第一章 线性空间与线性映射

1.1 线性空间

定义 1.1.1 (线性空间)
给定非空集合V及数域$\mathbb{F}$. 若有映射
σ: $V\times V\to V$
$(v_1,v_2)\mapsto\sigma(v_1,v_2)$
称为V上的加法，并记$\sigma(v_1,v_2)=v_1+v_2$；及映射
τ: $V\times\mathbb{F}\to V$
$(v,k)\mapsto\tau(v,k)$
称为V和$\mathbb{F}$之间的数乘法，并记$\tau(v,k)=v\cdot k$，且这两运算满足"通常的运算规则"，则称V关于此$+$和$\cdot$是$\mathbb{F}$上的线性空间. 在无混淆时，也简称V是线性空间. 表示数乘法的记号"$\cdot$"有时也省略. 线性空间中的元素也称为向量，线性空间也称为向量空间.

注 1.1.1 (数域$\mathbb{F}$)
这里可暂时理解为全体复数的集合$\mathbb{C}$的非空的，对加减乘除四则运算封闭的子集. 例如有理数集$\mathbb{Q}$，实数集$\mathbb{R}$，及复数集$\mathbb{C}$自身. 然而，更一般的域也是有的，并且讨论一般的域上的线性空间也是有实际需要的. 本书第二章将要涉及到全体有理分式构成的域.

注 1.1.2 (集合的积)
给定集合S_1及S_2，定义它们的积
$S_1\times S_2=\{(s_1,s_2):s_1\in S_1,\text{且 }s_2\in S_2\}$
或
$$S_1\times S_2=\left\{\begin{bmatrix}s_1\\s_2\end{bmatrix}:s_1\in S_1,\text{且 }s_2\in S_2\right\}.$$
将一个偶(有序对)写成行(s_1,s_2)，或写成列
$$\begin{bmatrix}s_1\\s_2\end{bmatrix}$$
依赖于约定，或有时只是为了节省书写篇幅. 这里应注意顺序：$(s_1,s_2)\neq(s_2,s_1)$. 类似地，还可定义多个集合的积. 特别地，我们有
$$\mathbb{F}^n=\underbrace{\mathbb{F}\times\mathbb{F}\times\cdots\times\mathbb{F}}_{n\text{个}}=\left\{\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}:x_i\in\mathbb{F},i=1,2,\ldots,n\right\}.(1.1.1)$$

注 1.1.3 (表示映射的带尾与不带尾的箭头的区别)
从集合S_1到集合S_2的映射f这一事实用不带尾的箭头表示，即
$f:S_1\to S_2$;
而映射f把集合S_1中的元素$a\in S_1$映为集合S_2中的元素$b\in S_2$这一事实，用带尾的箭头表示，即
$f:a\mapsto b$.
这种记号约定在数学文献中已很普遍.
另外，要习惯于将运算视为映射的观点. 例如，整数集合$\mathbb{Z}$上的加法运算+就决定了如下映射
σ: $\mathbb{Z}\times\mathbb{Z}\to\mathbb{Z}$
$(s,t)\mapsto s+t,$
如：$\sigma:(1,2)\mapsto1+2$，即$\sigma(1,2)=3$.

注 1.1.4 ("通常的运算规则")

这里指的是加法满足

(1). 交换律: $(\forall v_1, v_2 \in V)(v_1 + v_2 = v_2 + v_1)$.

(2). 结合律: $(\forall v_1, v_2, v_3 \in V)((v_1 + v_2) + v_3 = v_1 + (v_2 + v_3))$.

(3). 有零元: V 中有零元. 这里, 元素 $v \in V$ 若满足

$(\forall w \in V)((v + w = w) \wedge (w + v = w))$,

则称为零元. 易证, 零元若存在, 必唯一. 我们记这唯一的零元为 0.

(4). 有负元: V 中任意元素有负元. 这里, 元素 $w \in V$ 称为 $v \in V$ 的负元, 如果

$(v + w = 0) \wedge (w + v = 0)$.

易证, 一元素 $v \in V$ 的负元若存在, 必唯一. 今后, 记此唯一的负元为 $-v$.

此外, 数乘法还满足

(5). 数乘法对 V 中加法的分配律: $(\forall v_1, v_2 \in V \forall k \in \mathbb{F})((v_1 + v_2) \cdot k = v_1 \cdot k + v_2 \cdot k)$.

(6). 数乘法对 $\mathbb{F}$ 中加法的分配律: $(\forall v \in V \forall k_1, k_2 \in \mathbb{F})(v \cdot (k_1 + k_2) = v \cdot k_1 + v \cdot k_2)$.

(7). 数乘法与 $\mathbb{F}$ 中乘法的关系: $(\forall v \in V \forall k_1, k_2 \in \mathbb{F})(v \cdot (k_1 k_2) = (v \cdot k_1) \cdot k_2)$.

(8). 用数 $1 \in \mathbb{F}$ 作数乘法: $(\forall v \in V)(v \cdot 1 = v)$.

注 1.1.5 (数乘法中的数写在左边与写在右边)

首先容易体会, 这只是取决于习惯和规定, 没有本质区别. 可以利用同构的概念严格说明“没有本质区别”的含义. 另一方面, 本书将数乘法写在右边, 绝非为了标新立异. 历史上人们将数乘法写在左边, 是因为人们将向量的坐标(表现为数组或坐标向量)总是写成行的形式. 这样, 数与坐标向量的数乘法

$k \cdot [x_1 \quad x_2 \quad \cdots \quad x_n] = [k x_1 \quad k x_2 \quad \cdots \quad k x_n]$

可理解为矩阵乘法(数 k 理解为 1×1 矩阵), 即

$k \cdot [x_1 \quad x_2 \quad \cdots \quad x_n] = [k]_{1\times 1} [x_1 \quad x_2 \quad \cdots \quad x_n]_{1\times n}$.

在一些稍早的著作中, 如[1, p. ??], [2, pp.??], 这么写数乘法很常见. 而现代的习惯是将坐标向量默认为列的形式. 这样, 按许多现代教科书中的习惯, 将数写在左边的数乘法

$$k \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} k x_1 \\ k x_2 \\ \vdots \\ k x_n \end{bmatrix}$$

就不再能视为矩阵乘法了. 后面将会看到, 大量的技巧是通过矩阵分块的手段, 将所研究问题用矩阵乘法来表示为核心的. 而这些技巧之所以有时被初学者视为玄妙而难以掌握, 很大程度上乃是因为数乘法的这个现代的“错误”写法导致的.

例 1.1.1 ($\mathbb{F}$ 上的标准线性空间 $\mathbb{F}^n$)

这里集合 $\mathbb{F}^n$ 如(1.1.1)式定义. 加法和数乘法如下定义:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot k = \begin{bmatrix} x_1 k \\ x_2 k \\ \vdots \\ x_n k \end{bmatrix}. \quad (1.1.2)$$

特别注意, 数乘法可理解为矩阵乘法:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} [k]_{1\times 1}$$

读者需验证那些运算规则.

例 1.1.2 (几何空间作为线性空间)

[1] 华罗庚, 万哲先, 典型群,

[2] 贾可布森, 抽象代数学, 第二卷,

集合

$V = \{\text{几何空间中的有向线段}\}$

$\quad = \{\overrightarrow{AB} : A, B\text{两点都取遍几何空间}\}$,

这里, 经过平移能够重合的两个有向线段视为同一个抽象元素.

$\mathbb{F} = \mathbb{R}$.

加法: 依平行四边形法则或三角形法则进行.

数乘法: 依数的正负号将有向线段同向或反向伸缩.

读者需验证, 这样的加法和数乘法满足那些"通常的运算规则"(见注 1.1.4). 验证过程远非一蹴而就的. 这样的加法和数乘法满足 8 条"通常的运算规则", 这一事实可视为用代数运算的语言描述了大量的几何命题.

正是因为这个例子, 人们将任意线性空间中的元素也称为向量, 线性空间也称为向量空间.

例 1.1.3 (函数空间 $\mathcal{F}(I,\mathbb{R}^n)$)

这里, 记号 I 表示数轴 $\mathbb{R}$ 上的一个区间. 集合

$\mathcal{F}(I,\mathbb{R}^n) = \{f : f\text{是定义在}I\text{上, 取值于}\mathbb{R}^n\text{的函数}\}$.

给定 $f, g \in \mathcal{F}(I,\mathbb{R}^n)$, 定义 $f + g \in \mathcal{F}(I,\mathbb{R}^n)$ 如下

$f + g: \quad t \mapsto f(t) + g(t)$,

这里第二个加号"+"按例 1.1.1 中的(1.1.2)式来理解; 给定 $f \in \mathcal{F}(I,\mathbb{R}^n)$, $k \in \mathbb{R}$ 定义 $f \cdot k \in \mathcal{F}(I,\mathbb{R}^n)$ 如下

$f \cdot k: \quad t \mapsto f(t) \cdot k$,

这里第二个数乘符号" · "按例 1.1.1 中的(1.1.2)式来理解. 易见, $\mathcal{F}(I,\mathbb{R}^n)$ 是 $\mathbb{R}$ 上的线性空间. 特别地, $\mathcal{F}(I,\mathbb{R}^n)$ 的子集合

$\mathcal{C}(I,\mathbb{R}^n) = \{f : f \in \mathcal{F}(I,\mathbb{R}^n), \text{且对任意}t \in I,\ f\text{在}t\text{点连续}\}$

按上面定义的加法和数乘法也是 $\mathbb{R}$ 上的线性空间.

定义 1.1.2 (向量组及向量组拼成的抽象矩阵)

设 V 是 $\mathbb{F}$ 上的线性空间. V 中的有限序列 $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$, 或简记为 $\alpha_1, \alpha_2, \ldots, \alpha_p$, 称为 V 中的一个向量组, 简称为向量组. 向量组按顺序排成的行置于方括号中, 即以向量为元素的 1 行 p 列"矩阵"

$\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_p \end{bmatrix}$

称为向量组 $\alpha_1, \alpha_2, \ldots, \alpha_p$ 拼成的抽象矩阵.

定义 1.1.3 (向量组的线性相关性)

设 V 是 $\mathbb{F}$ 上的线性空间, $\alpha_1, \alpha_2, \ldots, \alpha_p$ 是 V 中的一个向量组.

(1) 向量组 $\alpha_1, \alpha_2, \ldots, \alpha_p$ 称为线性相关的, 如果存在不全为零的 p 个数 $k_i \in \mathbb{F}$, $i = 1, 2, \ldots, p$, 使得

$\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_p k_p = 0$,

这里, 等号右端的"0"为零向量, 见注 1.1.4. 即

$(\exists k_i \in \mathbb{F}, i = 1, 2, \ldots, p)((\neg(\wedge_{i=1}^{p}(k_i = 0))) \wedge (\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_p k_p = 0))$.

(2) 向量组 $\alpha_1, \alpha_2, \ldots, \alpha_p$ 称为线性无关的, 如果它不是线性相关的. 换言之, 如果只有当 p 个数 $k_i \in \mathbb{F}$, $i = 1, 2, \ldots, p$, 全为零时, 才有

$\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_p k_p = 0$.

或者说, 可由上式推出 p 个数 $k_i \in \mathbb{F}$, $i = 1, 2, \ldots, p$, 全为零. 亦即:

$(\forall k_i \in \mathbb{F}, i = 1, 2, \ldots, p)(((\wedge_{i=1}^{p}(k_i = 0))) \vee (\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_p k_p \neq 0))$.

向量组的线性相关性(及其后续概念, 如向量组的秩)是线性空间理论中的核心概念. 这一概念的来源, 在代数上可追溯到对线性方程组"一般理论"的探索: 即不满足于用消元法去具体地求解一个一个的线性方程组, 而要进一步回答解的存在性, 唯一性, 以及"刻画所有解"(或曰"解的结构")等问题. 在几何上则

明显地可追溯到对笛卡尔的坐标系方法的进一步理性研究：一个向量组何时能作为坐标系来使用.

