在 $V/\pi^{-1}(0)$ 上诱导出一个线性映射 \bar{T} 。我们现在给出其定义

3.106 记号

设 $T \in \mathcal{L}(V, W)$ 。 $\bar{T}: V/\pi^{-1}(0) \rightarrow W$ 由下式定义：

$$\bar{T}(\nu + \pi^{-1}(0)) = T\nu$$

良定义证

为了说明上述 T 的定义是合理的, 假设 $u, v \in V$ 满足 $u + nT = v + nT$. 由 3.101, 我们有 $u - v \in nT$. 于是 $T(u - v) = 0$. 所以 $Tu = Tv$. 于是 T 的定义确实是合理的. 验证 T 是一个从 V/nT 到 W 的线性映射的过程很常规, 留给读者完成

下面的结果说明, 我们可以将 T 看成 T 的修改版——换了定义空间, 从而变为单射

3.107 T 的零空间和值域

设 $T \in \mathcal{L}(V, W)$, 则

(a) $T \circ T = T$, 其中 T 是将 V 映成 $V/(nT)$ 的商映射 (b) T 是单射

(c) $\text{range } T = \text{ange } T$

(d) $V/(nT)$ 和 $\text{range } T$ 是同构的向量空间

证

(a) 如果 $v \in V$, 那么 $(T \circ T)(v) = T((v)) = T(v + nT) = Tv$, 命题得证;

(b) 设 $v \in V$ 且 $T(v + nT) = 0$. 那么 $Tv = 0$. 于是 $v \in nT$. 所以 3.101 表明 $v + nT = 0 + nT$. 由此推出 $nT = \{0 + nT\}$. 所以 T 是单射, 命题得证;

(c) T 的定义即表明 $\text{range } T = \text{ange } T$;

(d) 现在由 (b) 和 (c) 可得, 如果我们将 T 视作映射到 $\text{range } T$, 那么 T 就是将 $V/(nT)$ 映成 $\text{range } T$ 的同构

3F 对偶

对偶空间和对偶映射

映射到标量域 \mathbf{F} 中的线性映射在线性代数中扮演着特殊的角色, 于是, 我们赋予它们特殊的名称

3.108 线性泛函

V 上的线性泛函是从 V 到 \mathbf{F} 的线性映射. 换言之, 线性泛函是 $\mathcal{L}(V, \mathbf{F})$ 的元素

3.109 线性泛函

定义: \mathbf{R}^3 为 $(x, y, z) = 4x - 5y + 2z$. 那么 \mathbf{R}^3 上的线性泛函

固定 $(c_1, \dots, c_n) \in \mathbf{F}^n$. 定义: $\mathbf{F}^n \rightarrow \mathbf{F}$ 为 $(x_1, \dots, x_n) = c_1x_1 + \dots + c_nx_n$. 那么 \mathbf{F}^n 上的线性泛函

定义: $\mathcal{P}(\mathbf{R})$ 如下:

$$(p) = 3p''(5) + 7p(4)$$

那么 $\mathcal{P}(\mathbf{R})$ 上的线性泛函

定义: $\mathcal{P}(\mathbf{R})$ 如下: 对每个 $p \in \mathcal{P}(\mathbf{R})$,

$$(p) = \int_0^1 p$$

那么 $\mathcal{P}(\mathbf{R})$ 上的线性泛函

3.110 对偶空间

V 的对偶空间记作 V' , 是 V 上全体线性泛函所构成的向量空间. 换言之, $V' = \mathcal{L}(V, \mathbf{F})$

3.111 $\dim V' = \dim V$

V 是有限维的. 那么 V' 也是有限维的, 且 $\dim V' = \dim V$

证

由 3.72 $\dim V' = \dim \mathcal{L}(V, \mathbf{F}) = (\dim V)(\dim \mathbf{F}) = \dim V$

