第七章 线性变换

线性变换是最简单最基本的变换，是线性代数研究的主要对象之一，它从映射的角度来讨论线性空间的向量之间的内在联系及线性空间的结构.通过线性变换的学习，有利于提高对几何空间的旋转、投影、反射以及分析中的微商变换的实质.

从同态的角度来看，线性变换是线性空间到自身的一个同态映射，而同态映射是高等代数的后继课程近世代数的一种重要方法，因此，线性变换是同态映射的特例，是学习近世代数的基础.

线性变换是线性空间到自身的映射，实际上是线性空间的一个线性算子，是一般线性算子的特殊情形.所以通过线性变换的学习，将为以后学习线性泛函、算子理论等要做必要的准备.

7.1 线性变换的定义

教学的时间：**年**月**日** **教学学时数:2 学时**

一、教学目标

1. 理解线性变换的定义；
2. 掌握线性变换的性质.

二、教学重点

线性变换的定义及其性质

三、教学难点

线性变换的定义

四、教学过程

（一）引入

上一章我们看到，数域 P 上任意一个 n 维线性空间都与 P^n 同构，因之，有限维线性空间的结构可以认为是完全清楚了.线性空间是某一事物从量的方面的一个抽象.我们认识客观事物，固然要弄清它们单个的和总体的性质，但是更重要的是研究它们之前各种各样的联系.在线性空间中，事物之间的联系就反映为线性空间的映射.线性空间 V 到自身的映射通常称为 V 的一个变换.

这一章中要讨论的线性变换就是最简单的，同时也可以认为是最基本的一种变换，正如线性函数是最简单和最基本的函数一样.线性变换是线性代数的一个主要研究对象.

下面如果不特别声明，所考虑的都是某一固定的数域 P 上的线性空间.

（二）线性变换的定义

在讨论线性空间的同构时，我们考虑的是一种保持向量的加法和数量乘法的一一对应. 我们常称**两线性空间之间保持加法和数量乘法的映射为线性映射**. 本节要讨论的是在线性空间 V 上的线性映射——线性变换.

线性空间 V 到自身的映射称为 V 的一个变换.

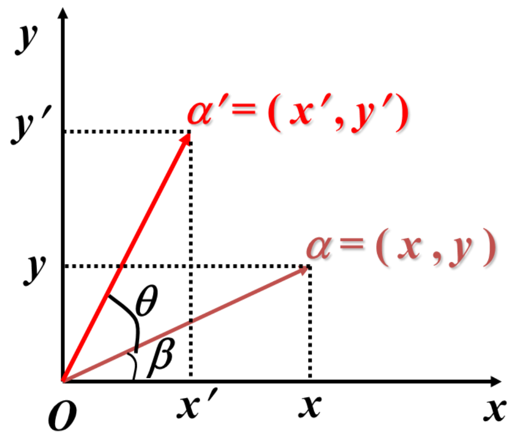
定义 1 线性空间 V 的一个变换 $\mathscr{A}$ 称为线性变换，如果对于 V 中任意的元素 α,β 和数域 P 中任意数 k，都有

$$\mathscr{A}(\alpha+\beta)=\mathscr{A}(\alpha)+\mathscr{A}(\beta);\mathscr{A}(k\alpha)=\mathscr{A}k(\alpha). \qquad (1)$$

用花体拉丁字母 $\mathscr{A}$，$\mathscr{B}$，…表示 V 的线性变换，$\mathscr{A}(\alpha)$ 或 $\mathscr{A}\alpha$ 代表元素 α 在变换 $\mathscr{A}$ 下的像.

定义中等式(1)所表示的性质，有时也说成**线性变换保持向量的加法与数量乘法.**

例 1 平面上的向量构成实数域上的二维线性空间. 把平面围绕坐标原点按反时针方向旋转 θ 角，用 $\mathscr{I}_\theta$ 表示. 若平面上一个向量 α 在直角坐标系下的坐标是 (x,y)，像 $\mathscr{I}_\theta(\alpha)$ 的坐标 (x',y').



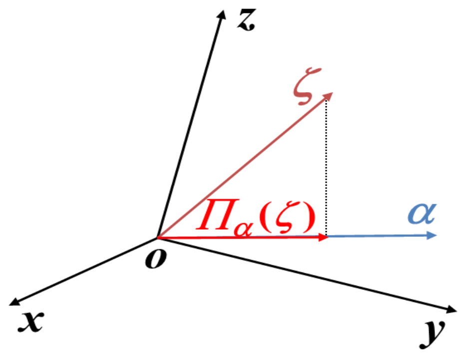
问题 1 求 $\mathscr{I}_\theta(\alpha)=\alpha'$ 的坐标.

设 α 旋转 θ 角之后的坐标（x'，y'），则按照公式 $\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}$.

问题 2 旋转变换 $\mathscr{I}_\theta(\alpha)=\alpha'$ 是不是线性变换？

同样地，空间中绕轴的旋转也是一个线性变换.

例 2 设 α 是几何空间中一固定非零向量，把每个向量 ξ 变到它在 α 上的内射影的变换，以 $\prod_\alpha$ 表示它.



问题 1 求证 $\prod_\alpha(\xi)=\dfrac{(\alpha,\xi)}{(\alpha,\alpha)}\alpha$，$(\alpha,\xi),(\alpha,\alpha)$ 表示内积.

问题 2 $\prod_\alpha$ 也是线性变换.

例 3 线性空间 V 中的恒等变换或称单位变换 E，即

$$\mathrm{E}(\alpha)=\alpha \qquad (\alpha\in V)$$

以及零变换$\mathcal{O}$，即

$$\mathcal{O}(\alpha)=0 \quad (\alpha\in V)$$

都是线性变换.

例 4 设V是数域P上的线性空间，k是P中的某个数，定义V的变换如下：

$$\mathrm{K}:\ \alpha\to k\alpha, \qquad \alpha\in V.$$

这是一个线性变换，称为由数k决定的数乘变换，可用 K 表示. 显然当$k=1$时，便得恒等变换，当$k=0$时，便得零变换.

例 5 在线性空间$P[x]$或者$P[x]_n$中，求导数是一个线性变换. 这个变换通常用 D 代表，即

D（$f(x)$）= $f'(x)$.

例 6 定义在闭区间$[a,b]$上的全体连续函数组成实数域上一线性空间，以$C(a,b)$代表. 在这个空间中变换

$$\mathcal{J}(f(x))=\int_a^x f(t)dt$$

是一线性变换.

（三）线性变换的简单性质

1．设 A 是V的线性变换，则$\mathcal{A}(0)=0$，$\mathcal{A}(-\alpha)=-\mathcal{A}(\alpha)$.

2．线性变换保持线性组合与线性关系式不变. 换句话说，如果β是$\alpha_1,\alpha_2,\cdots,\alpha_r$的线性组合：

$$\beta=k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r,$$

那么经过线性变换 A 之后，$\mathcal{A}(\beta)$是$\mathcal{A}(\alpha_1),\mathcal{A}(\alpha_2),\cdots,\mathcal{A}(\alpha_r)$同样的线性组合：

$$\mathcal{A}(\beta)=k_1\mathcal{A}(\alpha_1)+k_2\mathcal{A}(\alpha_2)+\cdots+k_r\mathcal{A}(\alpha_r).$$

又如果$\alpha_1,\alpha_2,\cdots,\alpha_r$之间有一线性关系式

$$k_1\alpha_1+k_2\alpha_2+\cdots+k_r\alpha_r=0$$

那么它们的像之间也有同样的关系式

$$k_1\mathcal{A}(\alpha_1)+k_2\mathcal{A}(\alpha_2)+\cdots+k_r\mathcal{A}(\alpha_r)=0.$$

3．线性变换把线性相关的向量组变成线性相关的向量组.