注 1.1.6 (向量组线性相关性的矩阵表达)
V 中的向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性相关，就是以该向量组拼成的抽象矩阵为系数矩阵的抽象齐次线性方程组

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=0,\ (1.1.3)$$

有非零解

Ax=0有非零解iff R(A)<n

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\begin{bmatrix}k_1\\k_2\\\vdots\\k_p\end{bmatrix}\in\mathbb{F}^p,\quad \begin{bmatrix}k_1\\k_2\\\vdots\\k_p\end{bmatrix}\neq 0,$$

这里，符号 $0\in\mathbb{F}^p$ 为 $\mathbb{F}^p$ 中的零元素. 在方程(1.1.3)中，很自然地，左边的抽象矩阵乘法运算按

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\alpha_1x_1+\alpha_2x_2+\cdots+\alpha_px_p \quad (1.1.4)$$

理解；右边的 0 为线性空间 V 中的零元素.
另一方面，V 中的向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性无关，就是以该向量组拼成的抽象矩阵为系数矩阵的抽象齐次线性方程组

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=0$$

R(A)=n

仅有零解

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\begin{bmatrix}k_1\\k_2\\\vdots\\k_p\end{bmatrix}=0\in\mathbb{F}^p.$$

能够将向量组的线性相关性用矩阵表达，应归功于将一般线性空间中的元素想像成列数组，并将数乘法的数写在右边这一用心做出的规定(见注 1.1.5)，它导致了“抽象分块矩阵乘法”—公式(1.1.4)—的运用.

定义 1.1.4 (两个向量组之间的线性表示关系)
设 V 是 $\mathbb{F}$ 上的线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_p$ 和 $\beta_1,\beta_2,\ldots,\beta_q$ 是 V 中的两个向量组，$\beta\in V$.
(1) 称向量 β 可由向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，如果存在 p 个数 $k_i\in\mathbb{F}$，$i=1,2,\ldots,p$，使得
$\alpha_1k_1+\alpha_2k_2+\cdots+\alpha_pk_p=\beta$.
亦即
$(\exists k_i\in\mathbb{F}, i=1,2,\ldots,p)(\alpha_1k_1+\alpha_2k_2+\cdots+\alpha_pk_p=\beta)$.
(2) 称向量组 $\beta_1,\beta_2,\ldots,\beta_q$ 可由向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，如果每个 β_j 都可由向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，$j=1,2,\ldots,q$. 为方便，$\beta_1,\beta_2,\ldots,\beta_q$ 可由 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，用符号记为
$\{\beta_1,\beta_2,\ldots,\beta_q\}\leq_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$.

注 1.1.7 (两个向量组之间线性表示关系的矩阵表达)
向量 β 可由向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，就是以向量 β 为右端项(非奇次项)，以向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 拼成的抽象矩阵为系数矩阵的抽象非奇次线性方程组

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\beta,$$

Ax=b解的结构：
1.有解iff R(A)=R(A,b)
2.唯一解 iff R(A)=R(A,b)=r
3.无穷多解iff R(A)=R(A,b)<r

有解

$$\begin{bmatrix}x_1\\x_2\\\vdots\\x_p\end{bmatrix}=\begin{bmatrix}k_1\\k_2\\\vdots\\k_p\end{bmatrix}\in\mathbb{F}^p;$$

向量组 $\beta_1,\beta_2,\ldots,\beta_q$ 可由向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性表示，就是如下的以矩阵为未知量的抽象线性方程组

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}\begin{bmatrix}x_{11} & x_{12} & \cdots & x_{1q}\\x_{21} & x_{22} & \cdots & x_{2q}\\\vdots & \vdots & \cdots & \vdots\\x_{p1} & x_{p2} & \cdots & x_{pq}\end{bmatrix}=\begin{bmatrix}\beta_1 & \beta_2 & \cdots & \beta_q\end{bmatrix}$$

有解

$$\begin{bmatrix}x_{11} & x_{12} & \cdots & x_{1q}\\x_{21} & x_{22} & \cdots & x_{2q}\\\vdots & \vdots & \cdots & \vdots\\x_{p1} & x_{p2} & \cdots & x_{pq}\end{bmatrix}=\begin{bmatrix}k_{11} & k_{12} & \cdots & k_{1q}\\k_{21} & k_{22} & \cdots & k_{2q}\\\vdots & \vdots & \cdots & \vdots\\k_{p1} & k_{p2} & \cdots & k_{pq}\end{bmatrix}\in\mathbb{F}^{p\times q}.$$

这里，符号 $\mathbb{F}^{p\times q}$ 表示 $\mathbb{F}$ 上的 p 行 q 列矩阵的集合.

命题 1.1.1 (线性表示关系的传递性)

设 V 是 $\mathbb{F}$ 上的线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_p$；$\beta_1,\beta_2,\ldots,\beta_q$；$\gamma_1,\gamma_2,\ldots,\gamma_t$ 是 V 中的三个向量组. 若

$$\{\alpha_1,\alpha_2,\ldots,\alpha_p\}\le_{lin}\{\beta_1,\beta_2,\ldots,\beta_q\},\ \{\beta_1,\beta_2,\ldots,\beta_q\}\le_{lin}\{\gamma_1,\gamma_2,\ldots,\gamma_t\}$$

lin:线性表示
a<=lin b：a可以用b线性表示

则

$$\{\alpha_1,\alpha_2,\ldots,\alpha_p\}\le_{lin}\{\gamma_1,\gamma_2,\ldots,\gamma_t\}.$$

证明：我们利用线性表示关系的矩阵表达(注 1.1.6). 由存在 $T\in\mathbb{F}^{q\times p}, S\in\mathbb{F}^{t\times q}$ 满足

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}=\begin{bmatrix}\beta_1 & \beta_2 & \cdots & \beta_q\end{bmatrix}T,\ \begin{bmatrix}\beta_1 & \beta_2 & \cdots & \beta_q\end{bmatrix}=\begin{bmatrix}\gamma_1 & \gamma_2 & \cdots & \gamma_t\end{bmatrix}S,$$

知

$$\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_p\end{bmatrix}=\begin{bmatrix}\gamma_1 & \gamma_2 & \cdots & \gamma_t\end{bmatrix}(ST).$$

由 $ST\in\mathbb{F}^{t\times p}$ 及注 1.1.6 得证结论. 证毕

定义 1.1.5 (向量组的极大线性无关子组)

设 V 是 $\mathbb{F}$ 上的线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_p$ 是 V 中一个向量组.

(1) 有限序列 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的任一子序列 $\beta_1,\beta_2,\ldots,\beta_s$ 都称为向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的一个子向量组，简称子组. 为方便，$\beta_1,\beta_2,\ldots,\beta_s$ 是 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的一个子组，简记为

$$\{\beta_1,\beta_2,\ldots,\beta_s\}\le_{sub}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}.$$

(2) 向量组 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的一个子组 $\beta_1,\beta_2,\ldots,\beta_s$ 称为一个极大线性无关子组，如果：1) $\beta_1,\beta_2,\ldots,\beta_s$ 是线性无关的; 2) 若有

$$\{\beta_1,\beta_2,\ldots,\beta_s\}\le_{sub}\{\gamma_1,\gamma_2,\ldots,\gamma_t\}\le_{sub}\{\alpha_1,\alpha_2,\ldots,\alpha_p\},$$

其中 $t>s$，则 $\gamma_1,\gamma_2,\ldots,\gamma_t$ 为线性相关.

回忆在标准线性空间中，向量组的极大线性无关子组的一个找法：将向量组拼成矩阵，用初等行变换将矩阵化为行阶梯型，则各台阶位置处的向量构成极大线性无关子组.

易知，一向量组(“母组”)的极大线性无关子组不必是唯一的. 然而，不同的极大线性无关子组所含向量的

个数是唯一的: 这个数是由母组决定的, 是母组的一个内在属性, 称它是母组的秩. 下面来证明这个结论.

引理 1.1.1 (扁的齐次线性方程组必有非零解)
设 $A\in\mathbb{F}^{m\times n}$，$1\le m<n$ 则齐次线性方程组 $Ax=0$ 必有非零解

$$x=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}\in\mathbb{F}^n,\quad \begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}\ne 0.$$

这里，$m<n$，也就是方程的个数少于未知数的个数(不定方程), 系数矩阵呈扁形.

证明: 对 m 用数学归纳法.
(1) $m=1$. 此时，$n\ge 2$. 根据 a_{11} 是否为 0 分两种情况. 1) 若 $a_{11}=0$，则取

$$\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},$$

易知其为一非零解. 2) 若 $a_{11}\ne 0$，由 $n\ge 2$，可取

$$\begin{bmatrix}c_1\\c_2\\c_3\\\vdots\\c_n\end{bmatrix}=\begin{bmatrix}-\frac{a_{12}}{a_{11}}\\1\\0\\\vdots\\0\end{bmatrix}.$$

则由

$$Ac=\begin{bmatrix}a_{11}&a_{12}&a_{13}&\cdots&a_{1n}\end{bmatrix}\begin{bmatrix}-\frac{a_{12}}{a_{11}}\\1\\0\\\vdots\\0\end{bmatrix}=a_{11}\left(-\frac{a_{12}}{a_{11}}\right)+a_{12}\cdot 1+a_{13}\cdot 0+\cdots+a_{1n}\cdot 0=0,$$

易知这样取得 c 为一非零解.
(2) 假设 $m\le p$ 时，命题已成立，下证 $m=p+1$ 的情形. 此时 $n\ge p+2$. 同样根据 x_1 的系数，也就是系数矩阵 A 的第一列

$$\alpha_1=\begin{bmatrix}a_{11}\\a_{21}\\\vdots\\a_{p+1,1}\end{bmatrix}$$

是否为零分两种情况: 1) 若 $\alpha_1=0$，则取

$$c=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix}=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}.$$

由

$$Ac=\begin{bmatrix}\alpha_1&\alpha_2&\cdots&\alpha_n\end{bmatrix}\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}=\alpha_1\cdot 1+\alpha_2\cdot 0+\cdots+\alpha_n\cdot 0=0$$

知 c 为非零解. 2) 若 $\alpha_1 \neq 0$，不妨设 $a_{11} \neq 0$，则有与 $Ax=0$ 同解的线性方程组

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22}-\frac{a_{21}}{a_{11}}\cdot a_{12} & \cdots & a_{2n}-\frac{a_{21}}{a_{11}}\cdot a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & a_{p+1,2}-\frac{a_{p+1,1}}{a_{11}}\cdot a_{12} & \cdots & a_{p+1,n}-\frac{a_{p+1,1}}{a_{11}}\cdot a_{1n} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=0. \quad (1.1.5)$$

由归纳假设，如下 p 个方程，$n-1(\geq p+1>p)$ 个未知数的线性方程组

$$\begin{bmatrix} a_{22}-\frac{a_{21}}{a_{11}}\cdot a_{12} & \cdots & a_{2n}-\frac{a_{21}}{a_{11}}\cdot a_{1n} \\ \vdots & \vdots & \vdots \\ a_{p+1,2}-\frac{a_{p+1,1}}{a_{11}}\cdot a_{12} & \cdots & a_{p+1,n}-\frac{a_{p+1,1}}{a_{11}}\cdot a_{1n} \end{bmatrix}\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}=0$$

有非零解

$$\begin{bmatrix} x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} c_2 \\ \vdots \\ c_n \end{bmatrix}\neq 0.$$

再取

$x_1=c_1=-\frac{1}{a_{11}}(a_{12}c_2+\cdots+a_{1n}c_n)$，易知

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}\neq 0$$

是方程组(1.1.5)，从而也是 $Ax=0$ 的非零解．证毕

定理 1.1.1 (线性表示与线性无关性)
设 V 是 $\mathbb{F}$ 上的线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_p$ 和 $\beta_1,\beta_2,\ldots,\beta_q$ 是 V 中两个向量组．若 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性无关，且 $\{\alpha_1,\alpha_2,\ldots,\alpha_p\}\leq_{lin}\{\beta_1,\beta_2,\ldots,\beta_q\}$，
则 $p\leq q$.