下面定义中, 3.4表明每个 j 都具有明确的定义

定义3.112 对偶基

ν_1, \dots, ν_n 是 V 的基, 那么 ν_1, \dots, ν_n 的对偶基是 V' 中的元素 j_1, \dots, j_n 所构成的组, 其中各 j 是 V 上满足下式的线性泛函:

$$j(\nu_k) = \begin{cases} 1 \\ 0 \end{cases}$$

3.113 \mathbf{F}^n 的标准基的对偶基

n 是正整数, 对 $1 \leq j \leq n$, 将 j 定义为将 \mathbf{F}^n 中的向量对应至它的第 k 个坐标的 \mathbf{F}^n 上的线性泛函. 于是对于各 $(x_1, \dots, x_n) \in \mathbf{F}^n$,

$$j(x_1, \dots, x_n) = x_j$$

证

令 ν_1, \dots, ν_n 是 \mathbf{F}^n 的标准基. 那么

$$j(k) = \begin{cases} 1 & k = j \\ 0 & k \neq j \end{cases}$$

于是 j_1, \dots, j_n 就是 \mathbf{F}^n 的标准基 ν_1, \dots, ν_n 的对偶基

下面结论说明了, 对于 V 的基而言, 组成其对偶基的线性泛函, 给出了用基向量的线性组合来表示 V 中向量时所取的系数

3.114 对偶基给出了线性组合的系数

ν_1, \dots, ν_n 是 V 的基, 且 j_1, \dots, j_n 是其对偶基. 那么对每个 $\nu \in V$, 有

$$\nu = j_1(\nu)\nu_1 + \dots + j_n(\nu)\nu_n$$

证

设 $\nu \in V$. 那么存在 $c_1, \dots, c_n \in \mathbf{F}$ 使得 $\nu = c_1\nu_1 + \dots + c_n\nu_n$

如果 $j \in \{1, \dots, n\}$, 那么在等式两端同时作用 j 可得 $j(\nu) = c_j$

将由上式给出的 c_1, \dots, c_n 表达式代入式(3.115)得 $\nu = j_1(\nu)\nu_1 + \dots + j_n(\nu)\nu_n$

接下来的结论表明对偶基的确是对偶空间的基. 于是使用术语“对偶基”就是恰当的

3.116 对偶基是对偶空间的基

V 是有限维的. 那么 V 的基的对偶基是 V' 的基

证

设 ν_1, \dots, ν_n 是 V 的基. 令 ν_1, \dots, ν_n 表示其对偶基

为证明 ν_1, \dots, ν_n 是由 V' 中元素构成的线性无关组, 设 $a_1, \dots, a_n \in \mathbf{F}$ 满足

$$a_{11} + \dots + a_{nn} = 0$$

而对各 $k \in 1, \dots, n$, 有

$$(a_{11} + \dots + a_{nn})(\nu_k) = a_k$$

于是式 (3.117) 表明 $a_1 = \dots = a_n = 0$. 所以 ν_1, \dots, ν_n 是线性无关的

因为 ν_1, \dots, ν_n 是 V' 中的线性无关组且其长度等于 $\dim V'$ (由 3.111), 我们可得出结论 ν_1, \dots, ν_n 是 V' 的一个基 (见 2.38)

在接下来的定义中, 注意如果 T 是从 V 到 W 的线性映射, 那么 T' 是从 W' 到 V' 的线性映射

3.118 对偶映射

$T \in \mathcal{L}(V, W)$, T 的对偶映射是由下式定义的线性映射 $T' \in \mathcal{L}(W', V')$: 对每个 $\phi \in W'$

$$T'(\phi) = \phi \circ T$$

$T \in \mathcal{L}(V, W)$ 且 $\phi \in W'$, 那么 $T'(\phi)$ 就是线性映射 ϕ 和 T 的复合. 于是 $T'(\phi)$ 的确是一个从 V 到 \mathbf{F} 的线性映射; 换言之, $T'(\phi) \in V'$

有些书籍中用于表示对偶的记号是 V^* 和 T^* 而不是 V' 和 T' . 但我们将 T^* 这个记号预留给伴随. 到了第7章, 我们学习内积空间上的线性映射时, 将引入伴随这个概念

T' 是从 W' 到 V' 的线性映射

若 $\phi, \psi \in W'$, 那么

$$T'(\phi + \psi) = (\phi + \psi) \circ T = \phi \circ T + \psi \circ T = T'(\phi) + T'(\psi)$$