但是线性变换把线性无关的向量组有可能变成线性相关的向量组，也有可能变成线性无关的向量组.比如零变换把线性空间的任意一组基变成了零向量，从而零变换把线性无关的向量组变成了线性无关的向量组. 单位变换把线性空间的任意一组基变成这组基，从而单位变换把线性无关的向量组变成线性无关的向量组.

五、小结

1. 线性变换的定义；

2.掌握线性变换的性质.

六、作业（习题册第 152—157 题，160、161、165—171,173,174、187、188、210）

七、教学反思

7.2 线性变换的运算

教学的时间：**年**月**日　　教学学时数:2 学时**

一、教学目标

1. 掌握线性变换的运算及其性质

2. 理解线性变换的多项式.

二、教学重点

线性变换的运算

三、教学难点

线性变换的乘法运算

四、教学过程

一、线性变换的乘法

1. 线性变换乘积的定义

设 $\mathcal{A}$，$\mathcal{B}$ 是线性空间 V 的两个线性变换，定义它们的乘积为.

$$(\mathcal{AB})(\alpha)=\mathcal{A}(\mathcal{B}(\alpha)) \qquad (\alpha\in V).$$

2. 线性变换乘积的性质

（1）线性变换的乘积也是线性变换.

证明 由线性变换乘积的定义知，$\mathcal{AB}$ 是 V 上的一个变换.

且 $\forall\alpha,\beta\in V,\forall k\in P,$ 有

$$\begin{aligned}(\mathcal{AB})(\alpha+\beta)&=\mathcal{A}(\mathcal{B}(\alpha+\beta))\\&=\mathcal{A}(\mathcal{B}(\alpha)+\mathcal{B}(\beta))=\mathcal{A}(\mathcal{B}(\alpha))+\mathcal{A}(\mathcal{B}(\beta))\\&=(\mathcal{AB})(\alpha)+(\mathcal{AB})(\beta)\end{aligned}$$

$$(\mathcal{AB})(k\alpha)=\mathcal{A}(\mathcal{B}(k\alpha))=\mathcal{A}(k\,\mathcal{B}(\alpha))=k\,\mathcal{A}(\mathcal{B}(\alpha))=k\,(\mathcal{AB})(\alpha)$$

线性变换 $\mathcal{A}$，$\mathcal{B}$ 的乘积 $\mathcal{AB}$ 也是线性变换.

例如，在实数域上的线性空间 $R[x]$ 中，线性变换

$$\mathrm{D}(f(x))=f'(x).$$

$$\mathcal{J}(f(x))=\int_a^x f(t)dt$$

的乘积 D $\mathcal{J}$=$\mathcal{E}$，但一般 $\mathcal{J}$D≠$\mathcal{E}$.

一般情况下，线性变换的乘积不满足交换律.

（2）线性变换的乘法适合结合律，即$(\mathcal{AB})\mathcal{C}=\mathcal{A}(\mathcal{BC})$.

（3）对于任意线性变换 $\mathcal{A}$，都有 $\mathcal{AE}=\mathcal{EA}=\mathcal{A}$.

二、线性变换的加法

1. **线性变换的加法的定义**

设 $\mathcal{A}$，$\mathcal{B}$ 是线性空间V的两个线性变换，定义它们的和 $\mathcal{A}+\mathcal{B}$ 为

$$(\mathcal{A}+\mathcal{B})(\alpha)=\mathcal{A}(\alpha)+\mathcal{B}(\alpha)\qquad(\alpha\in V).$$

易证：线性变换 $\mathcal{A}$，$\mathcal{B}$ 的和 $\mathcal{A}+\mathcal{B}$ 还是线性变换.

$\forall\alpha,\beta\in V,\forall k\in P,$有 $\qquad(\mathcal{A}+\mathcal{B})(\alpha+\beta)=\mathcal{A}(\alpha+\beta)+\mathcal{B}(\alpha+\beta)$

$=\mathcal{A}(\alpha)+\mathcal{A}(\beta)+\mathcal{B}(\alpha)+\mathcal{B}(\beta)$

$=\mathcal{A}(\alpha)+\mathcal{B}(\alpha)+\mathcal{A}(\beta)+\mathcal{B}(\beta)$

$=(\mathcal{A}+\mathcal{B})(\alpha)+(\mathcal{A}+\mathcal{B})(\beta)$

$(\mathcal{A}+\mathcal{B})(k\alpha)=\mathcal{A}(k\alpha)+\mathcal{B}(k\alpha)=k\mathcal{A}(\alpha)+k\mathcal{B}(\alpha)=k(\mathcal{A}(\alpha)+\mathcal{B}(\alpha))=k(\mathcal{A}+\mathcal{B})(\alpha)$

线性变换 $\mathcal{A}$，$\mathcal{B}$ 的和 $\mathcal{A}+\mathcal{B}$ 也是线性变换.

2. 线性变换的加法适合结合律与交换律，即

$$\mathcal{A}+(\mathcal{B}+\mathcal{C})=(\mathcal{A}+\mathcal{B})+\mathcal{C}.$$

$$\mathcal{A}+\mathcal{B}=\mathcal{B}+\mathcal{A}.$$

3. 对于加法，零变换 $\mathcal{O}$ 与所有线性变换 $\mathcal{A}$ 的和仍等于 $\mathcal{A}$:

$$\mathcal{A}+\mathcal{O}=\mathcal{A}.$$

4. 对于每个线性变换 $\mathcal{A}$，可以定义它的负变换（$-\mathcal{A}$）：$(-\mathcal{A})(\alpha)=-\mathcal{A}(\alpha)\qquad(\alpha\in V)$.

则负变换（$-\mathcal{A}$）也是线性变换，且 $\mathcal{A}+(-\mathcal{A})=\mathcal{O}$.

5. 线性变换的乘法对加法有左右分配律，即 $\mathcal{A}(\mathcal{B}+\mathcal{C})=\mathcal{AB}+\mathcal{AC}$，$(\mathcal{B}+\mathcal{C})\mathcal{A}=\mathcal{BA}+\mathcal{CA}$.

三、线性变换的数量乘法

1. 数域 P 中的数 k 与线性变换 $\mathscr{A}$ 的数量乘法定义为

$$k\mathscr{A}=\mathrm{K}\mathscr{A}$$

即 $k\mathscr{A}(\alpha)=\mathrm{K}(\mathscr{A}(\alpha))=\mathrm{K}\mathscr{A}(\alpha)$，当然 $k\mathscr{A}$ 还是线性变换.

2. 线性变换的数量乘法适合以下的规律：

$$(kl)\mathscr{A}=k(l\mathscr{A}),$$

$$(k+l)\mathscr{A}=k\mathscr{A}+l\mathscr{A},$$

$$k(\mathscr{A}+\mathscr{B})=k\mathscr{A}+k\mathscr{B},$$

$$1\mathscr{A}=\mathscr{A}.$$

线性空间 V 上全体线性变换，对于如上定义的加法与数量乘法，也构成数域 P 上一个线性空间.