证明：用反证法，假设 $p>q$．由线性表示关系的矩阵表达(注 1.1.7)可知，存在矩阵 $T\in\mathbb{F}^{q\times p}$，使得
$[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_p]=[\beta_1 \quad \beta_2 \quad \cdots \quad \beta_q]T$. (1.1.6)
由引理 1.1.1，存在 $c\in\mathbb{F}^p$ 非零，使得
$Tc=0$. (1.1.7)
等式(1.1.6)两边右乘 c，由(1.1.7)得

$$\begin{aligned}[\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_p]c &= [\beta_1 \quad \beta_2 \quad \cdots \quad \beta_q]Tc \\ &= 0.\end{aligned} \quad (1.1.8)$$

由 $c\in\mathbb{F}^p$ 非零，此结论与 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 线性无关矛盾．证毕

注 1.1.8 在上述证明中，V 中的元素(抽象向量)与 $\mathbb{F}$ 中的元素(系数)一起作矩阵运算(见(1.1.6)和(1.1.8)). 这种极富启发性的形式运算完全是严格的(见注 1.1.4)，它得益于将数乘法中的数写在右边这一细心的约定(见注 1.1.5).

<=sub 子组

命题 1.1.2 (子组关系与线性表示关系)
(1) 若 $\{\beta_1,\beta_2,\ldots,\beta_s\}\leq_{sub}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$，则 $\{\beta_1,\beta_2,\ldots,\beta_s\}\leq_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$；
(2) 进而若 $\beta_1,\beta_2,\ldots,\beta_s$ 是 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的一个极大线性无关子组，则还有 $\{\alpha_1,\alpha_2,\ldots,\alpha_p\}\leq_{lin}\{\beta_1,\beta_2,\ldots,\beta_s\}$.
证明: (1) 显然.
(2) 从 $\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$ 中任取一元 α_j，要说明它可由 $\beta_1,\beta_2,\ldots,\beta_s$ 线性表示．分两种情况：1) α_j 是 $\beta_1,\beta_2,\ldots,\beta_s$ 中一元．此时结论显然. 2) α_j 不在 $\beta_1,\beta_2,\ldots,\beta_s$ 中．此时，必可将 α_j 插入序列 $\beta_1,\beta_2,\ldots,\beta_s$ 中某位置，不妨设在最前面，使得
$\{\beta_1,\beta_2,\ldots,\beta_s\}\leq_{sub}\{\alpha_j,\beta_1,\beta_2,\ldots,\beta_s\}\leq_{sub}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$.

由定义 1.1.5, 向量组 $\{\beta_1,\beta_2,\ldots,\beta_s\}$ 线性无关. 而 $\{\alpha_j,\beta_1,\beta_2,\ldots,\beta_s\}$ 线性相关. 由此易知 α_j 可由 $\beta_1,\beta_2,\ldots,\beta_s$ 线性表示. 证毕

定理 1.1.2 (向量组的秩)

向量组的任意两个极大线性无关子组所含向量的数目相同. 该数目称为向量组的秩.

证明: 设 $\beta_1,\beta_2,\ldots,\beta_s$ 和 $\gamma_1,\gamma_2,\cdots,\gamma_t$ 是 $\alpha_1,\alpha_2,\ldots,\alpha_p$ 的两个极大线性无关子组. 则由子组与线性表示的关系(命题 1.1.2),

$\{\beta_1,\beta_2,\ldots,\beta_s\}\le_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_p\}$, $\{\alpha_1,\alpha_2,\ldots,\alpha_p\}\le_{lin}\{\gamma_1,\gamma_2,\ldots,\gamma_t\}$.

由此及命题 1.1.1 有

$\{\beta_1,\beta_2,\ldots,\beta_s\}\le_{lin}\{\gamma_1,\gamma_2,\ldots,\gamma_t\}$.

再由 $\beta_1,\beta_2,\ldots,\beta_s$ 线性无关及定理 1.1.1 即得 $s\le t$. 在上述推理中, 调换 $\beta_1,\beta_2,\ldots,\beta_s$ 和 $\gamma_1,\gamma_2,\cdots,\gamma_t$ 的位置, 即得 $t\le s$. 证毕

1.2 基与坐标

定义 1.2.1 (有限维线性空间, 基与坐标)

设 V 是 $\mathbb{F}$ 上的线性空间. 若有正整数 n 及 V 中的向量组 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 使得

(1) $\alpha_1,\alpha_2,\ldots,\alpha_n$ 线性无关,

(2) 任取向量 $\alpha\in V$ 均可由 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 线性表示

$$\begin{aligned}\alpha &= \alpha_1 k_1+\alpha_2 k_2+\cdots+\alpha_n k_n\\ &= \begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}\begin{bmatrix}k_1\\ k_2\\ \vdots\\ k_n\end{bmatrix}, \quad (1.2.1)\end{aligned}$$

则称 V 是 $\mathbb{F}$ 上的有限维线性空间. 向量组 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 称为 V 的一个基或坐标系; 基向量组中向量(简称为基向量)的个数 n 称为 V 的维数, 记为 $\dim V$; 在表达式(1.2.1)中,

$$\begin{bmatrix}k_1\\ k_2\\ \vdots\\ k_n\end{bmatrix}\in\mathbb{F}^n$$

称为向量 α 在基 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 下的坐标; 而表达式(1.2.1)自身则称为 α 沿坐标系 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 的展开式.

注 1.2.1 (零维线性空间)

规定仅含一个元素的线性空间(零线性空间)为零维线性空间, 其维数规定为 0. 零维线性空间也算作是有限维线性空间.

注 1.2.2 (基矩阵)

由基向量组 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 拼成的矩阵 $\begin{bmatrix}\alpha_1 & \alpha_2 & \cdots & \alpha_n\end{bmatrix}$ 称为基矩阵. 为说话方便, 称一般的线性空间中的向量 $\alpha\in V$ 为抽象向量, 并记展开式(1.2.1)为

$$\begin{bmatrix}\text{抽}\\ \text{象}\\ \text{向}\\ \text{量}\end{bmatrix}=\begin{bmatrix}\text{基矩阵}\end{bmatrix}\begin{bmatrix}\text{坐}\\ \text{标}\\ \text{向}\\ \text{量}\end{bmatrix}. \quad (1.2.2)$$

命题 1.2.1 (维数的唯一性)

若 V 是 $\mathbb{F}$ 上的有限维线性空间, $\alpha_1,\alpha_2,\ldots,\alpha_n$ 和 $\beta_1,\beta_2,\ldots,\beta_m$ 是 V 的两个基, 则 $m=n$.

证明: 由 $\alpha_1,\alpha_2,\ldots,\alpha_n$ 和 $\beta_1,\beta_2,\ldots,\beta_m$ 都是 V 基, 定义 1.2.1 中的条件(2), 及定义 1.1.4, 可知

$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}\le_{lin}\{\beta_1,\beta_2,\ldots,\beta_m\}$, 同时 $\{\beta_1,\beta_2,\ldots,\beta_m\}\le_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$.

再由定义 1.2.1 中的条件(1)及定理 1.1.1 知 $m=n$. 证毕

命题 1.2.2 (坐标系实现有限维抽象线性空间与标准线性空间的一一对应)
设 V 是 $\mathbb{F}$ 上的 n 维线性空间，$\alpha_1,\alpha_2,\ldots,\alpha_n$ 是 V 的一个选定的坐标系. 则由下式

$$\mathcal{A}:\begin{bmatrix}k_1\\k_2\\\vdots\\k_n\end{bmatrix}\mapsto \alpha_1k_1+\alpha_2k_2+\cdots+\alpha_nk_n$$

决定的映射 $\mathcal{A}:\mathbb{F}^n\to V$ 是一一对应(可逆映射).

注 1.2.3 (映射可逆的条件)
回忆映射 $\tau:S\to T$ 为一一对应(可逆映射)当且仅当下列两条件同时成立:
(1) τ 是满的(或称映上的): 对任意 $t\in T$，存在 $s\in S$ 使得 $\tau(s)=t$;
(2) τ 是单的(或称一对一的的): 对任意 $s_1,s_2\in S$，$s_1\neq s_2$，都有 $\tau(s_1)\neq\tau(s_2)$.

命题 1.2.2 的证明: 由定义 1.2.1 中条件(2)知，上述映射是满射. 下证它是单射. 假设 $k,l\in\mathbb{F}^n$ 使得 $\mathcal{A}(k)=\mathcal{A}(l)$，即
$\alpha_1k_1+\alpha_2k_2+\cdots+\alpha_nk_n=\alpha_1l_1+\alpha_2l_2+\cdots+\alpha_nl_n$.
此即:
$\alpha_1(k_1-l_1)+\alpha_2(k_2-l_2)+\cdots+\alpha_n(k_n-l_n)=0$,
由定义 1.2.1 条件(1),
$(k_1-l_1)=(k_2-l_2)=\cdots=(k_n-l_n)=0$.
此即 $k=l$. 证毕

注 1.2.4 (笛卡尔坐标系，解析几何的基础)
笛卡尔正是借助于坐标系，将欧几里德几何空间中的问题("几何问题")一对一(既不增多，也不减少)地转化为标准线性空间 $\mathbb{R}^3$ 中的问题("代数问题"). 参见例 1.1.2. 这是数学史上的伟大事件. 我们看到，坐标系这一观念，可视为线性无关性这一代数概念的几何渊源.

例 1.2.1 (无限维线性空间的例)
存在不是有限维的线性空间(简称为无限维的线性空间). 记
$\mathbb{R}[x]=\{f\mid f\in\mathcal{F}(\mathbb{R},\mathbb{R}),\text{且}f\text{可写成实系数多项式}\}$.
这里，函数"f 可写成实系数多项式"，是指存在 $k\geq 0$ 及 $k+1$ 个实数 $a_0,a_1,\ldots,a_k$，使得对任意 $x\in\mathbb{R}$,
$f(x)=a_0+xa_1+\cdots+x^ka_k$.
注意这里我们将系数写在右边，这也是为了与将数乘法的数写在右边的约定相配合，其好处下面立见.
又记
$\mathbb{R}[x]_n=\{f\mid f\in\mathcal{F}(\mathbb{R},\mathbb{R}),\text{且}f\text{可写成次数}<n\text{的实系数多项式}\}$.
按照线性空间 $\mathcal{F}(\mathbb{R},\mathbb{R})$ (见例 1.1.3)的加法和数乘法，易知 $\mathbb{R}[x]$ 和 $\mathbb{R}[x]_n$ 都是 $\mathbb{R}$ 上的线性空间. 下证 $\mathbb{R}[x]_n$ 是 n 维线性空间；而 $\mathbb{R}[x]$ 不是有限维的线性空间. 在此之前，需要下面的引理

引理 1.2.1 (实多项式函数的实系数多项式表达式是唯一的)
设 $f\in\mathcal{F}(\mathbb{R},\mathbb{R})$. 如果是指存在 $k\geq 0$ 及 $k+1$ 个实数 $a_0,a_1,\ldots,a_k$，以及 $k'\geq 0$ 及 $k'+1$ 个实数 $b_0,b_1,\ldots,b_{k'}$ 使得对任意 $x\in\mathbb{R}$,
$f(x)=a_0+xa_1+\cdots+x^ka_k$，同时，$f(x)=b_0+xb_1+\cdots+x^{k'}b_{k'}$,
不妨设 $k'\geq k$，则必有
$a_i=b_i, i=0,1,\ldots,k$，且 $b_i=0$ 若有 $k+1\leq i\leq k'$.
证明: 由已知，对任意 $x\in\mathbb{R}$,
$(b_0-a_0)+x(b_1-a_1)+\cdots+x^k(b_k-a_k)+x^{k+1}b_{k+1}+\cdots+x^{k'}b_{k'}=0$.
分别将 $k'+1$ 个不同的自变量值，例如，$x=0,1,\ldots,k'$，带入上式，则得

$$\begin{bmatrix} 1 & 0^1 & \cdots & 0^{k'} \\ 1 & 1^1 & \cdots & 1^{k'} \\ \vdots & \vdots & & \vdots \\ 1 & (k')^1 & \cdots & (k')^{k'} \end{bmatrix} \begin{bmatrix} b_0 - a_0 \\ b_1 - a_1 \\ \vdots \\ b_k - a_k \\ b_{k+1} \\ \vdots \\ b_{k'} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

由上式中 Vandermond 矩阵的非奇异即得结论. 证毕

下面证明例 1.2.1 中的结论.