若 $\lambda \in \mathbf{F}$ 且 $\phi \in W'$,

$$T'(\lambda\phi) = (\lambda\phi) \circ T = \lambda(\phi \circ T) = \lambda T'(\phi)$$

下例中撇号 (') 有两种不相关的意思: 在 D' 中表示线性映射 D 的对偶, 在 p' 中则表示多项式 p 的导数

3.119 微分线性映射的对偶映射

$D: \mathcal{P}(\mathbf{R}) \rightarrow \mathcal{P}(\mathbf{R})$, $Dp = p'$

设 ϕ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函, 定义为 $\phi(p) = p(3)$. 那么 $D'(\phi)$ 是由下式定义的 $\mathcal{P}(\mathbf{R})$ 上的线性泛函

$$(D'(\phi))(p) = (\phi \circ D)(p) = (Dp)(3) = (p')(3)$$

于是 $D'(\phi)$ 是 $\mathcal{P}(\mathbf{R})$ 上将 p 对应到 $p'(3)$ 的线性泛函

设 ψ 是 $\mathcal{P}(\mathbf{R})$ 上的线性泛函, 定义为 $\psi(p) = \int_0^1 p$. 那么 $D'(\psi)$ 是由下式定义的 $\mathcal{P}(\mathbf{R})$ 上的线性泛函

$$\begin{aligned} (D'(\psi))(p) &= (\psi \circ D)(p) \\ &= (Dp) \\ &= (p') \\ &= \int_0^1 p' \\ &= p(1) - p(0) \end{aligned}$$

于是 $D'(\psi)$ 是 $\mathcal{P}(\mathbf{R})$ 上将 p 对应到 $p(1) - p(0)$ 的线性泛函

在下面结果中, (a) 和 (b) 表明将 T 对应到 T' 的函数是一个从 $\mathcal{L}(V, W)$ 到 $\mathcal{L}(W', V')$ 的线性映射
在下述(c) 中, 注意从等式左侧的 ST 到右侧的 $T'S'$, s 和 T 的顺序是相反的

3.120 对偶映射的代数性质

$T \in \mathcal{L}(V, W)$

(a) $\forall S \in \mathcal{L}(V, W), (S + T)' = S' + T'$

(b) $\forall \lambda \in \mathbf{F}, (\lambda T)' = \lambda T'$

(c) $\forall S \in \mathcal{L}(W, U), (ST)' = T'S'$

证

(a) 和(b) 证明留给读者

(c) 设 $u \in U'$

$$(ST)'(u) = u \circ (ST) = (u \circ S) \circ T = T'(u \circ S) = T'(S'(u)) = (T'S')(u)$$

其中第1、3、4 个等号成立是因为对偶映射的定义, 第2 个等号成立是因为函数的复合满足结合律, 而最后一个等号是由复合的定义得来

上式表明对于所有 $u \in U'$, 均有 $(ST)'(u) = (T'S')(u)$. 于是 $(ST)' = T'S'$

线性映射的对偶的零空间和值域

在本小节中, 我们的目标是用 $\text{range } T$ 和 $\text{null } T$ 来刻画 $\text{null } T'$ 和 $\text{range } T'$. 为达成这个目标, 我们需要给出下面这个定义

3.121 零化子

对 $U \subseteq V$, U 的零化子, 记作 U^0 , 定义为

3.122 一个零化子的元素

设 U 是 $\mathcal{P}(\mathbf{R})$ 的由 x^2 的多项式倍 11 所构成的子空间. 如果 f 是 $\mathcal{P}(\mathbf{R})$ 上定义为 $(f)(p) = p'(0)$ 的线性泛函, 那么 $f \in U^0$

对 $U \subseteq V$, 零化子 U^0 是对偶空间 V' 的一个子集. 于是 U^0 依赖于包含 U 的那个向量空间 V , 所以形如 U_V^0 的记号会更准确些. 然而, 根据上下文我们总能明确包含 U 的那个空间是什么, 所以我们常用 U^0 这个更简单的记号

证

令 e_1, e_2, e_3, e_4, e_5 表示 \mathbf{R}^5 的标准基, 并令 $e'_1, e'_2, e'_3, e'_4, e'_5 \in (\mathbf{R}^5)'$ 表示 e_1, e_2, e_3, e_4, e_5 的对偶基. 设

$$U = \text{span}(e_1, e_2) = \{(x_1, x_2, 0, 0, 0) \in \mathbf{R}^5 : x_1, x_2 \in \mathbf{R}\}$$

我们想要证明 $U^0 = \text{span}(e'_3, e'_4, e'_5)$

回忆一下 (参见例 3.113), e'_j 是 \mathbf{R}^5 上将向量对应到其第 j 个坐标的线性泛函: $e'_j(x_1, x_2, x_3, x_4, x_5) = x_j$.