四、线性变换的逆变换

1. V 的变换 $\mathscr{A}$ 称为可逆的，如果有 V 的变换 $\mathscr{B}$ 存在，使 $\mathscr{A}\mathscr{B}=\mathscr{B}\mathscr{A}=\mathrm{E}$. 变换 $\mathscr{B}$ 称为 $\mathscr{A}$ 的逆变换，记为 $\mathscr{A}^{-1}$.

2. 如果线性变换 $\mathscr{A}$ 是可逆的，那么它的逆变换 $\mathscr{A}^{-1}$ 也是线性变换.

证明 $\forall\alpha,\beta\in V,\forall k\in P$，有 $\mathscr{A}^{-1}(\alpha+\beta)=\mathscr{A}^{-1}((\mathscr{A}\mathscr{A}^{-1})(\alpha)+(\mathscr{A}\mathscr{A}^{-1})(\beta))$

$=\mathscr{A}^{-1}\{\mathscr{A}[\mathscr{A}^{-1}(\alpha)]+\mathscr{A}[\mathscr{A}^{-1}(\beta)]\}=\mathscr{A}^{-1}\{\mathscr{A}[\mathscr{A}^{-1}(\alpha)]+\mathscr{A}[\mathscr{A}^{-1}(\beta)]\}$

$=\mathscr{A}^{-1}\{\mathscr{A}[\mathscr{A}^{-1}(\alpha)+\mathscr{A}^{-1}(\beta)]\}=(\mathscr{A}^{-1}\mathscr{A})[\mathscr{A}^{-1}(\alpha)+\mathscr{A}^{-1}(\beta)]=\mathscr{A}^{-1}(\alpha)+\mathscr{A}^{-1}(\beta)$

$\mathscr{A}^{-1}(k\alpha)=\mathscr{A}^{-1}[k(\mathscr{A}\mathscr{A}^{-1})(\alpha)]=\mathscr{A}^{-1}\{k\mathscr{A}[\mathscr{A}^{-1}(\alpha)]\}=(\mathscr{A}^{-1}\mathscr{A})[k\mathscr{A}^{-1}(\alpha)]\}=k\mathscr{A}^{-1}(\alpha)$

因此，逆变换 $\mathscr{A}^{-1}$ 也是线性变换.

五、线性变换的多项式

既然线性变换的乘法满足结合律，当若干个线性变换 $\mathscr{A}$ 重复相乘时，其最终结果是完全确定的，与乘法的结合方法无关. 因此当 n 个（n 是正整数）线性变换 $\mathscr{A}$ 相乘时，就可以用 $\overbrace{AA\cdots A}^{n个}$ 来表示，称为 $\mathscr{A}$ 的 n 次幂，简记为 $\mathscr{A}^n$. 作为定义.

令

$$\mathcal{A}^0=\mathcal{E}.$$

根据线性变换幂的定义，可以推出指数法则：

$$\mathcal{A}^{m+n}=\mathcal{A}^m\mathcal{A}^n,\ (\mathcal{A}^m)^n=\mathcal{A}^{mn}\ (m,n\geq 0)$$

当线性变换 $\mathcal{A}$ 可逆时，定义 $\mathcal{A}$ 的负整数幂为

$$\mathcal{A}^{-n}=(\mathcal{A}^{-1})^n\ (n\text{ 是正整数}).$$

值得注意的是，线性变换乘积的指数法则不成立，即一般说来

$$(\mathcal{A}\mathcal{B})^n\neq\mathcal{A}^n\mathcal{B}^n.$$

设

$$f(x)=a_mx^m+a_{m-1}x^{m-1}+\cdots+a_0$$

是 $P[x]$ 中的一个多项式，$\mathcal{A}$ 是 V 的一个线性变换，定义

$$f(\mathcal{A})=a_m\mathcal{A}^m+a_{m-1}\mathcal{A}^{m-1}+\cdots+a_0\mathcal{E}$$

显然 $f(\mathcal{A})$ 是一线性变换，它称为线性变换 $\mathcal{A}$ 的多项式.

不难验证，如果在 $P[x]$ 中

$$h(x)=f(x)+g(x)\ ,p(x)=f(x)g(x),$$

那么

$$h(\mathcal{A})=f(\mathcal{A})+g(\mathcal{A}),\quad p(\mathcal{A})=f(\mathcal{A})g(\mathcal{A}).$$

特别地，

$$f(\mathcal{A})g(\mathcal{A})=g(\mathcal{A})f(\mathcal{A}).$$

即**同一个线性变换的多项式的乘法是可交换的**.

例 1 在三维几何空间中，对于某一向量 α 的内射影 $\prod_\alpha$ 是一个线性变换. ς 在以 α 为法向量的平面 x 上的内射影 $\prod_x(\zeta)$. ς 对于平面 x 的反射 $\mathcal{R}_x$

问题 1 求 $\prod_x(\zeta)$，$\mathcal{R}_x(\zeta)$ 的计算公式？

问题 2 求 $\prod_x(\zeta)$，$\mathcal{R}_x(\zeta)$ 是不是线性变换？

$\prod_\alpha$ 可以用下面的公式来表示：

$$\prod_\alpha(\xi)=\frac{(\alpha,\xi)}{(\alpha,\alpha)}\alpha.$$

其中 $(\alpha,\xi),(\alpha,\alpha)$ 表示向量的内积.

ζ 在以 α 为法向量的平面 x 上的内射影 $\prod_x(\zeta)$

$$\prod_x(\zeta)=\zeta-\prod_\alpha(\zeta)$$

表示. 因此

$$\prod_x=\mathscr{E}-\prod_\alpha.\ \text{这里}\mathscr{E}\text{是恒等变换.}$$

ζ 对于平面 x 的反射 $\mathscr{R}_x$ 也是一个线性变换，它的像由公式

$$\mathscr{R}_x(\zeta)=\zeta-2\prod_\alpha(\zeta)$$

给出. 因此

$$\mathscr{R}_x=\mathscr{E}-2\prod_\alpha.$$

设 α,β 是空间的两个向量. 显然，α 与 β 互相垂直的充要条件为

$$\prod_\alpha\cdot\prod_\beta=\mathscr{O}$$

例 2 在线性空间 $P[\lambda]_n$ 中，求微商是一个线性变换，用 $\mathscr{D}$ 表示. 显然有

$$\mathscr{D}^n=\mathscr{O}.$$

其次，变换的平移

$$f(\lambda)\to f(\lambda+a)\quad a\in P$$

也是一个线性变换，用 $\mathscr{G}_a$ 表示. 根据泰勒展开式

$$f(\lambda+a)=f(\lambda)+af'(\lambda)+\frac{a^2}{2!}f''(\lambda)+\cdots+\frac{a^{n-1}}{(n-1)!}f^{(n-1)}(\lambda),$$

因之 $\mathscr{G}_a$ 实质上是 $\mathscr{D}$ 的多项式：

$$\mathscr{G}_a=\mathscr{E}+a\mathscr{D}+\frac{a^2}{2!}\mathscr{D}^2+\cdots+\frac{a^{n-1}}{(n-1)!}\mathscr{D}^{n-1}.$$

五、小结

1.线性变换的运算有哪些？他们还是线性变换吗？

2.线性变换的运算有哪些性质？

六、作业（习题册）

七、教学反思

7.3 线性变换的矩阵

教学的时间：**年**月**日　　　　　　教学学时数:4 学时**

一、教学目标

1.掌握线性变换在某个基下的矩阵、矩阵相似的概念;

2.掌握同一个线性变换在不同基下的矩阵的关系.