证明: (1) $\mathbb{R}[x]_n$ 是 n 维线性空间. 只需找到恰有 n 个元素的一个基. 下证 $\mathbb{R}[x]_n$ 中的如下函数序列

$1, x^1, x^2, \ldots, x^{n-1}$

即是该线性空间的一个基. 1) $1, x^1, x^2, \ldots, x^{n-1}$ 线性无关. 否则, 存在

$$a = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{R}^n, a \neq 0,$$

使得

$$\begin{bmatrix} 1 & x & \cdots & x^{n-1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = a_0 + x a_1 + \cdots + x^{n-1} a_{n-1} = 0.$$

注意这里的等号“=”是线性空间 $\mathbb{R}[x]_n$ 中的等号, 即函数相等; 最右边的“0”是线性空间 $\mathbb{R}[x]_n$ 中的零元素, 即衡取 0 值的函数. 上式第二个等号两边视为同一函数的两个多项式表示, 因此由引理 1.2.1, $a = 0$, 矛盾. 2) 任一次数< n 的实系数多项式显然能由 $1, x^1, x^2, \ldots, x^{n-1}$ 线性表示.

(2) $\mathbb{R}[x]$ 不是有限维的线性空间. 反证法, 假设 $\mathbb{R}[x]$ 是有限维, 例如 N 维, 线性空间, 并设多项式函数序列

$f_1(x), f_2(x), \ldots, f_N(x)$

是线性空间 $\mathbb{R}[x]$ 的基. 用符号 $\partial(f(x))$ 表示多项式 $f(x)$ 的次数. 记

$d = \max\{\partial(f_i(x)) : i = 1, 2, \ldots, N\}$.

则由引理 1.2.1, 多项式函数 x^{d+1} 不可能写成 $f_1(x), f_2(x), \ldots, f_N(x)$ 的一个线性表示. 此与基的假设矛盾. 证毕

例 1.2.2 (标准线性空间的标准基与一般基)

如下的向量组

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \ldots, e_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

是标准线性空间 $\mathbb{F}^n$ 的一个基, 称为标准基. 特别注意, 标准基矩阵恰为 n 阶单位矩阵 I_n. 任一向量 $v \in \mathbb{F}^n$ 在标准基下的坐标就是它自己:

$v = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix} v = I_n v$.

向量组 $v_1, v_2, \ldots, v_n$ 构成 $\mathbb{F}^n$ 的基的充要条件是该向量组拼成的矩阵 $\begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$(这回是一个真正的矩阵!)是非奇异的矩阵. 此时, 任意向量 $v \in \mathbb{F}^n$ 在该基下的坐标是下面非齐次线性方程组的解:

$[v_1 \quad v_2 \quad \cdots \quad v_n]x = v$.
反过来说，我们可以将具有非奇异系数矩阵的非齐次线性方程组的求解这一代数问题解释成将右端向量沿系数矩阵的列向量组构成的坐标系展开这一几何问题．此观点有助于领悟许多“矩阵技巧”．

1.3 线性子空间

定义 1.3.1 设V是$\mathbb{F}$上的线性空间，$W \subseteq V$是非空子集．若
(1) 对任意$\alpha, \beta \in W$，均有$\alpha + \beta \in W$，
(2) 对任意$\alpha \in W, k \in \mathbb{F}$，均有$\alpha \cdot k \in W$，
则称W是V的一个线性子空间，简称子空间．

子空间W本身，按照V原有的加法及V和$\mathbb{F}$之间原有的数乘法，也是$\mathbb{F}$上线性空间：定义中的两个条件保证了原有的运算可对W进行．

只含零向量的子集合$\{0\}$和全集V自己，都是V的子空间．

例 1.3.1 (矩阵的核与像)

核：零空间：x

给定$A \in \mathbb{F}^{m\times n}$，$b \in \mathbb{F}^m$且$b \neq 0$．$A$的核定义为
$\ker A = \{x : x \in \mathbb{F}^n, Ax = 0\}$，
这也就是齐次线性方程组$Ax = 0$的解空间，它是$\mathbb{F}^n$的子空间；A的像定义为

像空间：AX :y

$\operatorname{im} A = \{y : y \in \mathbb{F}^m, \text{且存在} x \in \mathbb{F}^n \text{使得} y = Ax\} = \{Ax : x \in \mathbb{F}^n\}$，
也就是从$\mathbb{F}^n$到$\mathbb{F}^m$的映射$x \mapsto Ax$的值域，它是$\mathbb{F}^m$的子空间；然而，非齐次线性方程组的解集合
$S = \{x : x \in \mathbb{F}^n, Ax = b\}$
不是线性子空间．请根据子空间的定义验证上述三个结论．

例 1.3.2 (向量组张成的子空间)
设V是$\mathbb{F}$上的线性空间，$\alpha_1, \alpha_2, \ldots, \alpha_t$是$V$中的一个向量组．定义$V$的子集
$\operatorname{span}\{\alpha_1, \alpha_2, \ldots, \alpha_t\} = \{\alpha_1 k_1 + \alpha_2 k_2 + \cdots + \alpha_t k_t : k_i \in \mathbb{F}, i = 1,2,\ldots,t\}$.
容易验证，它是V的一个线性子空间，称为由向量组$\alpha_1, \alpha_2, \ldots, \alpha_t$张成的子空间．反之，给定$V$的一个线性子空间$W$，若能找到向量组$\beta_1, \beta_2, \ldots, \beta_r$使得恰有
$W = \operatorname{span}\{\beta_1, \beta_2, \ldots, \beta_r\}$，
则称向量组$\alpha_1, \alpha_2, \ldots, \alpha_t$为子空间$W$的一个生成向量组(简称生成组)．生成组的概念提供了子空间的一种表现方法，以后一般通过指定生成组来指定子空间．子空间本身作为线性空间也有基的概念．读者可以证明，子空间的生成组的一个极大线性无关子组是子空间的一个基．
对给定的$A \in \mathbb{F}^{m\times n}$，$\operatorname{im} A$也就是由$A$的$n$个列向量所构成的$\mathbb{F}^m$中的向量组(简称为$A$的列向量组)所张成的$\mathbb{F}^m$的子空间．因此，可以通过给定$\mathbb{F}$上的行数为$m$的一个矩阵$A$来给定标准线性空间$\mathbb{F}^m$的子空间—作为$\operatorname{im} A$.

下面讨论子空间本身作为线性空间的基．

引理 1.3.1 (无限维空间中有线性无关的无穷序列)
设V是$\mathbb{F}$上的无限维线性空间，则存在V中的无限序列
$\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$，
使得对任意的$k \geq 1$，有限序列(向量组)$\alpha_1, \alpha_2, \ldots, \alpha_k$是线性无关的(这样的无穷序列本身称为是线性无关的)
证明：V不是零维的，因此必有非零向量，记为$\alpha_1 \neq 0$在V中．如果
$\operatorname{span}\{\alpha_1\} = V$，
则V是 1 维线性空间，矛盾；如果
$\operatorname{span}\{\alpha_1\} \subset V, \text{但} \neq V$，
则必有$\alpha_2 \in V \setminus \operatorname{span}\{\alpha_1\}$，从而有限序列$\alpha_1, \alpha_2$线性无关．如果
$\operatorname{span}\{\alpha_1, \alpha_2\} = V$，

则V是 2 维线性空间, 矛盾; 如果

$\mathrm{span}\{\alpha_1,\alpha_2\}\subset V$, 但$\neq V$,

则必有$\alpha_3\in V\setminus\mathrm{span}\{\alpha_1,\alpha_2\}$, 从而有限序列$\alpha_1,\alpha_2,\alpha_3$线性无关. 如此不断继续, 可得定理. 证毕

上述推理中, 严格地说, 用了所谓的“选择公理”[1]: 粗略地说, 即总可以在一非空(无限)集合中“选出”一个元素来.

定理 1.3.1 (子空间的维数与基和全空间的维数与基的关系)
设V是$\mathbb{F}$上的有限维线性空间, W是V的一个子空间. 则
(1) W也是有限维的线性空间, 且$\dim W\leq\dim V$;
(2) W的任一基必可扩充为V的一个基.
证明: 设$\dim V=n$, $\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的一个基.
(1) 先证明 1) W是有限维的线性空间, 用反证法. 假设W是无限维的线性空间, 由引理 1.3.1, 可以找到W中的线性无关的无穷序列

$\beta_1,\beta_2,\ldots$.

显然, 由$\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的一个基及$W\subset V$, 我们有

$\{\beta_1,\beta_2,\ldots,\beta_{n+1}\}\leq_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$.

于是由定理 1.1.1, $n+1\leq n$, 矛盾. 再证明 2) $\dim W\leq\dim V$. 设$\dim W=m$, 并且$\gamma_1,\gamma_2,\ldots,\gamma_m$是$W$的一个基. 则由$\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的一个基及$W\subset V$, 我们有

$\{\gamma_1,\gamma_2,\ldots,\gamma_m\}\leq_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$.

于是由定理 1.1.1, $m\leq n$.
(2) 如上述(1), 2), 设$\dim W=m$, 并且$\gamma_1,\gamma_2,\ldots,\gamma_m$是$W$的一个基. 要说明$\gamma_1,\gamma_2,\ldots,\gamma_m$可扩充成$V$的基. 可采用引理 1.3.1 证明中的方法来扩充, 但我们不这么做, 而是证明下面更强的结论(斯坦尼兹替换定理[2]): 可在V的基$\alpha_1,\alpha_2,\ldots,\alpha_n$中选出$n-m$个向量来, 记为$\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_{n-m}}$,使得向量组

$\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_{n-m}}$

为V的一个基.
第一步先看α_1. 若$\alpha_1\notin\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m\}$, 则令$i_1=1$, 转入下一步; 若$\alpha_1\in\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m\}$, 则直接转入下一步.
第二步看α_2. 若$\alpha_2\notin\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1}\}$(第一步中的情形一), 则令$i_2=2$, 转入下一步; 或$\alpha_2\notin\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m\}$(第一步中的情形二), 则令$i_1=2$, 转入下一步. 若$\alpha_2\in\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1}\}$(第一步中的情形一), 或$\alpha_2\in\mathrm{span}\{\gamma_1,\gamma_2,\ldots,\gamma_m\}$(第一步中的情形二), 则直接转入下一步.
如此一直进行n步, 检查完$\alpha_1,\alpha_2,\ldots,\alpha_n$中的所有的向量, 设被筛选出的记为$\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}$(暂时并未排除$r=0$). 由筛选过程容易看出, 向量组$\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}$线性无关, 且

$\{\alpha_1,\alpha_2,\ldots,\alpha_n\}\leq_{lin}\{\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}\}$. (1.3.1)

另一方面, 由$\alpha_1,\alpha_2,\ldots,\alpha_n$是$V$的基显然有

$\{\gamma_1,\gamma_2,\ldots,\gamma_m,\alpha_{i_1},\alpha_{i_2},\ldots,\alpha_{i_r}\}\leq_{lin}\{\alpha_1,\alpha_2,\ldots,\alpha_n\}$. (1.3.2)

因此, 由(1.3.1), (1.3.2)及定理 1.1.1, 有$m+i_r=n$. 证毕

注 1.3.1 (标准线性空间中有限维子空间的基扩充全空间基的一个方法)
这等价于将一个列满秩的矩阵扩充列, 使成为一个方的非奇异矩阵. 设$A\in\mathbb{F}^{n\times m}$, $\mathrm{rank}\,A=m$(从而必有$m\leq n$), $B\in\mathbb{F}^{n\times n}$非奇异(例如取$B=I_n$). 将$A$和$B$拼成矩阵$\begin{bmatrix}A & B\end{bmatrix}$. 用初等行变换将$\begin{bmatrix}A & B\end{bmatrix}$化为行阶梯型. 则在后$n$列中一定恰有$n-m$处台阶. 在$B$的列向量中筛出位于该$n-m$个位置的, 记为$b_{i_1},b_{i_2},\ldots,b_{i_{n-m}}$. 则矩阵$\begin{bmatrix}A & b_{i_1} & b_{i_2} & \cdots & b_{i_{n-m}}\end{bmatrix}$即为非奇异的方阵.

