线性子空间

定义 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（矩阵的核与像）

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 $Ax = 0$ 的解集合， 它是 $\mathbb{F}^n$ 的子空间； A 的像定义为

$$\operatorname{im} A = \{y : y \in \mathbb{F}^m, \text{且存在 } x \in \mathbb{F}^n \text{使得 } y = Ax\} = \{Ax : x \in \mathbb{F}^n\},$$

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（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

$$\mathrm{span}\{\alpha_1,\alpha_2,\ldots,\alpha_t\}=\left\{\alpha_1k_1+\alpha_2k_2+\cdots+\alpha_tk_t : k_i\in\mathbb{F}, i=1,2,\ldots,t\right\},$$

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

$$W=\mathrm{span}\{\beta_1,\beta_2,\ldots,\beta_r\},$$

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．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.

进而，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 $\operatorname{im}A$.

定义（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

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

及

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

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

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

再从生成组中筛出极大线性无关子组(How?)，即为V_1+V_2的基.

定义 （子空间的直和）

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（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

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

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

定义（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（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：（对$V=\mathbb{F}^n$给出具体找法;对一般的V用坐标法）

线性映射

定义（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（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.

例（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

$$\mathcal{A}_A:x\mapsto Ax.$$

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

$$A=\begin{bmatrix}\mathcal{A}(\varepsilon_1) & \mathcal{A}(\varepsilon_2) & \cdots & \mathcal{A}(\varepsilon_n)\end{bmatrix},$$

也就是 $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_n$ 的像 $\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\ldots,\mathcal{A}(\varepsilon_n)$（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

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

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（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

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

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

$$A=\begin{bmatrix}a_1 & a_2 & \cdots & a_n\end{bmatrix}=\begin{bmatrix}a_{ij}\end{bmatrix}_{m\times n}$$

称为 $\mathcal{A}$ 在相应的入口基和出口基下的表示.

注（线性映射的矩阵表示定义的公式化）

记

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

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

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

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.

定理（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：

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

例（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．

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

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.

$$f(x)=5x^3+2x^2+1$$

$$=\begin{bmatrix}1&x&x^2&x^3\end{bmatrix}\begin{bmatrix}1\\0\\2\\5\end{bmatrix}.$$

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.

$$D\begin{bmatrix}1\\0\\2\\5\end{bmatrix}=\begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\end{bmatrix}\begin{bmatrix}1\\0\\2\\5\end{bmatrix}$$

$$=\begin{bmatrix}0\\4\\15\end{bmatrix}.$$

第三步：与出口基作线性组合，得$\mathcal{D}(f(x))$.

例（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.

$$\mathcal{A}[e_x \quad e_y \quad e_z] = [\cos\theta\cdot e_x + \sin\theta\cdot e_y \quad -\sin\theta\cdot e_x + \cos\theta\cdot e_y \quad e_z]$$
$$= [e_x\cos\theta + e_y\sin\theta \quad e_x(-\sin\theta) + e_y\cos\theta \quad e_z]$$
$$= [e_x \quad e_y \quad e_z]\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

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（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.

$$\mathcal{B}\begin{bmatrix} e_x & e_y & e_z \end{bmatrix} = \begin{bmatrix} e_x & e_y & -e_z \end{bmatrix}$$

$$= \begin{bmatrix} e_x & e_y & e_z \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

因此，$\mathcal{B}$的矩阵表示为

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

矩阵的等价与相似

矩阵等价

定义（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（矩阵等价的几何意义）

将 $A\in\mathbb{F}^{m\times n}$ 视为 $\mathbb{F}^n$ 到 $\mathbb{F}^m$ 的线性映射：$x\mapsto Ax$，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)．则定义中的 B 恰为线性映射 A 在这对基下的矩阵表示．特别的，线性映射 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.

问题（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： $x \mapsto Ax$，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 A 在这对基下的矩阵表示“尽可能简单”.

注（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

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

现在，变换矩阵 P 和 Q（或 Q^{-1}）有了新的意义（几何意义）：??

注（线性系统输入-输出解耦）

矩阵相似

定义（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（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（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（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 $A(W) = \{Ax : x \in W\}$

例（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（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

$$P = \left[P_1 \;\vdots\; P_2 \right], \quad B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right].$$

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).

注（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（基于不变子空间 $\ker A$ 和 $\mathrm{im}\, A$ 将 A 块三角化）

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

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

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

则

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

(2) A 的像 $\operatorname{im}A$ 的维数 $\dim(\operatorname{im}A)=r$. 取 $\operatorname{im}A$ 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

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

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 $W=\operatorname{span}\{p\}, p\neq 0$, 是 A 的一个不变子空间，这等价于说

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

或存在 $\lambda\in\mathbb{F}$ 使得

$$Ap=p\cdot\lambda.$$

定义（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

$$Ap=p\cdot\lambda,$$

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

熟知，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，为写出的元素为 0.

反复运用定理，或直接验证，可得如下的

定理（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 $P=\begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix}$可逆，$A \in \mathbb{F}^{n\times n}$．则

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

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

证明：等价于

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

这又等价于

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

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（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（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

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