二、教学重点

1.线性线性变换在某个基下的矩阵的概念;

2.同一个线性变换在不同基下的矩阵的关系.

三、教学难点

同一个线性变换在不同基下的矩阵的关系

四、教学过程

（一）、线性变换关于基的矩阵

设V是数域P上n维线性空间. $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ V的一组基，现在建立线性变换与矩阵关系.

空间V中任意一个向量ξ可以被基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$线性表出，即有关系式

$$\xi = x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n \qquad (1)$$

其中系数是唯一确定的，它们就是ξ在这组基下的坐标. 由于线性变换保持线性关系不变，因而在ξ的像$\mathcal{A}\xi$与基的像$\mathcal{A}\varepsilon_1$，$\mathcal{A}\varepsilon_2$，⋯，$\mathcal{A}\varepsilon_n$之间也必然有相同的关系：

$$\mathcal{A}\xi = \mathcal{A}（x_1\varepsilon_1 + x_2\varepsilon_2 + \cdots + x_n\varepsilon_n）$$

$$= x_1\mathcal{A}(\varepsilon_1) + x_2\mathcal{A}(\varepsilon_2) + \cdots + x_n\mathcal{A}(\varepsilon_n) \qquad (2)$$

上式表明，如果知道了基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$的像，那么线性空间中任意一个向量$\xi$的像也就知道了，或者说

1. 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间$V$的一组基,如果线性变换$\mathcal{A}$与$\mathcal{B}$在这组基上的作用相同，即$\mathcal{A}\varepsilon_i = \mathcal{B}\varepsilon_i$，　$i = 1, 2, \cdots, n$,那么$\mathcal{A}=\mathcal{B}$.

结论 1 的意义就是，一个线性变换完全被它在一组基上的作用所决定. 下面指出，基

向量的像却完全可以是任意的，也就是

2. 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间$V$的一组基，对于任意一组向量$\alpha_1,\alpha_2,\cdots,\alpha_n$一定有一个线性变换$\mathcal{A}$使

$$\mathcal{A}\varepsilon_i=\alpha_i \qquad i=1,2,\cdots,n.$$

这是因为$\forall\xi\in V$，有$\xi=x_1\varepsilon_1+x_2\varepsilon_2+\cdots+x_n\varepsilon_n$.

令 V 上的变换$A\xi=x_1\alpha_1+x_2\alpha_2+\cdots+x_n\alpha_n$，下证$\mathcal{A}$是线性变换即可.

定理 1 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是线性空间$V$的一组基，$\alpha_1,\alpha_2,\cdots,\alpha_n$是$V$中任意$n$个向量. 存在唯一的线性变换$\mathcal{A}$使

$$\mathcal{A}\varepsilon_i=\alpha_i \qquad i=1,2,\cdots,n.$$

定义 2 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$是数域$P$上$n$维线性空间$V$的一组基，$\mathcal{A}$是$V$中的一个线性变换.

基向量的像可以被基线性表出：$\begin{cases}A\varepsilon_1=a_{11}\varepsilon_1+a_{21}\varepsilon_2+\cdots+a_{n1}\varepsilon_n\ ,\\ A\varepsilon_2=a_{12}\varepsilon_1+a_{22}\varepsilon_2+\cdots+a_{n2}\varepsilon_n\ ,\\ \qquad\cdots\cdots\cdots\cdots\cdots\cdots\\ A\varepsilon_n=a_{1n}\varepsilon_1+a_{2n}\varepsilon_2+\cdots+a_{nn}\varepsilon_n\ .\end{cases}$

用矩阵表示就是

$$\mathcal{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\mathcal{A}(\varepsilon_1),\mathcal{A}(\varepsilon_2),\cdots,\mathcal{A}(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A \qquad (5)$$

其中$A=\begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdots & \cdots & \cdots & \cdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$，矩阵$A$称为线性变换$\mathcal{A}$在基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$下的矩阵.

单位变换在任意一组基下的矩阵是单位矩阵，零变换在任意一个基下的矩阵是零矩阵.

例 1 设$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m$是$n\ (n>m)$维线性空间$V$的子空间$W$的一组基，把它扩充为$V$的一组基$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$. 指定线性变换$\mathcal{A}$如下

$$\begin{cases}A\varepsilon_i=\varepsilon_i\ ,i=1,2,\cdots,m,\\ A\varepsilon_i=0\ ,\ i=m+1,\cdots,n.\end{cases}$$

如此确定的线性变换$\mathcal{A}$为子空间W的一个投影. 不难证明

$$\mathcal{A}^2=\mathcal{A}$$

投影 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是

$$\begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & 1 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{pmatrix}$$

这样，在取定一组基之后，就建立了由数域 P 上的 n 维线性空间 V 的线性变换到数域 P 上的 $n\times n$ 矩阵的一个映射. 前面结论 1 说明这个映射是单射，结论 2 说明这个映射是满射. 换句话说，在这二者之间建立了一个双射. 这个对应的重要性表现在它保持运算，即有

定理 2 设 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是数域 P 上 n 维线性空间 V 的一组基，在这组基下，每个线性变换按公式（5）对应一个 $n\times n$ 矩阵，这个对应具有以下性质：

1）线性变换的和对应于矩阵的和；

2）线性变换的乘积对应于矩阵的乘积；

3）线性变换的数量乘积对应于矩阵的数量乘积；

4）可逆的线性变换与可逆矩阵对应，且逆变换对应于逆矩阵.

定理 2 说明数域 P 上 n 维线性空间 V 的全体线性变换组成的集合 $L(V)$ 对于线性变换的加法与数量乘法构成 P 上一个线性空间，与数域 P 上 n 级方阵构成的线性空间 $P^{n\times n}$ 同构.

定理 3 设线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是 A，向量 ξ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标是 $(x_1,x_2,\cdots,x_n)$，则 $\mathscr{A}\xi$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的坐标 $(y_1,y_2,\cdots,y_n)$ 可以按公式 $\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_n\end{pmatrix}=A\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix}$ 计算.

二、同一个线性变换在不同基下的矩阵的关系.

线性变换的矩阵是与空间中一组基联系在一起的. 一般说来，随着基的改变，同一个线性变换就有不同的矩阵. 为了利用矩阵来研究线性变换，有必要弄清楚线性变换的矩阵是如何随着基的改变而改变的.

定理 4 设线性空间 V 中线性变换 A 在两组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n, \tag{6}$$

$$\eta_1,\eta_2,\cdots,\eta_n \tag{7}$$

下的矩阵分别为A和B从基（6）到（7）的过渡矩阵是X，于是$B=X^{-1}AX$.

定理 4 告诉我们，同一个线性变换 A 在不同基下的矩阵之间的关系.

定义 3 设A，B为数域P上两个n级方阵，如果可以找到数域P上的n级可逆方阵X，使得$B=X^{-1}AX$，就说A相似于B，记作$A\sim B$.

相似是矩阵之间的一种关系，这种关系具有下面三个性质：

1. 反身性：$A\sim A$
2. 对称性：如果$A\sim B$，那么$B\sim A$.
3. 传递性：如果$A\sim B$，$B\sim C$，那么$A\sim C$.