先设 $f \in \text{span}(e'_3, e'_4, e'_5)$. 那么存在 $c_3, c_4, c_5 \in \mathbf{R}$ 使得 $f = c_3 e'_3 + c_4 e'_4 + c_5 e'_5$ 成立. 如果 $(x_1, x_2, 0, 0, 0) \in U$, 那么

$$(x_1, x_2, 0, 0, 0) = (c_3 e'_3 + c_4 e'_4 + c_5 e'_5)(x_1, x_2, 0, 0, 0) = 0$$

于是 $f \in U^0$. 由此我们证明了 $\text{span}(e'_3, e'_4, e'_5) \subseteq U^0$

为了证明这个包含关系反过来也成立, 设 $f \in U^0$. 因为对偶基是 $(\mathbf{R}^5)'$ 的基, 所以存在 $c_1, c_2, c_3, c_4, c_5 \in \mathbf{R}$ 使得 $f = c_1 e'_1 + c_2 e'_2 + c_3 e'_3 + c_4 e'_4 + c_5 e'_5$ 成立. 因为 $e_1 \in U$ 且 $f \in U^0$, 所以我们有

$$0 = f(e_1) = (c_1 e'_1 + c_2 e'_2 + c_3 e'_3 + c_4 e'_4 + c_5 e'_5)(e_1) = c_1.$$

类似地, $c_2 \in U$ 从而 $c_2 = 0$. 所以 $c = c_{33} + c_{44} + c_{55}$. 于是 $c \in \text{span}(3, 4, 5)$, 说明 $U^0 \subseteq \text{span}(3, 4, 5)$. 于是, $U^0 = \text{span}(3, 4, 5)$

3.124 零化子是子空间

设 $U \subseteq V$. 那么 U^0 是 V' 的子空间

证

注意到 $0 \in U^0$ (此处 0 是 V 上的零线性泛函), 因为将零线性泛函作用于 U 中每个向量都会得到 $0 \in \mathbf{F}$

设, $\alpha \in U^0$. 于是, $\alpha \in V'$ 且对每个 $u \in U$ 都有 $(\alpha)(u) = (\alpha)(u) = 0$. 如果 $u \in U$, 那么

$$(\alpha + \beta)(u) = (\alpha)(u) + (\beta)(u) = 0 + 0 = 0.$$

于是 $\alpha + \beta \in U^0$

类似可证 U^0 对于标量乘法封闭. 于是 1.34 表明 U^0 是 V' 的一个子空间

下面结果说明 $\dim U^0 = \dim V - \dim U$. 例如, 若 U 是 \mathbf{R}^5 的一个二维子空间, 那么 U^0 就是 $(\mathbf{R}^5)'$ 的一个三维子空间——见例 3.123. 下面的结论可以遵循例 3.123 的思考方式来证明: 选取 U 的基 u_1, \dots, u_m , 将其扩充成 V 的基 $u_1, \dots, u_m, \dots, u_n$, 令 $\alpha_1, \dots, \alpha_m, \dots, \alpha_n$ 为 V' 的对偶基, 然后证明 $\alpha_{m+1}, \dots, \alpha_n$ 是 U^0 的基, 即可得到欲证的结论. 即便下面给出了一个更取巧的证明, 你还是应该按上面概述的思路自行构造一个证明