定义 1.3.2 (子空间的交与和)
设V是$\mathbb{F}$上的线性空间, V_1,V_2是V的两个子空间. 则子集合

[1] 贾可布森
[2] 范德瓦尔登

$V_1 \cap V_2 = \{v : v \in V_1, v \in V_2\}$

子空间的交与和：集合的交与并

及

$V_1 + V_2 = \{v : \text{存在} v_1 \in V_1, v_2 \in V_2 \text{使得} v = v_1 + v_2\} = \{v_1 + v_2 : v_1 \in V_1, v_2 \in V_2\}$

都是V 的子空间，分别称为V_1与V_2的交与和．类似地，可定义多个子空间的交与和．

设$V_1 = \mathrm{span}\{\alpha_1, \ldots, \alpha_s\}$，$V_1 = \mathrm{span}\{\beta_1, \ldots, \beta_t\}$，则显然有

$V_1 + V_2 = \mathrm{span}\{\alpha_1, \ldots, \alpha_s, \beta_1, \ldots, \beta_t\}$．

再从生成组中筛出极大线性无关子组，即为$V_1 + V_2$ 的基．

定义 1.3.3 (子空间的直和)

V1和V2没有交集时的和称：直和

设V 是$\mathbb{F}$ 上的线性空间，V_1, V_2 及W 是V 的子空间．若

(1) $V_1 + V_2 = W$，

(2) $V_1 \cap V_2 = \{0\}$，

则称W 是V_1与V_2的直和；或V_1, V_2 为W 的一个直和分解．记为$V_1 \oplus V_2 = W$．

子空间的直和与子空间的和相比，不同之处在于，将和空间中的元素表为该两子空间中元素之和时，表法是唯一的，即有下面的

命题 1.3.1 (直和中元素的唯一分解性)

设$V_1 \oplus V_2 = W$．若$w = v_1 + v_2$，$w = v'_1 + v'_2$，其中，$w \in W$，$v_1, v'_1 \in V_1$，及$v_2, v'_2 \in V_2$．则必有$v_1 = v'_1, v_2 = v'_2$．

证明：由$w = v_1 + v_2$，$w = v'_1 + v'_2$ 有

$v_1 - v'_1 = v'_2 - v_2$；

由$v_1, v'_1 \in V_1$ 及$v_2, v'_2 \in V_2$ 有$v_1 - v'_1 \in V_1$ 及$v'_2 - v_2 \in V_2$．因此

$v_1 - v'_1 = v'_2 - v_2 \in V_1 \cap V_2$，

从而由$V_1 \cap V_2 = \{0\}$ 知$v_1 = v'_1, v_2 = v'_2$．

定义 1.3.4 (直和分解决定的投影)

设$V_1 \oplus V_2 = W$．则映射$\mathcal{A}_i : W \to V_i$，$i = 1, 2$，由下式决定

$w = \mathcal{A}_1(w) + \mathcal{A}_2(w)$，$w \in W$．

称映射$\mathcal{A}_1$是将W 沿V_2向V_1的投影，$\mathcal{A}_2$是将W 沿V_1向V_2投影．

设$\mathbb{F}^n = \mathrm{im}A \oplus \mathrm{im}B$，不妨假设$A, B$ 都是列满秩的．下面我们具体算出上述两个投影．其实就是对任意的$w \in \mathbb{F}^n$，求解x_1, x_2 使得

$w = Ax_1 + Bx_2$，

注意，$Ax_1 \in \mathrm{im}A$，$Bx_2 \in \mathrm{im}B$．记$T = \begin{bmatrix} A & B \end{bmatrix}$，$T^{-1} = \begin{bmatrix} C \\ D \end{bmatrix}$．于是

$$w = Ax_1 + Bx_2 = T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = TT^{-1}w = \begin{bmatrix} A & B \end{bmatrix}\begin{bmatrix} C \\ D \end{bmatrix} w = ACw + BDw .$$

因此沿$\mathrm{im}B$ 向$\mathrm{im}A$ 的投影为

$\mathcal{A}_1(w) = ACw$，

沿$\mathrm{im}A$ 向$\mathrm{im}B$ 的投影为

$\mathcal{A}_2(w) = BDw$，

AC 和BD 分别是一对投影矩阵，满足

$$AC + BD = \begin{bmatrix} A & C \end{bmatrix}\begin{bmatrix} B \\ D \end{bmatrix} = TT^{-1} = I_n .$$

又由$T^{-1}T = I$ 有

$CA = I_p, DB = I_q, CB = 0, DA = 0$．

这里 p,q 分别是 A,B 的列数. 于是

$$(AC)^2 = ACAC = AI_pC = AC,$$

$$(BD)^2 = BDBD = BI_qD = BD.$$

即投影矩阵是幂等的.

任给一个幂等的 n 阶方阵 P, 则

$$\mathrm{im}P \bigcap \ker P = \{0\},$$ 进而由

$$\dim(\mathrm{im}P) + \dim(\ker P) = n$$

知

$$\mathrm{im}P \oplus \ker P = \mathbb{F}^n.$$

令 $Q = I - P$, 则 Q 也是幂等的, 且

$$\mathrm{im}Q = \ker P.$$

于是

$$\mathbb{F}^n = \mathrm{im}P \oplus \mathrm{im}Q.$$

小结: 任给一个幂等的 n 阶方阵 P, 则

(1) $\ker P$ 与 $\mathrm{im}\, P$ 互补;

(2) P 是沿 $\ker P$ 向 $\mathrm{im}\, P$ 的投影矩阵, $I-P$ 是沿 $\mathrm{im}\, P$ 向 $\ker P$ 的投影矩阵.

注: 幂等矩阵并不只是给定了子空间 $\mathrm{im}\, P$, 它同时给定了该子空间的补子空间 $\ker P$. 因此任何一个 $\mathbb{F}^n$ 的任何一个(有序的)直和分解可由一个幂等矩阵来表示, 有序直和分解的全体与幂等矩阵的全体之间一一对应.

定义 1.3.5 (补子空间)

设 V 是 $\mathbb{F}$ 上的线性空间, V_1, V_2 是 V 的子空间. 若 $V_1 \oplus V_2 = V$, 则称 V_1, V_2 是互补的子空间, 或 V_2 是 V_1 的补子空间.

命题 1.3.2 (任一子空间必有补子空间)

设 V 是 $\mathbb{F}$ 上的有限维线性空间, V_1 是 V 的子空间. 则存在子空间 V_2, 使得 $V_1 \oplus V_2 = V$.

证明: 请由定理 1.3.1 来证明. 证毕

1.4 线性映射

定义 1.4.1 (线性映射与线性变换)

设 V_1, V_2 是 $\mathbb{F}$ 上的两个线性空间. 映射 $\mathcal{A}: V_1 \to V_2$ 称为从 V_1 到 V_2 的线性映射, 如果它保持加法和数乘法:

(1) 对任意 $\alpha_1, \alpha_2 \in V_1$ 有 $\mathcal{A}(\alpha_1 + \alpha_2) = \mathcal{A}(\alpha_1) + \mathcal{A}(\alpha_2)$ (和的像等于像的和);

(2) 对任意 $\alpha \in V_1, k \in \mathbb{F}$ 有 $\mathcal{A}(\alpha \cdot k) = \mathcal{A}(\alpha) \cdot k$ (倍数的像等于像的倍数).

注意在上述两公式中, 左边的加法和数乘法是 V_1 上的运算, 右边的是 V_2 上的. 若从一个线性空间到该线性空间自己的线性映射也称为该线性空间上的线性变换.

例 1.4.1 (线性映射及非线性映射的例)

请验证, 映射

$$\mathcal{A}: \mathbb{R}^2 \to \mathbb{R}^2$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} x_1 + x_2 \\ x_1 x_2 \end{bmatrix}$$

不是线性映射; 映射

$$\mathcal{B}: \mathbb{R}^3 \to \mathbb{R}^2$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_1 + 3x_2 \end{bmatrix}$$

是线性映射.

定义 1.4.2 (线性同构)
设V_1,V_2是$\mathbb{F}$上的两个线性空间. 线性映射$\mathcal{A}:V_1\to V_2$称为从V_1到V_2的线性同构, 如果它作为映射是可逆映射(或一一对应).

命题 1.4.1 (可逆线性映射的逆映射也是线性映射)
设V_1,V_2是$\mathbb{F}$上的两个线性空间. 映射$\mathcal{A}:V_1\to V_2$是从V_1到V_2的线性同构. 则对任意$\beta_1,\beta_2\in V_2$, $k\in\mathbb{F}$有
$\mathcal{A}^{-1}(\beta_1\cdot k+\beta_2)=\mathcal{A}^{-1}(\beta_1)\cdot k+\mathcal{A}^{-1}(\beta_2)$.

证明: 由$\mathcal{A}$是单射, 只需证明
$\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1\cdot k+\beta_2)\right)=\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1)\cdot k+\mathcal{A}^{-1}(\beta_2)\right)$.
由逆映射定义,
$\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1\cdot k+\beta_2)\right)=\beta_1\cdot k+\beta_2$;
由$\mathcal{A}$是线性的及逆映射定义,
$$\begin{aligned}\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1)\cdot k+\mathcal{A}^{-1}(\beta_2)\right)&=\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1)\cdot k\right)+\mathcal{A}\left(\mathcal{A}^{-1}(\beta_2)\right)\\&=\mathcal{A}\left(\mathcal{A}^{-1}(\beta_1)\right)\cdot k+\beta_2\\&=\beta_1\cdot k+\beta_2.\end{aligned}$$
证毕

注 1.4.1 (坐标系实现有限维抽象线性空间与标准线性空间的线性同构)
从命题 1.2.2 易知.

由定义 1.4.2 及命题 1.4.1 知, 同构的线性空间, 它们的任何与线性空间结构有关的性质都是一样的. 因此只需在线性空间的同构类中选出一个代表来研究即可. 而同样维数的有限维线性空间都是同构的, 可用同维数的标准线性空间作为代表.

例 1.4.2 (矩阵与标准线性空间之间的线性映射两事物的等同性)
由任给矩阵$A\in\mathbb{F}^{m\times n}$通过右乘列向量, 能决定线性映射
$\mathcal{A}_A:\mathbb{F}^n\to\mathbb{F}^m$
如下
$\mathcal{A}_A:x\mapsto Ax$. (1.4.1)
请读者作为练习验证, 若$A_1,A_2\in\mathbb{F}^{m\times n}$, $A_1\neq A_2$, 必有$\mathcal{A}_{A_1}\neq\mathcal{A}_{A_2}$, 即按此方式, 不同的矩阵决定不同的线性映射. 反之, 上述构造过程还是可逆的, 即任意给定线性映射$\mathcal{A}:\mathbb{F}^n\to\mathbb{F}^m$, 可找到矩阵$A\in\mathbb{F}^{m\times n}$使得恰有$\mathcal{A}=\mathcal{A}_A$, 找法如下: 记$\mathbb{F}^n$的标准基, 即单位矩阵$I_n$的列向量组为$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$, 定义
$A=\begin{bmatrix}\mathcal{A}(\varepsilon_1) & \mathcal{A}(\varepsilon_2) & \cdots & \mathcal{A}(\varepsilon_n)\end{bmatrix}$,
也就是$\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n$的像$\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\dots,\mathcal{A}(\varepsilon_n)$ (它们在$\mathbb{F}^m$中)拼成的矩阵. 下面验证, 此矩阵即为所欲找. 任给
$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}\in\mathbb{F}^n,$$
由$x=\varepsilon_1\cdot x_1+\varepsilon_2\cdot x_2+\cdots+\varepsilon_n\cdot x_n$易见

$$
\begin{aligned}
\mathcal{A}_A(x) &= Ax \\
&= [\mathcal{A}(\varepsilon_1) \quad \mathcal{A}(\varepsilon_2) \quad \cdots \quad \mathcal{A}(\varepsilon_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \\
&= \mathcal{A}(\varepsilon_1) \cdot x_1 + \mathcal{A}(\varepsilon_2) \cdot x_2 + \cdots + \mathcal{A}(\varepsilon_n) \cdot x_n \\
&= \mathcal{A}(\varepsilon_1 \cdot x_1 + \varepsilon_2 \cdot x_2 + \cdots + \varepsilon_n \cdot x_n) \\
&= \mathcal{A}(x).
\end{aligned}
$$

注意到在上述构造中，向量 $\mathcal{A}(\varepsilon_j) \in \mathbb{F}^m$ 本身也可视为它在 $\mathbb{F}^m$ 的标准基，即矩阵 I_m 的列向量组，下的坐标. 由此观察，可将上述构造推广到一般的线性映射 $\mathcal{A}: V_1 \to V_2$ 的矩阵表示.