定理 5 线性变换在不同基下所对应的矩阵是相似的；反过来，如果两个矩阵相似，那么它们可以看作同一个线性变换在两组基下所对应的矩阵.

矩阵的相似对于运算有下面的性质.

如果$B_1=X^{-1}A_1X$，$B_2=X^{-1}A_2X$，那么$B_1+B_2=X^{-1}(A_1+A_2)X$，$B_1B_2=X^{-1}(A_1A_2)X$

由此可知，如果$B=X^{-1}AX$，且$f(x)$是数域P上一多项式，那么$f(B)=X^{-1}f(A)X$

利用矩阵相似的这个性质可以简化矩阵的计算.

例 2 设V是数域P上一个二维线性空间，$\varepsilon_1,\varepsilon_2$是一组基，线性变换 A 在$\varepsilon_1,\varepsilon_2$下的矩阵是$A=\begin{pmatrix}2&1\\-1&0\end{pmatrix}$若$(\eta_1,\eta_2)=(\varepsilon_1,\varepsilon_2)\begin{pmatrix}1&-1\\-1&2\end{pmatrix}$. 求（1）$\mathcal{A}$在$V$的另一组基$\eta_1,\eta_2$下的矩阵$B$；

（2）求B^k；（3）求A^k.

五、小结

1. 回顾一下线性变换在某组基下的矩阵的概念、矩阵相似的概念；
2. 同一个线性变换在不同基下的矩阵的关系.

六、作业（习题册）

七、教学反思

7.4 特征值与特征向量

教学的时间：**年**月**日　　　　教学学时数:3 学时**

一、教学目标

1.掌握矩阵的特征值、特征向量、特征多项式的概念和性质；

2.会求线性变换的特征值和特征向量、会求矩阵的特征值和特征向量；

3.掌握相似矩阵与它们的特征多项式的关系及哈密尔顿-凯莱定理.

二、教学重点

1.矩阵的特征值、特征向量、特征多项式的概念；

2.矩阵的特征值和特征向量

三、教学难点

求矩阵的特征值和特征向量的方法

四、教学过程

一、线性变换的特征值和特征向量的定义

定义 4 设 $\mathcal{A}$ 是数域 P 上线性空间 V 的一个线性变换, 如果对于数域 P 中一数 λ_0, 存在一个非零向量 ξ, 使得

$$\mathcal{A}\xi = \lambda_0 \xi \tag{1}$$

那么 λ_0 称为 $\mathcal{A}$ 的一个特征值, 而 ξ 叫做 $\mathcal{A}$ 的属于特征值 λ_0 的一个特征向量.

例 1 求零变换和数乘变换的特征值和特征向量.

二、线性变换的特征值和特征向量的几何意义

几何意义：特征向量的方向经过线性变换后，保持在同一条直线上，这时或者方向不变 $(\lambda_0 > 0)$ 或者方向相反 $(\lambda_0 < 0)$，至于 $(\lambda_0 = 0)$ 时，特征向量就被线性变换变成 0.

三、特征值与特征向量的求法

设 V 是数域 P 上 n 维线性空间， $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 是它的一组基，线性变换 $\mathcal{A}$ 在这组基下的矩阵是 A. 设 λ_0 是特征值，它的一个特征向量 ξ 在 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标是 $x_{01}, x_{02}, \cdots, x_{0n}$, 则 $\mathcal{A}\xi$ 的坐标是

$$A\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix}.$$

$\lambda_0\xi$ 的坐标是

$$\lambda_0\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix}$$

因此（1）式相当于坐标之间的等式

$$A\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix} = \lambda_0\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix} \qquad (2)$$

或

$$(\lambda_0 E - A)\begin{pmatrix} x_{01} \\ x_{02} \\ \vdots \\ x_{0n} \end{pmatrix} = 0$$

这说明特征向量ξ的坐标$(x_{01}, x_{02}, \cdots, x_{0n})$满足齐次方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = \lambda_0 x_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = \lambda_0 x_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = \lambda_0 x_n, \end{cases}$$

即

$$\begin{cases} (\lambda_0 - a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n = 0, \\ -a_{21}x_1 + (\lambda_0 - a_{22})x_2 - \cdots - a_{2n}x_n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ -a_{n1}x_1 - a_{n2}x_2 - \cdots + (\lambda_0 - a_{nn})x_n = 0, \end{cases} \qquad (3)$$

由于$\xi \neq 0$，所以它的坐标$x_{01}, x_{02}, \cdots, x_{0n}$不全为零，即齐次线性方程组有非零解. 而齐次线性方程组有非零解的充要条件是它的系数行列式为零，即

$$|\lambda_0 E - A| = \begin{vmatrix} \lambda_0 - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda_0 - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda_0 - a_{nn} \end{vmatrix} = 0.$$

定义 5 设 A 是数域 P 上一个 n 级矩阵,λ 是一个文字. 矩阵 $\lambda E - A$ 的行列式

$$|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix}. \tag{4}$$

叫做**矩阵 A 的特征多项式**,这是数域 P 上的一个 n 次多项式.

上面的分析说明，**如果 λ_0 是线性变换 $\mathscr{A}$ 的特征值，那么 λ_0 一定是矩阵 A 的特征多项式的一个根**；反过来，**如果 λ_0 是矩阵 A 的特征多项式在数域 P 中的一个根**，即 $|\lambda_0 E - A| = 0$，那么齐次线性方程组 $(\lambda E - A)X = 0$ 就有非零解.这时，如果 $(x_{01}, x_{02}, \cdots, x_{0n})$ 是方程组（3）的一个非零解，那么非零向量

$$\xi = x_{01}\varepsilon_1 + x_{02}\varepsilon_2 + \cdots + x_{0n}\varepsilon_n$$

满足（1），即 λ_0 是线性变换 $\mathscr{A}$ 的一个特征值，ξ 就是属于特征值 λ_0 的一个特征向量.

因此确定一个线性变换 $\mathscr{A}$ 的一个特征值与特征向量的方法可以分成以下几步：

1. 在线性空间 V 中取一组基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$，写出 $\mathscr{A}$ 在这组基下的矩阵 A；

2. 求出 A 的特征多项式 $|\lambda E - A|$ 在数域 P 中全部的根，它们也就是线性变换 $\mathscr{A}$ 的全部特征值；

3. 把所求得的特征值逐个地代入方程组 $(\lambda_i E - A)X = 0$，对于每一个特征值，解方程组 $(\lambda_i E - A)X = 0$，求出一组基础解系，它们就是属于特征值 λ_i 的几个线性无关的特征向量在基 $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ 下的坐标，这样，也就求出了属于每个特征值的全部线性无关的特征向量.

矩阵 A 的特征多项式的根有时也称为 **A 的特征值**，而相应的线性方程组（3）的解也就称为 **A 的属于这个特征值的特征向量.**

例 2 （1）设τ是R^2上的一个线性变换, 对于$\forall\begin{pmatrix}x_1\\x_2\end{pmatrix}\in R^2$，有$\tau\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}2x_1+x_2\\2x_1+3x_2\end{pmatrix}$. 求$\tau$的特征值与特征向量.

（2）设线性变换$\mathscr{A}$在基$\varepsilon_1,\varepsilon_2,\varepsilon_3$下的矩阵是$A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$, 求$\mathscr{A}$的特征值与特征向量.