3.125 零化子的维数

设 V 是有限维的且 U 是 V 的一个子空间. 那么

$$\dim U^0 = \dim V - \dim U$$

证

令 $i \in \mathcal{L}(U, V)$ 是包含映射 (inclusion map) 12, 定义为 $i(u) = u$ (对于任一 $u \in U$ 均成立). 于是 i' 是从 V' 到 U' 的线性映射. 将线性映射基本定理 (3.21) 应用于 i' 可得

$$\dim \text{range } i' + \dim \text{null } i' = \dim V'$$

而 $\text{null } i' = U^0$ (思考定义即可想出此式) 且 $\dim V' = \dim V$ (由 3.111), 所以我们可以将上述等式改写成

$$\dim \text{range } i' + \dim U^0 = \dim V.$$

如果 $\alpha \in U'$, 那么 α 可以被扩充为 V 上的一个线性泛函 (实例见于 3A 节习题 13). i' 的定义表明 $i'(\alpha) = \alpha$. 由此 $\alpha \in \text{range } i'$, 这表明 $\text{range } i' = U'$, 所以

$$\dim \text{range } i' = \dim U' = \dim U,$$

那么式 (3.126) 就可改写成 $\dim U + \dim U^0 = \dim V$, 原命题得证

下面结果可作一个有力工具, 用以证明某个子空间是最大子空间【见 (a)】或是最小子空间【见 (b)】

3.127 零化子等于 0 或整个空间的条件

V 是有限维的, U 是 V 的一个子空间. 那么
 (a) $U^0 = \{0\} \iff U = V$; (b) $U^0 = V' \iff U = \{0\}$

证

为证明(a), 我们有

$$\begin{aligned} U^0 = \{0\} &\iff \dim U^0 = 0 \\ &\iff \dim U = \dim V \\ &\iff U = V \end{aligned}$$

其中, 第二个等价来自3.125 而第三个等价来自2.39
 类似地, 为证明(b), 我们有

$$\begin{aligned} U^0 = V' \iff \dim U^0 &= \dim V' \\ \dim U^0 &= \dim V \\ \dim U &= 0 \\ U &= \{0\}, \end{aligned}$$

其中, 第一个等价中的一个方向源于2.39, 第二个等价来自3.111, 而第三个等价来自3.125
 下面结果中(a) 的证明无需假设 V 和 W 是有限维的

3.128 T^* 的零空间

设 V 和 W 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么
 (a) $\text{n}T' = (\text{ange}T)^0$; (b) $\dim \text{null } T' = \dim \text{n}T + \dim W - \dim V$.

证

(a) 先设 $\alpha \in \text{n}T'$. 于是 $0 = T'(\alpha) = \alpha \circ T$. 因此

于是 $\alpha \in (\text{ange}T)^0$. 这意味着 $\text{n}T' \subseteq (\text{ange}T)^0$

为了证明这个包含关系反过来也成立, 现设 $\alpha \in (\text{ange}T)^0$. 于是对每个 $\nu \in V$ 都有 $(T\nu) = 0$. 所以 $0 = \alpha \circ T = T'(\alpha)$. 换言之, $\alpha \in \text{n}T'$, 这表明 $(\text{ange}T)^0 \subseteq \text{n}T'$, 这样就完成了(a) 的证明

(b) 我们有

$$\begin{aligned} \dim \text{n}T' &= \dim(\text{ange}T)^0 \\ &= \dim W - \dim \text{ange}T \\ &= \dim W - (\dim V - \dim \text{n}T) \\ &= \dim \text{n}T + \dim W - \dim V, \end{aligned}$$

其中第一个等号源于 (a), 第二个等号源于 3.125, 第三个等号则源于线性映射基本定理 (3.21)

下面结果在某些情况下有用处, 因为有时验证 T' 是单射比直接验证 T 是满射更容易

3.129 T 是满射等价于 T^* 是单射

设 V 和 W 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么 T 是满射 $\iff T'$ 是单射

证

我们有

$$\iff (\text{range } T)^0 = \{0\}$$

$$\iff \text{n}T' = \{0\}$$

其中第二个等价来源于 3.127 (a) 而第三个等价源于 3.128 (a)

3.130 T' 的值域

设 V 和 W 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

(a) $\dim \text{range } T' = \dim \text{range } T$; (b) $\text{range } T' = (\text{n}T)^0$

证

(a) 我们有

$$\begin{aligned} \dim \text{range } T' &= \dim W' - \dim \text{n}T' \\ &= \dim W - \dim (\text{range } T)^0 \\ &= \dim \text{range } T, \end{aligned}$$