定义 1.4.3 (线性映射的矩阵表示)
给定 $\mathbb{F}$ 上的线性空间 V_1, V_2，及线性映射 $\mathcal{A}: V_1 \to V_2$. 设 $\dim V_1 = n$， $\dim V_n = m$，并设

$\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$

为 V_1 的一个基(称为入口基);

$\eta_1, \eta_2, \ldots, \eta_m$

为 V_2 的一个基(称为出口基). 记第 j 个入口基向量 $\varepsilon_j \in V_1$ 在 $\mathcal{A}$ 下的像 $\mathcal{A}(\varepsilon_j) \in V_2$ 在出口基 $\eta_1, \eta_2, \ldots, \eta_m$ 下的坐标为

$$
a_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \in \mathbb{F}^m,
$$

即

$$
\mathcal{A}(\varepsilon_j) = [\eta_1 \quad \eta_2 \quad \cdots \quad \eta_m] \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, \quad j = 1, 2, \ldots, n,
$$

则由 $\mathbb{F}^m$ 中的向量组 $a_1, a_2, \ldots a_n$ 拼成的矩阵

$A = [a_1 \quad a_2 \quad \cdots \quad a_n] = [a_{ij}]_{m\times n}$

称为 $\mathcal{A}$ 在相应的入口基和出口基下的表示.

注 1.4.2 (线性映射的矩阵表示定义的公式化)
定义记号

$\mathcal{A}[\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n] = [\mathcal{A}(\varepsilon_1) \quad \mathcal{A}(\varepsilon_2) \quad \cdots \quad \mathcal{A}(\varepsilon_n)]$,

即线性映射作用在向量组拼成的矩阵上，定义为向量组中每个向量的像按原顺序所成的向量组(简称为向量组的像)，拼成的矩阵. 则有下面的公式

$\mathcal{A}[\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n] = [\eta_1 \quad \eta_2 \quad \cdots \quad \eta_m]A$.

用文字表示，读作

$$
\begin{bmatrix} \text{线性} \\ \text{映射} \end{bmatrix} \begin{bmatrix} \text{入口基} \\ \text{矩阵} \end{bmatrix} = \begin{bmatrix} \text{出口基} \\ \text{矩阵} \end{bmatrix} \begin{bmatrix} \text{表示} \\ \text{矩阵} \end{bmatrix}. \quad (1.4.2)
$$

aij是第j个入口基向量的像
在第i个出口基向量上的坐标分量

为方便记忆，我们进一步指出：表示矩阵的第 j 列是第 j 个入口基向量的像在出口基下的坐标；表示矩阵的第 i 行是所有入口基向量(共 n 个)在第 i 个出口基向量上的坐标分量拼成的行.

读者也可以类似例 1.4.2, 通过固定入口基和出口基, 建立并证明, 矩阵与一般线性空间之间的线性映射的等同性的结论.

定理 1.4.1 (用坐标计算线性映射)
设线性映射 $\mathcal{A}: V_1 \to V_2$ 在入口基 $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ 和出口基 $\eta_1, \eta_2, \ldots, \eta_m$ 下的矩阵表示为 A. 又设 $\alpha \in V_1$ 在入口基下的坐标为 $x \in \mathbb{F}^n$, 则 $\mathcal{A}(\alpha) \in V_2$ 在出口基下的坐标为 $Ax \in \mathbb{F}^m$.
证明: 也就是要验证
$\mathcal{A}(\alpha) = [\eta_1 \quad \eta_2 \quad \cdots \quad \eta_m](Ax)$.
这由如下计算可见:

$$
\begin{aligned}
[\eta_1 \quad \eta_2 \quad \cdots \quad \eta_m](Ax) &= ([\eta_1 \quad \eta_2 \quad \cdots \quad \eta_m]A)x \text{ (由线性空间运算法则)} \\
&= (\mathcal{A}[\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n])x \text{ (由线性映射的矩阵表示的定义)} \\
&= \mathcal{A}([\varepsilon_1 \quad \varepsilon_2 \quad \cdots \quad \varepsilon_n]x) \text{ (由线性映射的定义)} \\
&= \mathcal{A}(\alpha) \text{ (由}x\text{为}\alpha\text{的坐标).}
\end{aligned}
$$

证毕

例 1.4.3 (微分算子的矩阵表示)
微分算子 $\mathcal{D}: \mathbb{R}[x]_4 \to \mathbb{R}[x]_3$ 表示导数运算, 即 $\mathcal{D}: f(x) \mapsto f'(x)$. 容易验证, 它是线性映射. 选 $\mathbb{R}[x]_4$ 的基(入口基): $1, x, x^2, x^3$; 选 $\mathbb{R}[x]_3$ 的基(出口基): $1, x, x^2, x^3$. 现计算微分算子在这对基下的矩阵表示.

$$
\begin{aligned}
\mathcal{D}[1 \quad x \quad x^2 \quad x^3] &= [\mathcal{D}(1) \quad \mathcal{D}(x) \quad \mathcal{D}(x^2) \quad \mathcal{D}(x^3)] \\
&= [0 \quad 1 \quad 2x \quad 3x^2] \\
&= [1 \quad x \quad x^2]\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.
\end{aligned}
$$

于是, $\mathcal{D}$ 的矩阵表示为

$$
D = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}.
$$

对任给 $f(x) \in \mathbb{R}[x]_4$, 计算 $\mathcal{D}(f(x))$ 可以有两种方法: 一种是根据映射 $\mathcal{D}$ 的具体意义(求导数)来计算; 另一种是通过坐标和矩阵来计算, 下面以 $f(x) = 5x^3 + 2x^2 + 1$ 为例来说明.
第一步: 计算 $f(x)$ 在入口基下的坐标.

$$
\begin{aligned}
f(x) &= 5x^3 + 2x^2 + 1 \\
&= [1 \quad x \quad x^2 \quad x^3]\begin{bmatrix} 1 \\ 0 \\ 2 \\ 5 \end{bmatrix}.
\end{aligned}
$$

第二步: 用矩阵乘以上述坐标, 得到 $\mathcal{D}(f(x))$ 在出口基下的坐标.

$$
\begin{aligned}
D\begin{bmatrix} 1 \\ 0 \\ 2 \\ 5 \end{bmatrix} &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 2 \\ 5 \end{bmatrix} \\
&= \begin{bmatrix} 0 \\ 4 \\ 15 \end{bmatrix}.
\end{aligned}
$$

第三步: 与出口基作线性组合, 得 $\mathcal{D}(f(x))$.

$$\mathcal{D}(f(x)) = \begin{bmatrix} 1 & x & x^2 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 15 \end{bmatrix}$$
$$= 4x + 15x^2.$$
对于多项式求导这个具体问题来说，上述计算很显然麻烦. 但它表面，用特定的线性空间的特定的运算弄出来的特定的线性映射，都可以归结为矩阵乘法这一“通用运算”.

例 1.4.4 (绕指定了正方向的固定轴旋转角度θ的变换$\mathcal{A}$的矩阵表示)
注意，这里指定了轴的方向后，角度的正负号的意义随之确定(依右手系规则)
旋转变换首先是几何空间中的点到点的变换. 在旋转轴上选定一点O(“坐标原点”)，按例 1.1.2 将几何空间建模成线性空间V后，旋转变换遂成为V上的线性变换.
以O为起点沿旋转轴正方向取单位长有向线段，记为e_z，再取以O为起点的另两单位长有向线段e_x, e_y，使得e_x, e_y, e_z构成线性空间V中的右手直角坐标系. 入口基和出口基都选为e_x, e_y, e_z，来计算$\mathcal{A}$的矩阵表示.
$$\begin{aligned}\mathcal{A}\begin{bmatrix} e_x & e_y & e_z \end{bmatrix} &= \begin{bmatrix} \cos\theta\cdot e_x + \sin\theta\cdot e_y & -\sin\theta\cdot e_x + \cos\theta\cdot e_y & e_z \end{bmatrix} \\ &= \begin{bmatrix} e_x\cos\theta + e_y\sin\theta & e_x(-\sin\theta) + e_y\cos\theta & e_z \end{bmatrix} \\ &= \begin{bmatrix} e_x & e_y & e_z \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.\end{aligned}$$
因此，$\mathcal{A}$的矩阵表示为
$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

例 1.4.5 (几何空间中的镜面反射变换$\mathcal{B}$的矩阵表示)
镜面反射首先是几何空间中的点到点的变换：在几何空间中选定一个平面(作为反射面)，在该平面上的点不变，不在该平面上的点变成它关于该平面的对称点.
在反射面上选定一点O，按例 1.1.2 将几何空间建模成线性空间V后，镜面反射变换$\mathcal{B}$也成为线性空间V上的线性变换.
以O为起点在反射平面内建立平面直角坐标系e_x, e_y. 再选e_z，使e_x, e_y, e_z构成线性空间V中的右手直角坐标系. 入口基和出口基都选为e_x, e_y, e_z，来计算$\mathcal{B}$的矩阵表示.
$$\begin{aligned}\mathcal{B}\begin{bmatrix} e_x & e_y & e_z \end{bmatrix} &= \begin{bmatrix} e_x & e_y & -e_z \end{bmatrix} \\ &= \begin{bmatrix} e_x & e_y & e_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.\end{aligned}$$
因此，$\mathcal{B}$的矩阵表示为
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

1.5 矩阵的等价与相似

1.5.1 矩阵等价

定义 1.5.1 (矩阵等价)
两个矩阵$A, B \in \mathbb{F}^{m\times n}$称为等价，如果存在$n$阶非奇异矩阵$P \in \mathbb{F}^{n\times n}$和$m$阶非奇异矩阵$Q \in \mathbb{F}^{m\times m}$使得$AP = QB$.

这与通常的定义(指存在n阶非奇异矩阵$P\in\mathbb{F}^{n\times n}$和$m$阶非奇异矩阵$Q\in\mathbb{F}^{m\times m}$使得$QAP=B$)没有区别：只需将定义 1.5.1 中的$Q$写成$Q^{-1}$. 我们知道，矩阵等价的通常的定义，其目的是用矩阵乘法语言来描述和记录对矩阵实施左右初等变换的信息. 现在，按我们这里的定义，可以看见矩阵等价概念的另一个深刻的几何解释.

注 1.5.1 (矩阵等价的几何意义)
由$A\in\mathbb{F}^{m\times n}$决定$\mathbb{F}^n$到$\mathbb{F}^m$的线性映射$\mathcal{A}_A: x\mapsto Ax$，将$n$阶非奇异矩阵$P\in\mathbb{F}^{n\times n}$的列向量组视为$\mathbb{F}^n$中的一个基(入口基)，$m$阶非奇异矩阵$Q\in\mathbb{F}^{m\times m}$视为$\mathbb{F}^m$中的一个基(出口基). 注意到$\mathcal{A}_A(P)=AP$，则定义 1.5.1 中的$B$恰为线性映射$\mathcal{A}_A$在这对基下的矩阵表示. 特别地，线性映射$\mathcal{A}_A$在入口基和出口基都是标准基时(即$P=I_n$，$Q=I_m$)的矩阵表示即为$A$自己.