例 3 在空间$P[x]_n$中，线性变换

$$\mathscr{D}f(x)=f'(x)$$

在基$1,x,\frac{x^2}{2!},\cdots,\frac{x^{n-1}}{(n-1)!}$下的矩阵是

$$D=\begin{pmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&1\\0&0&0&\cdots&0\end{pmatrix}$$

D的特征多项式是

$$|\lambda E-D|=\begin{vmatrix}\lambda&-1&0&\cdots&0\\0&\lambda&-1&\cdots&0\\\cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&-1\\0&0&0&\cdots&\lambda\end{vmatrix}=\lambda^n .$$

因此，D的特征值只有 0. 通过解相应的齐次线性方程组知道，属于特征值 0 的线性无关的特征向量组只能是任一非零常数. 这表明微商为零的多项式只能是零或非零的常数.

例 4 平面上全体向量构成实数域上一个二维线性空间，§1 例 1 中旋转$\mathscr{T}_\theta$在直角坐标系下的矩阵为

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

它的特征多项式为

$$\begin{vmatrix}\lambda-\cos\theta&\sin\theta\\-\sin\theta&\lambda-\cos\theta\end{vmatrix}=\lambda^2-2\lambda\cos\theta+1$$

当$\theta\neq k\pi$时，这个多项式没有实根. 因之，当$\theta\neq k\pi$时，$\mathscr{T}_\theta$没有特征值. 从几何上看，

这个结论是明显的.

四、特征值与特征向量的性质

问题 1 特征向量是不是被特征值所唯一决定的?

问题 2 特征值是不是被特征向量所唯一决定的?

问题 3 数域 P 上的线性空间 V 上的线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的特征向量和零向量作成的集合是不是 V 的子空间?

1. 特征子空间

如果 ξ 是线性变换 $\mathscr{A}$ 的属于特征值 λ_0 的特征向量，那么 ξ 的任何一个非零倍数 $k\xi$ 也是 $\mathscr{A}$ 的属于特征值 λ_0 的特征向量. 这说明**特征向量不是被特征值所唯一决定的**. 相反，**特征值却是被特征向量所唯一决定的**，这是因为，如果特征向量 ξ 是同时属于特征值 λ_1,λ_2，那么 $\mathscr{A}\xi=\lambda_1\xi$，且 $\mathscr{A}\xi=\lambda_2\xi$，于是 $\lambda_1\xi=\lambda_2\xi$，从而（$\lambda_1-\lambda_2$）$\xi=0$，因此 $\lambda_1=\lambda_2$.

对于线性变换 $\mathscr{A}$ 的任一个特征值 λ_0，全部适合条件

$$\mathscr{A}\alpha=\lambda_0\alpha$$

的向量 α 所成的集合，也就是 $\mathscr{A}$ 的属于 λ_0 的全部特征向量再添上零向量所成的集合，是 V 的一个子空间，称为 $\mathscr{A}$ 的一个特征子空间，记为 V_{λ_0}.

用集合记号可写为 $V_{\lambda_0}=\{\alpha\mid A\alpha=\lambda_0\alpha,\alpha\in V\}$.

显然，V_{λ_0} 的维数就是属于 λ_0 的线性无关的特征向量的最大个数.

问题 4 同一个线性变换关于不同基下的矩阵多项式的关系如何?

2. 同一个线性变换关于不同基下的矩阵多项式的关系

特征值自然是被线性变换所决定的. 但是在有限维空间中，**任取**一组基后，特征值就是线性变换在这组基下矩阵的特征多项式的根. 随着基的不同，线性变换的矩阵一般是不同的. 下面研究同一个线性变换关于不同基下的矩阵多项式的关系如何?

设线性空间 V 中线性变换 $\mathscr{A}$ 在两组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，$\eta_1,\eta_2,\cdots,\eta_n$ 下的矩阵分别为 A 和 B 从基前一组基到后一组基的过渡矩阵是 X，于是 $B=X^{-1}AX$.

$$|\lambda E-B|=|\lambda E-X^{-1}AX|=|\lambda X^{-1}EX-X^{-1}AX|=|X^{-1}(\lambda E-A)X|=|X^{-1}||\lambda E-A||X|=|\lambda E-A|$$

定理 6 相似矩阵有相同的特征多项式.**(同一个线性变换关于不同基下的矩阵多项式相等)**

定理 6 说明，**线性变换的矩阵的特征多项式与基的选取无关，它直接被线性变换所决定的.** 因此，以后就可以说线性变换的特征多项式了.

既然相似的矩阵有相同的特征多项式，当然特征多项式的各项系数对于相似的矩阵来说都是相同的. 考虑特征多项式的常数项，得到相似矩阵有相同的行列式. 因此，以后就可以说线性变换的行列式.

问题 5 相似矩阵有相同的特征多项式，反之是否成立？

定理 6 的逆是不对的，特征多项式相同的矩阵不一定是相似的. 例如

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

它们的特征多项式都是$(\lambda-1)^2$，但 A 和 B 不相似，因为和 A 相似的矩阵只能是 A 本身.

问题 6 特征多项式具有哪些性质？

3. 特征多项式的性质

在线性变换的研究中，矩阵的特征多项式是重要的. 下面先来看一下它的系数. 在

$$|\lambda E-A|=\begin{vmatrix}\lambda-a_{11}&-a_{12}&\cdots&-a_{1n}\\-a_{21}&\lambda-a_{22}&\cdots&-a_{2n}\\\cdots&\cdots&\cdots&\cdots\\-a_{n1}&-a_{n2}&\cdots&\lambda-a_{nn}\end{vmatrix}.$$

的展开式中，有一项是主对角线上元素的连乘积

$$(\lambda-a_{11})(\lambda-a_{22})\cdots(\lambda-a_{nn})$$

展开式中的其余项，至多包含$n-2$个主对角线上的元素，它对λ的次数最多是$n-2$. 因此特征多项式中含λ的n次与$n-1$次的项只能在主对角线上元素的连乘积中出现，它们是

$$\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}.$$

在特征多项式中令$\lambda=0$，即得常数项$|-A|=(-1)^n|A|$.

因此，如果只写特征多项式的前两项与常数项，就有

$$|\lambda E-A|=\lambda^n-(a_{11}+a_{22}+\cdots+a_{nn})\lambda^{n-1}+\cdots+(-1)^n|A|. \tag{5}$$

若$|\lambda E-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$，则由根与系数的关系可知，$A$ 的全体特征值的和

为$\lambda_1+\lambda_2+\cdots+\lambda_n=a_{11}+a_{22}+\cdots+a_{nn}$（称为$A$的迹）. 而的$A$全体特征值的积$\lambda_1\lambda_2\cdots\lambda_n=|A|$.

设A是数域P上一个$n\times n$矩阵，$f(\lambda)=|\lambda E-A|$是A的特征多项式，问： $f(A)=?$

哈密顿-凯莱(Hamilton-Caylay)定理 设A是数域P上一个$n\times n$矩阵，$f(\lambda)=|\lambda E-A|$是A的特征多项式，则

$$f(A)=A^n-(a_{11}+a_{22}+\cdots+a_{nn})A^{n-1}+\cdots+(-1)^n|A|E=0$$

推论 设$\mathcal{A}$是有限维空间V的线性变换，$f(\lambda)$是$\mathcal{A}$的特征多项式，那么$f(\mathcal{A})=o$.