其中第一个等号源于 3.21, 第二个等号源于 3.111 和 3.128 (a), 第三个等号则源于 3.125

(b) 先设 $\nu \in \text{range } T'$. 于是存在 $\nu' \in W'$ 使得 $\nu' = T'(\nu)$ 成立. 如果 $\nu \in \text{n}T$, 那么

$$(\nu' = T'(\nu))\nu = (\nu' \circ T)(\nu) = (T\nu) = (0) = 0.$$

因此 $\nu \in (\text{n}T)^0$. 这意味着 $\text{range } T' \subseteq (\text{n}T)^0$

我们将通过说明 $\text{range } T'$ 和 $(\text{n}T)^0$ 的维数相同来完成证明. 为此, 注意到

$$\begin{aligned} \dim \text{range } T' &= \dim \text{range } T \\ &= \dim V - \dim \text{n}T \\ &= \dim (\text{n}T)^0, \end{aligned}$$

其中第一个等号来源于 (a), 第二个等号来源于 3.21, 第三个等号则源于 3.125

下个结果应和 3.129 对照起来看

3.131 T 是单射等价于 T' 是满射

设 V 和 W 是有限维的且 $T \in \mathcal{L}(V, W)$. 那么

T 是单射 $\iff T'$ 是满射

证

我们有

$$\begin{aligned} (\text{n}T)^0 &= V' \\ \text{range } T' &= V' \end{aligned}$$

其中第二个等价来源于 3.127 (b) 而第三个等价源于 3.130 (b)

线性映射的对偶的矩阵

下面结论假设我们有 V 的基 ν_1, \dots, ν_n 及其在 V' 中的对偶基 ν_1, \dots, ν_n ，也有 W 的基 w_1, \dots, w_m 及其在 W' 中的对偶基 ν_1, \dots, ν_m 。于是可按上述 V 和 W 的基计算出 (T) ，按上述 V' 和 W' 中的对偶基计算出 (T') 。利用这些基可以得出

3.132 T' 的矩阵是 T 的矩阵的转置

V 和 W 是有限维的， $T \in \mathcal{L}(V, W)$ ，则

$$(T') = ((T))$$

证

令 $A = (T)$ 且 $B = (T')$ 。设 $1 \leq j \leq m$ 且 $1 \leq k \leq n$
由 (T') 的定义

$$T'(j) = \sum_{i=1}^n \nu_i(j) \nu_i$$

上式左侧等于 $j \circ T$ 。于是将上式两侧均作用于 ν_k 上，可得

$$\begin{aligned} (j \circ T)(\nu_k) &= \sum_{i=1}^n \nu_i(j) (T\nu_k) \\ &= \sum_{i=1}^n A_{i,j} \nu_i \end{aligned}$$

还可写出

$$\begin{aligned} (j \circ T)(\nu_k) &= j(T\nu_k) \\ &= j \left(\sum_{i=1}^m A_{i,k} w_i \right) \\ &= \sum_{i=1}^m A_{i,k} j(w_i) \\ &= A_{j,k} \end{aligned}$$

比照上面两组等式的最后一行，我们有 $A_{k,j} = A_{j,k}$ 。于是 $B = A$ 。换言之， $(T') = ((T))$
现在，我们利用对偶给出矩阵的列秩等于行秩的另一种证法。我们之前用别的工具证明了该结论——见3.57

3.133 列秩等于行秩

设 $A \in \mathbf{F}^{m,n}$ 。那么 A 的列秩等于 A 的行秩

证

定义 $T: \mathbf{F}^{n,1} \rightarrow \mathbf{F}^{m,1}$ 为 $Tx = Ax$ 。于是 $(T) = A$ ，其中 (T) 是关于 $\mathbf{F}^{n,1}$ 和 $\mathbf{F}^{m,1}$ 的标准基计算得到的。这样一来
其中第二个等号源于3.78，第三个等号源于3.130 (a)，第四个等号源于3.78，第五个等号源于3.132，最后一个等号则源于列秩和行秩的定义

上述结果的另一种证明方法见于7A 节的习题8