下面为简洁起见，直接将“由矩阵A决定的线性映射$\mathcal{A}_A$”称为“线性映射A”. 换言之，根据上下文，符号A即可表示矩阵，也可表示映射.

问题 1.5.1 (矩阵的等价“最简型”)
给定矩阵$A\in\mathbb{F}^{m\times n}$，找非奇异矩阵$P\in\mathbb{F}^{n\times n}$，$Q\in\mathbb{F}^{m\times m}$，使得
$$B=Q^{-1}AP$$
“尽可能简单”. 这问题的几何实质是：给定从$\mathbb{F}^n$到$\mathbb{F}^m$的线性映射$\mathcal{A}_A: x\mapsto Ax$，寻找入口基$p_1,p_2,\ldots,p_n$和出口基$q_1,q_2,\ldots,q_m$，使得线性映射$\mathcal{A}_A$在这对基下的矩阵表示“尽可能简单”.

引理 1.5.1 (矩阵核的补子空间与像同构)
设矩阵$A\in\mathbb{F}^{m\times n}$，$W\oplus\ker A=\mathbb{F}^n$. 记
$$A(W)=\{Ax: x\in W\}.$$
dim(ker$A$)=n-rank$A$

则$A(W)=\operatorname{im}A$，且$\dim W=\dim(A(W))=\operatorname{rank}A$.

证明：易知子空间$A(W)\subseteq\operatorname{im}A$，因此只需证$\dim A(W)=\dim(\operatorname{im}A)$. 由$W\oplus\ker A=\mathbb{F}^n$可知
$$\dim W=n-\dim(\ker A)=n-(n-\operatorname{rank}A)=\operatorname{rank}A.$$
设子空间W的基为$p_1,p_2,\ldots,p_r$，这里$r=\operatorname{rank}A=\dim(\operatorname{im}A)$. 下证向量组$Ap_1,Ap_2,\ldots,Ap_r$为$A(W)$的一个基. 若
$$Ap_1\cdot x_1+Ap_2\cdot x_2+\cdots+Ap_r\cdot x_r=0,$$
则
$$A(p_1\cdot x_1+p_2\cdot x_2+\cdots+p_r\cdot x_r)=0,$$
此即
$$(p_1\cdot x_1+p_2\cdot x_2+\cdots+p_r\cdot x_r)\in\ker A.$$
于是
$$(p_1\cdot x_1+p_2\cdot x_2+\cdots+p_r\cdot x_r)\in W\bigcap\ker A=\{0\},$$
进而由$p_1,p_2,\ldots,p_r$线性无关有
$$x_1=x_2=\cdots=x_r=0,$$
因此向量组$Ap_1,Ap_2,\ldots,Ap_r$线性无关，进而易知它为$A(W)$的一个基. 于是$\dim(A(W))=r$. 证毕

定理 1.5.1 (由核与像确定的入口基和出口基给出等价标准型)
给定矩阵$A\in\mathbb{F}^{m\times n}$，$\operatorname{rank}A=r$. 设子空间$W\subseteq\mathbb{F}^n$是$\ker A$的补子空间；子空间$U\subseteq\mathbb{F}^m$是$\operatorname{im}A$的补子空间，并取$W$的基$p_1,p_2,\ldots,p_r$；$\ker A$的基$p_{r+1},\ldots,p_n$；$U$的基$q_{r+1},\ldots,q_m$. 则

(1) $p_1,p_2,\ldots,p_r,p_{r+1},\ldots,p_n$是入口空间$\mathbb{F}^n$的基，$A(p_1),A(p_2),\ldots,A(p_r),q_{r+1},\ldots,q_m$是出口空间$\mathbb{F}^m$的基.

(2) A在(1)中的入口基和出口基下的表示为

$$\left[\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & 0_{m-r,n-r} \end{array}\right].$$

即

$$A\underbrace{[p_1, p_2, \ldots, p_r, p_{r+1}, \ldots, p_n]}_{P}=\underbrace{[A(p_1), A(p_2), \ldots, A(p_r), q_{r+1}, \ldots, q_m]}_{Q}\left[\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & 0_{m-r,n-r} \end{array}\right].$$

证明: 由引理 1.5.1 可得(1); 可直接验证(2). 证毕

还可以进一步写成上述定理的如下形式的逆定理: 若已有可逆矩阵 P,Q 使得

$$AP=Q\left[\begin{array}{ccc|c} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \\ \hline & 0 & & 0_{m-t,n-t} \end{array}\right],$$

则必有 $t=\operatorname{rank}A$, 且可由 P 决定出 $\ker A$ 及其补, 由 Q 决定出 $\operatorname{im}A$ 及其补. 请读者细化.

注 1.5.2 (初等行、列变换下的标准型)
我们知道, 经过一系列初等行、列变换, 可将任一矩阵 $A\in\mathbb{F}^{m\times n}$, $\operatorname{rank}A=r$ 化为标准型, 即有可逆矩阵 P,Q 使得

$$QAP=\begin{bmatrix} I_r & 0 \\ 0 & 0_{m-r,n-r} \end{bmatrix},$$

而定理 1.5.1 则给出了变换矩阵 P 和 Q (或 Q^{-1})的几何意义.

有时我们也需要考虑只用初等行变换作用在矩阵上的所能达到的“最简型”问题. 这问题的几何意义是: 将 $A\in\mathbb{F}^{m\times n}$ 视为 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性映射: $x\mapsto Ax$, 在 $\mathbb{F}^n$ 中选标准基($P=I_n$)作为入口基, 在 $\mathbb{F}^m$ 中选 m 阶非奇异矩阵 $Q\in\mathbb{F}^{m\times m}$ 的一个列向量组作为出口基, 使线性映射 A 在这对基下的表示尽可能简单.

例 1.5.1 (用初等变换化矩阵为行阶梯型的几何证明)
记

$$A=[a_1 \quad a_2 \quad \cdots \quad a_n].$$

考虑 $\mathbb{F}^m$ 中的子空间序列

$$\{0\}\subseteq\operatorname{span}\{a_1\}\subseteq\operatorname{span}\{a_1,a_2\}\subseteq\cdots\subseteq\operatorname{span}\{a_1,a_2,\ldots,a_n\}=\operatorname{im}A.$$

将那些为严格包含关系处的 a_j 选出, 记为

$$a_{j_1},a_{j_2},\ldots,a_{j_r}.$$

这里, $r=\operatorname{rank}A$. 于是对 $1\le j<j_1$ (如果有), $a_j=0$; 对 $j_k\le j<j_{k+1}$, $k=1,\ldots,r-1$, a_j 可由 $a_{j_1},a_{j_2},\ldots,a_{j_k}$ 线性表示; 对 $j_r\le j\le n$ (如果有), a_j 可由 $a_{j_1},a_{j_2},\ldots,a_{j_r}$ 线性表示. 再将 $a_{j_1},a_{j_2},\ldots,a_{j_r}$ 扩充成 $\mathbb{F}^m$ 的一个基:

$$a_{j_1},a_{j_2},\ldots,a_{j_r},q_{r+1},\ldots,q_m.$$

则有

$$A=\Big[\underbrace{\overbrace{a_{j_1},a_{j_2},\ldots,a_{j_r}}^{Q_r},q_{r+1},\ldots,q_m}_{Q}\Big]\left[\begin{array}{c|c|c|c|c}
0\ \cdots\ 0 & 1\ *\ \cdots\ * & 0\ *\ \cdots\ *\ \cdots & 0\ *\ \cdots\ * & 0 \\ \hline
 & & 1\ *\ \cdots\ *\ \cdots & 0\ *\ \cdots\ * & 0 \\ \hline
 & & & \vdots & \vdots \\ \hline
 & & & 0 & 0 \\ \hline
 & & & 1\ *\ \cdots\ * & 0 \\ \hline
 & & & & 0_{m-r,n-r}
\end{array}\right].$$

其中，第 (i, j_i) 位置上的元素均为 1，$i=1,2,\ldots,r$.
还可以证明，若可逆矩阵 Q 化 A 为上述行阶梯型，则 Q 的前 $r=\operatorname{rank}A$ 列由 A 唯一决定的(按上面的子空间序列方法)；若 A 是行满秩的，则化 A 为行阶梯型的可逆矩阵 Q 还是唯一的.

1.5.2 矩阵相似

定义 1.5.2 (矩阵相似)
两个矩阵 $A,B\in\mathbb{F}^{n\times n}$ 称为相似，如果存在 n 阶非奇异矩阵 $P\in\mathbb{F}^{n\times n}$ 使得 $AP=PB$，或 $P^{-1}AP=B$.

注 1.5.3 (矩阵相似的几何意义)
将 $A\in\mathbb{F}^{n\times n}$ 视为 $\mathbb{F}^n$ 上的线性变换：$x\mapsto Ax$，将 n 阶非奇异矩阵 $P\in\mathbb{F}^{n\times n}$ 的列向量组视为 $\mathbb{F}^n$ 中的一个基. 由于出口空间和入口空间相同，都是 $\mathbb{F}^n$，所以可将入口基和出口基选得一样，其基矩阵都选为 P. 则定义 1.5.2 中的 B 恰为线性变换 A 在这对基下的矩阵表示. 特别的，线性变换 A 在入口基和出口基都是 $\mathbb{F}^n$ 的标准基时(即 $P=I_n$)的矩阵表示即为 A 自己.

问题 1.5.2 (矩阵的相似“最简型”)
给定矩阵 $A\in\mathbb{F}^{n\times n}$，找非奇异矩阵 $P\in\mathbb{F}^{n\times n}$，使得
$B=P^{-1}AP$
“尽可能简单”. 这问题的几何实质是：给定 $\mathbb{F}^n$ 上的线性变换：$x\mapsto Ax$，寻找 $\mathbb{F}^n$ 的基 $p_1,p_2,\ldots,p_n$，将它同时作为入口基和出口基，使得线性变换 A 在这对基下的矩阵表示“尽可能简单”.

定义 1.5.3 (方矩阵的不变子空间)
设矩阵 $A\in\mathbb{F}^{n\times n}$. 子空间 $W\subseteq\mathbb{F}^n$ 称为 A 的不变子空间，如果
$A(W)\subseteq W$.
这里，记号 $A(W)=\{Ax:x\in W\}$.

x属于W，则 A(x)属于W

例 1.5.2 (不变子空间的例)
设矩阵 $A\in\mathbb{F}^{n\times n}$. 容易验证，$A$ 的核 $\ker A$，A 的像 $\operatorname{im}A$，零空间 $\{0\}$，及全空间 $\mathbb{F}^n$ 都是 A 的不变子空间.

定理 1.5.2 (方阵的不变子空间与方阵的块三角化两事物的等同性)
设 $P^{-1}AP=B$，且 P，B 相应的分块如下
$$P=[P_1 \mid P_2],\quad B=\left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22}\end{array}\right].$$
则
(1) $B_{21}=0$，当且仅当 $\operatorname{im}P_1$ 为 A 的不变子空间.
(2) $B_{12}=0$，当且仅当 $\operatorname{im}P_2$ 为 A 的不变子空间.
(3) $B_{21}=0$，同时 $B_{12}=0$，当且仅当 $\operatorname{im}P_1$ 及 $\operatorname{im}P_2$ 均为 A 的不变子空间(从而 $\operatorname{im}P_1$ 及 $\operatorname{im}P_2$ 为互补的子空间).