五、小结

1.请同学叙述特征值和特征向量的概念；

2.如何求线性变换的特征值和特征向量？

3.同一个线性变换关于不同基下的特征值相同与否？

4.叙述特征子空间的定义和特征多项式具有的性质？

5.叙述 Hamilton-Caylay)定理及其推论.

六、作业（习题册）

七、教学反思

7.5 对角矩阵

教学的时间：**年**月**日 教学学时数:2 学时**

一、教学目标

1. 了解线性变换可对角化、矩阵可对角化的概念.

2. 掌握线性变换在某一组基下的矩阵为对角矩阵的充要条件.

2. 熟练掌握线性变换在某一个基下的矩阵为对角矩阵以及相应的矩阵对角化的方法.

二、教学重点

线性变换在某一组基下的矩阵为对角矩阵的充要条件以及矩阵对角化的方法.

三、教学难点

线性变换在某一组基下的矩阵是对角矩阵的判断与计算.

四、教学过程

（一）引入

在 7.3 节中，我们学过：在 L(v)与 $P^{n\times n}$ 建立了一个同构映射，即 V 上的线性变换 $\mathcal{A}$ 关于某组基下的矩阵为 A, 若线性变换 $\mathcal{A}$ 关于另一组基下的矩阵为 B, 则 $B=X^{-1}AX$.

问：此时矩阵 A、B 它们有那些相同的性质？从秩、行列式、行/列向量的线性相关性、特征多项式、特征值等来考虑。

说明相似矩阵具有那么多相同的性质特征，在相似矩阵中选取一类矩阵使得它比较简单，且能具备刚才的性质特征。

问：零矩阵、单位矩阵可以吗？零矩阵、单位矩阵与哪些矩阵相似？我们发现他们只能与自己相似，所以我们选取的矩阵不可以选我们认为最简单的零矩阵、单位矩阵。接下来我们就可以考虑比这两个矩阵稍微复杂一点的矩阵——对角矩阵。我们暂时认为对角矩阵选取对角矩阵的是比较简单合适。

（二）可对角化的充要条件

问题 1 线性变换在一组基下的矩阵是对角矩阵的充要条件？

设线性变换 $\mathcal{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵为对角矩阵 $A=diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$.

于是有 $A(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)diag(\lambda_1,\lambda_2,\cdots,\lambda_n)$, 即 $A(\varepsilon_i)=\lambda_i\varepsilon_i \quad (i=1,2,\cdots,n)$,

也就是 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是线性变换 $\mathcal{A}$ 的 n 个线性无关的特征向量.

反过来，线性变换 $\mathscr{A}$ 的 n 个线性无关的特征向量 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$，那么线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 下的矩阵是不是对角矩阵？

定理 7 设 $\mathscr{A}$ 是 n 维线性空间 V 的一个线性变换，线性变换 $\mathscr{A}$ 的矩阵可以在某一基下为对角矩阵的充要条件 $\mathscr{A}$ 有 n 个线性无关的特征向量.

例 1 设线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵是 $A=\begin{pmatrix}1&2&2\\2&1&2\\2&2&1\end{pmatrix}$，由上节课的**例 2** 的第 2 小题（课本 p293）得：线性变换 $\mathscr{A}$ 的特征值 -1 对应的特征向量 $\xi_1=\begin{pmatrix}1\\0\\-1\end{pmatrix}$, $\xi_2=\begin{pmatrix}0\\1\\-1\end{pmatrix}$，特征值 5 对应的特征向量 $\xi_3=\begin{pmatrix}1\\1\\1\end{pmatrix}$.

问：(1) ξ_1,ξ_2,ξ_3 是不是线性无关？他们能否做成 R^3 的一组基？

(2) 求线性变换 $\mathscr{A}$ 在基 ξ_1,ξ_2,ξ_3 下的矩阵.

(3) 属于不同特征值的特征向量 ξ_3,ξ_1 的线性相关性？ ξ_3,ξ_2 的线性相关性？

问题 2 属于不同特征值的特征向量是不是线性无关的？

特征值的个数是自然数个，考虑利用数学归纳法来证明.

当 k=1 时，由于特征向量是非零向量，因此，线性无关.

假设命题对于 k 个不同的特征值时，相对应的特征向量线性无关，下证对于 k+1 个不同的特征值时，相对应的特征向量线性无关即可.

定理 8 属于不同特征值的特征向量是线性无关的.

例 2 设线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵是 $A=\begin{pmatrix}1&4&2\\0&-3&4\\0&4&3\end{pmatrix}$，容易得线性变换 $\mathscr{A}$ 的特征值-5, 1, 5，相应的特征向量为：$\xi_1=\begin{pmatrix}1\\-2\\1\end{pmatrix},\xi_2=\begin{pmatrix}1\\0\\0\end{pmatrix},\xi_3=\begin{pmatrix}2\\1\\2\end{pmatrix}$.（课本 p320 习题 22 改编）

(1) ξ_1,ξ_2,ξ_3 是不是线性无关？他们能否做成 R^3 的一组基？

(2)求线性变换 $\mathscr{A}$ 在基 ξ_1,ξ_2,ξ_3 下的矩阵.

问题 3 若线性空间的维数与线性变换的不同的特征值的个数相同，则该线性变换在某个基下的矩阵一定是对角形吗？

推论 1 如果在 n 维线性空间 V 中，线性变换 $\mathscr{A}$ 的特征多项式在数域 P 中有 n 个不同的根，即 $\mathscr{A}$ 有 n 个不同的特征值，那么 $\mathscr{A}$ 在某组基下的矩阵是对角形的.

例如， 例 2 知：线性变换 $\mathscr{A}$ 在基 $\varepsilon_1,\varepsilon_2,\varepsilon_3$ 下的矩阵是 $A=\begin{pmatrix}1&4&2\\0&-3&4\\0&4&3\end{pmatrix}$，线性变换 $\mathscr{A}$ 的特征值-5, 1, 5. 由**推论 1 知：**线性变换 $\mathscr{A}$ 在基 ξ_1,ξ_2,ξ_3 下的矩阵是对角矩阵.

推论 2 在复数上的线性空间中，如果线性变换 $\mathscr{A}$ 的特征多项式没有重根，那么 $\mathscr{A}$ 在某组基下的矩阵是对角形的.

在一个线性变换没有 n 个不同的特征值的情形，要判断这个线性变换的矩阵能不能成为对角形，问题就要复杂些.

问题 4 若维线性空间的维数与线性变换的不同的特征值的个数不一定相同，则该线性变换属于不同特征值的线性无关的特征向量作成的向量组是否线性无关？

定理 9 如果 $\lambda_1,\cdots,\lambda_k$ 是线性变换 $\mathscr{A}$ 的不同的特征值，而 $\alpha_{i1},\cdots,\alpha_{ir_i}$ 是属于特征值 λ_i 的线性无关的特征向量，$i=1,2,\cdots,k$，那么向量组 $\alpha_{11},\cdots,\alpha_{ir_1},\cdots,\alpha_{k1},\cdots,\alpha_{kr_k}$ 也线性无关.