证明：记

$$P=\left[\underbrace{p_1 \ \cdots \ p_{n_1}}_{P_1} \;\vdots\; \underbrace{p_{n_1+1} \ \cdots \ p_n}_{P_2}\right].$$

将 $P^{-1}AP=B$ 写成 $AP=PB$:

$$A[p_1 \ \cdots \ p_{n_1} \mid p_{n_1+1} \ \cdots \ p_n]=[p_1 \ \cdots \ p_{n_1} \mid p_{n_1+1} \ \cdots \ p_n]\left[\begin{array}{ccc|ccc} b_{11} & \cdots & b_{1,n_1} & b_{1,n_1+1} & \cdots & b_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{n_1,1} & \cdots & b_{n_1,n_1} & b_{n_1,n_1+1} & \cdots & b_{n_1,n} \\ \hline b_{n_1+1,1} & \cdots & b_{n_1+1,n_1} & b_{n_1+1,n_1+1} & \cdots & b_{n_1+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{n,1} & \cdots & b_{n,n_1} & b_{n,n_1+1} & \cdots & b_{n,n} \end{array}\right]$$

(1) 若 $B_{21}=0$，则对 $j=1,\ldots,n_1$，都有

$$\begin{aligned} Ap_j &= p_1\cdot b_{1j}+\cdots+p_{n_1}\cdot b_{n_1,j}+p_{n_1+1}\cdot 0+\cdots+p_n\cdot 0 \\ &\in \operatorname{im} P_1, \end{aligned}$$

因此，$A(\operatorname{im} P_1)\subseteq \operatorname{im} P_1$，即 $\operatorname{im} P_1$ 为 A 的不变子空间．反之，若 $\operatorname{im} P_1$ 为 A 的不变子空间，即 $A(\operatorname{im} P_1)\subseteq \operatorname{im} P_1$，则对 $j=1,\ldots,n_1$，由

$$Ap_j = p_1\cdot b_{1j}+\cdots+p_{n_1}\cdot b_{n_1,j}+p_{n_1+1}\cdot b_{n_1+1,j}+\cdots+p_n\cdot b_{n,j}$$

知

$$\begin{aligned} p_{n_1+1}\cdot b_{n_1+1,j}+\cdots+p_n\cdot b_{n,j} &= Ap_j-(p_1\cdot b_{1j}+\cdots+p_{n_1}\cdot b_{n_1,j}) \\ &\in \operatorname{im} P_1. \end{aligned}$$

又显然

$$p_{n_1+1}\cdot b_{n_1+1,j}+\cdots+p_n\cdot b_{n,j}\in \operatorname{im} P_2,$$

因此,

$$p_{n_1+1}\cdot b_{n_1+1,j}+\cdots+p_n\cdot b_{n,j}\in (\operatorname{im} P_1)\cap(\operatorname{im} P_2)=\{0\}. \text{(注意 } \operatorname{im} P_1\oplus \operatorname{im} P_2=\mathbb{F}^n\text{)}$$

由 P 的列线性无关，知

$$b_{n_1+1,j}=\cdots=b_{n,j}=0,\quad j=1,\ldots,n_1,$$

此即 $B_{21}=0$.

(2)及(3)的证明与(1)同理．证毕

注 1.5.4 (基于不变子空间将矩阵块三角化、块对角化的方法)

(1) 设子空间 $W\subseteq\mathbb{F}^n$ 是 n 阶方阵 A 的非平凡的不变子空间(即 $W\neq\{0\},W\neq\mathbb{F}^n$)，下面我们构造基矩阵 P，将 A 上三角化．设 $\dim W=n_1$，任选 W 的基 $p_1,\ldots,p_{n_1}$，并扩充为 $\mathbb{F}^n$ 的基如下(这等价于找 W 的一个补子空间):

$$p_1,\ldots,p_{n_1},p_{n_1+1},\ldots,p_n,$$

拼成的基矩阵记为 P，则 $P^{-1}AP$ 即为上三角矩阵;

(2) 设子空间 $W,U\subseteq\mathbb{F}^n$ 是 n 阶方阵 A 的两个互补的非平凡的不变子空间，即 $W\oplus U=\mathbb{F}^n$，则分别任选 W 和 U 的基，它们合成全空间 $\mathbb{F}^n$ 的基．用该基拼成的基矩阵作相似变换，即可把 A 块对角化.

例 1.5.3 (基于不变子空间 $\ker A$ 和 $\operatorname{im} A$ 将 A 块三角化)

设矩阵 $A\in\mathbb{F}^{n\times n}$，$\operatorname{rank} A=r$.

(1) $\ker A$ 的维数 $\dim(\ker A)=n-r$．取 $\ker A$ 的基 $p_1,\ldots,p_{n-r}$，并扩充为 $\mathbb{F}^n$ 的基如下:

$$\underbrace{p_1,\ldots,p_{n-r}}_{P_1},\underbrace{p_{n-r+1},\ldots,p_n}_{P_2},$$

则

$$A[P_1 \mid P_2]=[P_1 \mid P_2]\left[\begin{array}{c|c} 0 & B_{12} \\ \hline 0 & B_{22} \end{array}\right].$$

(2) A 的像 $\mathrm{im}\,A$ 的维数 $\dim(\mathrm{im}\,A)=r$.取 $\mathrm{im}\,A$ 的基 $q_1,\ldots,q_r$，并扩充为 $\mathbb{F}^n$ 的基如下:

$$\underbrace{q_1,\ldots,q_r}_{Q_1},\underbrace{q_{r+1},\ldots,q_n}_{Q_2},$$

则

$$A[Q_1 \mid Q_2]=[Q_1 \mid Q_2]\left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline 0 & 0 \end{array}\right].$$

1.5.3 一维不变子空间(特征值与特征向量)

给定矩阵 $A\in\mathbb{F}^{n\times n}$. 设 $W=\mathrm{span}\{p\}, p\neq 0$，是 A 的一个不变子空间，这等价于说

$Ap\in\mathrm{span}\{p\}$,

或存在 $\lambda\in\mathbb{F}$ 使得

$Ap=p\cdot\lambda$.

定义 1.5.4 (特征值与特征向量)

给定矩阵 $A\in\mathbb{F}^{n\times n}$，数 $\lambda\in\mathbb{F}$，及非零向量 $p\in\mathbb{F}^n, p\neq 0$. 若有

$Ap=p\cdot\lambda$, (1.5.4)

则称 λ 是 A 的一个特征值， p 为相应于该特征值的一个特征向量，或更明确地，称数与向量的对 (λ,p) 为 A 的一个特征值与特征向量对.

孰知, (1.5.4)可写为矩阵形式

$(\lambda I_n-A)p=0$.

这里

$$\lambda I_n=\begin{bmatrix}\lambda & & \\ & \ddots & \\ & & \lambda\end{bmatrix},$$

其中，未写出的元素为 0.

反复运用定理 1.5.2, (3)，或直接验证，可得如下的

定理 1.5.3 (n 阶矩阵可相似对角化等价于存在 n 个互补的一维不变子空间，即 n 个线性无关的特征向量)

设矩阵 $P=[p_1\quad p_2\quad\cdots\quad p_n]$ 可逆， $A\in\mathbb{F}^{n\times n}$. 则

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}, \quad (1.5.5)$$

当且仅当每个 p_j 均为 A 的特征向量， $j=1,2,\ldots,n$.

证明: (1.5.5)等价于

$$A[p_1\quad p_2\quad\cdots\quad p_n]=[p_1\quad p_2\quad\cdots\quad p_n]\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},$$

这又等价于

$Ap_j=p_j\lambda_j, j=1,2,\ldots,n$. 证毕

1.5.6 可对角化矩阵的谱分解

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix},\quad P=\begin{bmatrix}p_1 & p_2 & \cdots & p_n\end{bmatrix},\quad P^{-1}=\begin{bmatrix}q_1\\ q_2\\ \vdots\\ q_n\end{bmatrix}$$

$$A=P\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}P^{-1}$$

$$=\begin{bmatrix}p_1 & p_2 & \cdots & p_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n\end{bmatrix}\begin{bmatrix}q_1\\ q_2\\ \vdots\\ q_n\end{bmatrix}$$

$$=\lambda_1 p_1 q_1+\lambda_2 p_2 q_2+\cdots+\lambda_n p_n q_n$$

称

$$A=\lambda_1 p_1 q_1+\lambda_2 p_2 q_2+\cdots+\lambda_n p_n q_n$$

为 A 的谱分解. 注意到 $p_1q_1, p_2q_2, \ldots, p_nq_n$ 是与直和分解

$$\mathbb{F}^n=\operatorname{im}p_1\oplus\operatorname{im}p_2\oplus\cdots\oplus\operatorname{im}p_n$$

相应的(斜)投影矩阵, 谱分解有如下的几何意义: 算子 A 的作用相当于向 A 的特征向量系作投影, 再按特征值加权.

但是, 一般情况下, 对一给定矩阵 $A\in\mathbb{F}^{n\times n}$, 我们甚至不能保证一定存在一个一维不变子空间(特征向量). 因此上面的定理自然不能期待. 然而如果在复数域 $\mathbb{C}$ 上, 我们有

引理 1.5.2 (复矩阵恒有一维不变子空间, 即特征向量)

对任意正整数 n, 及 n 阶复矩阵 $A\in\mathbb{C}^{n\times n}$, 存在 $\lambda\in\mathbb{C}$ 及非零向量 $p\in\mathbb{C}^n, p\neq 0$, 使得

$Ap=p\cdot\lambda$.

于是, 我们有下面著名的 Schur 定理

定理 1.5.4 (Schur 定理: 复矩阵恒能相似上三角化)

任给 n 阶复矩阵 $A\in\mathbb{C}^{n\times n}$, 存在 n 阶可逆复矩阵 $P\in\mathbb{C}^{n\times n}$, 使得

$$P^{-1}AP=\begin{bmatrix}\lambda_1 & & & *\\ 0 & \lambda_2 & & \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & \lambda_n\end{bmatrix}.$$

证明: 对 n 用数学归纳法. $n=1$ 时, 显然. 假设 $n\le m$ 时命题成立, 下证 $n=m+1$ 时也成立.

设 $A\in\mathbb{C}^{(m+1)\times(m+1)}$. 由于 $m+1$ 次复特征多项式 $|\lambda I-A|$ 在复数域 $\mathbb{C}$ 中总有根, 任取其一根, 记为 λ_1, 相应的一个特征向量记为 p_1, 则

$Ap_1=p_1\lambda_1$. (1.5.7)

将 p_1 扩充成 $\mathbb{C}^{(m+1)}$ 的一个基如下:

$p_1, q_2, \ldots, q_{m+1}$,

则

$$A\underbrace{\begin{bmatrix} p_1 & q_2 & \cdots & q_{m+1} \end{bmatrix}}_{R} = \begin{bmatrix} p_1 & q_2 & \cdots & q_{m+1} \end{bmatrix} \underbrace{\begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}}_{\begin{bmatrix} \lambda_1 & \alpha \\ 0 & B \end{bmatrix}}. \quad (1.5.8)$$

由 $B \in \mathbb{C}^{m\times m}$ 及归纳假设，存在 m 阶可逆复矩阵 Q 使得

$$Q^{-1}BQ = \begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_{m+1} \end{bmatrix}.$$

于是,

$$\begin{bmatrix} \lambda_1 & \alpha \\ 0 & B \end{bmatrix} = \left[\begin{array}{c|c} \lambda_1 & \alpha \\ \hline 0 & Q\begin{bmatrix} \lambda_2 & & * \\ & \ddots & \\ 0 & & \lambda_{m+1} \end{bmatrix}Q^{-1} \end{array}\right]$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \left[\begin{array}{c|ccc} \lambda_1 & & \alpha Q & \\ \hline & \lambda_2 & & * \\ 0 & & \ddots & \\ & 0 & & \lambda_{m+1} \end{array}\right] \begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix}.$$

由此及(1.5.8)，有

$$R^{-1}AR = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} \left[\begin{array}{c|ccc} \lambda_1 & & \alpha Q & \\ \hline & \lambda_2 & & * \\ 0 & & \ddots & \\ & 0 & & \lambda_{m+1} \end{array}\right] \begin{bmatrix} 1 & 0 \\ 0 & Q^{-1} \end{bmatrix},$$

从而

$$\left(R\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}\right)^{-1} A \left(R\begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix}\right) = \left[\begin{array}{c|ccc} \lambda_1 & & \alpha Q & \\ \hline & \lambda_2 & & * \\ 0 & & \ddots & \\ & 0 & & \lambda_{m+1} \end{array}\right]$$

即为上三角矩阵．证毕