根据这个定理，对于一个线性变换，求出属于每个特征值的线性无关的特征向量，把它们合在一起还是线性无关的. 如果属于每个特征值的线性无关的特征向量总的个数等于空间的维数 n，那么这个线性变换在刚才这 n 个线性无关的特征向量作为线性空间 v 的基下的矩阵是对角矩阵；如果线性无关的特征向量的个数少于空间的维数，那么这个线性变换在任何一组基下的矩阵都不能是对角形. 换句话说，**设 A 全部不同的特征值是 $\lambda_1,\cdots,\lambda_r$，于是 A 在某一组基下的矩阵成对角形的充要条件是 A 的特征子空间 $V_{\lambda_1},\cdots,V_{\lambda_r}$ 的维数之和等于线性空间的维数.**

应该看到，当线性变换 $\mathscr{A}$ 在一组基下的矩阵 A 是对角形时：$A=\begin{pmatrix}\lambda_1&0&0&\cdots&0\\0&\lambda_2&0&\cdots&0\\ \cdots&\cdots&\cdots&\cdots&\cdots\\0&0&0&\cdots&\lambda_n\end{pmatrix}$

A 的特征多项式就是$|\lambda E-A|=(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

因此，如果线性变换$\mathcal{A}$在一组基下的矩阵是对角形，那么主对角线上的元素除排列次序外是确定的，它们正好是$\mathcal{A}$的特征多项式全部的根（重根按重数计算）.

根据§7.3 定理 5，一个线性变换的矩阵能不能在某一组基下是对角形的问题就相当于一个矩阵是不是相似于一个对角矩阵的问题.

例 3 设线性变换$\mathcal{A}$在基$\varepsilon_1,\varepsilon_2,\varepsilon_3$下的矩阵是$A=\begin{pmatrix}-2&1&1\\0&2&0\\-4&1&3\end{pmatrix}$.

求 (1)线性变换$\mathcal{A}$的 3 个线性无关的特征向量ξ_1,ξ_2,ξ_3.

(2)线性变换$\mathcal{A}$在基ξ_1,ξ_2,ξ_3.下的矩阵为对角矩阵 B.

(3)由$\varepsilon_1,\varepsilon_2,\varepsilon_3$到$\xi_1,\xi_2,\xi_3$.的过渡矩阵 X, 求$X^{-1}AX$.

(4)求B^k， A^k.

五、小结

1. $\mathcal{A}$是 n 维线性空间 $\mathcal{V}$ 的一个线性变换，$\mathcal{A}$在某个基下的矩阵是对角矩阵的充要条件？

2. 一个线性变换不同特征值的特征向量是线性相关还是线性无关？

3. 一个线性变换不同特征值的线性无关的特征向量作成的向量组是线性相关性？

4. 如何求一个线性变换在某个基下的对角矩阵?

六、作业（习题册）

七、教学反思

7.6 线性变换的值域和核

教学的时间：**年**月**日 教学学时数:2 学时**

一、教学目标

1. 掌握线性变换的值域、核、秩、零度等概念及其相关的性质；

2. 掌握线性变换的值域与基像组所生成的子空间的关系；

3. 能利用线性变换的值域的一组基的原像和与其核的基找线性空间的基；

4. 理解线性变换的秩和零度间的关系与线性空间的维数的关系；

5. 了解有限维线性空间上的线性变换是单射的充要条件。

二、 教学重点

1. 线性变换的值域、核、秩、零度等概念及其相关的性质；

2. 线性变换的值域与基像组所生成的子空间的关系；

3. 利用线性变换的值域的一组基的原像与其核的基找线性空间的基.

三、教学难点

线性变换的值域与基像组所生成的子空间的关系；

利用线性变换的值域的一组基的原像与其核的基找线性空间的基.

四、教学过程

（一）定义

定义 6 设 $\mathscr{A}$ 是线性空间 V 的一个线性变换，$\mathscr{A}$ 的全体像组成的集合称为 $\mathscr{A}$ 的值域，用 $\mathscr{A}V$ 表示.所有被 $\mathscr{A}$ 变成零向量的向量组成的集合称为 $\mathscr{A}$ 的核，用 $\mathscr{A}^{-1}(0)$ 表示.

若用集合的记号则 $\mathscr{A}V=\{A(\xi)\mid\xi\in V\}$,$\mathscr{A}^{-1}(0)=\{\xi\mid A(\xi)=0,\xi\in V\}$

（二）值域与核的性质

问题 1 线性变换的值域与核都是不是 V 的子空间？ Yes

$\mathscr{A}V$ 的维数称为 $\mathscr{A}$ **的秩**，$\mathscr{A}^{-1}(0)$ 的维数称为 $\mathscr{A}$ **的零度.**

例 1 设线性空间 $V=P[x]_n$ 中，令 $\mathscr{D}(f(x))=f'(x)$.求 $\mathscr{D}$ 的值域 DV 和核 $D^{-1}(0)$.

($\mathscr{D}$ 的值域就是 $P[x]_{n-1}$，$\mathscr{D}$ 的核就是子空间 P .)

问题 2 微商变换 D 的值域能否看作是线性空间基像生成的子空间？是不是任意的线

性空间上线性变换的值域看作是基像生成的子空间？

（三）线性变换值域的结构

定理 10 设 $\mathscr{A}$ 是 n 维线性空间 V 的线性变换，$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 是 V 的一组基，在这组基下 $\mathscr{A}$ 的矩阵是 A，则

1） $\mathscr{A}$ 的值域 $\mathscr{A}V$ 是由基像组生成的子空间，即

$$\mathscr{A}V = \boldsymbol{L}(\mathscr{A}\varepsilon_1,\mathscr{A}\varepsilon_2,\cdots,\mathscr{A}\varepsilon_n)$$

2） $\mathscr{A}$ 的秩= A 的秩.

定理 10 说明线性变换与线性变换的矩阵之间的对应关系保持秩不变.

例 2 设 V 是数域 $\boldsymbol{P}$ 上的 $\boldsymbol{n}$ 维的线性空间，求单位变换和零变换的值域和核，以及这两个线性变换的秩和零度.

问题 3 微商变换 $\mathscr{D}$ 的值域的一组基的原像与其核的基能否作成线性空间 V 的一组基？是不是所有的线性空间上的线性变换都具有这样的性质？

（四）线性变换的秩、零度与空间维数的关系

定理 11 设 $\mathscr{A}$ 是 n 维线性空间 V 的线性变换，则 $\mathscr{A}V$ 的一组基的原像及 $\mathscr{A}^{-1}(0)$ 的一组基合起来就是 V 的一组基.于是

$$\mathscr{A}\text{的秩}+\mathscr{A}\text{的零度}=n$$

证明 要证 $\mathscr{A}V$ 的一组基的原像及 $\mathscr{A}^{-1}(0)$ 的一组基合起来作成线性空间 $\boldsymbol{V}$ 的一组基，只需证明这组向量组一是线性无关，二是线性空间 $\boldsymbol{V}$ 中的每一个向量都可以由这组向量组线性表示即可.

例 1 中， 虽然微商变换的值域 $\mathscr{A}V$ 与核 $\mathscr{A}^{-1}(0)$ 的维数之和为 n，但 $\mathscr{A}V+\mathscr{A}^{-1}(0)$ 并不一定是整个空间.

推论 对于**有限维**线性空间的线性变换，它是单射的**充要条件**是它是满射.

(其证明留着下去思考)

例 3 设 A 是一个 $n\times n$ 矩阵, $A^2=A$.证明: A 相似于一个对角矩阵 $diag(1,1,\cdots 1,0,0,\cdots,0)$.

证明 设矩阵 A 是 $\boldsymbol{n}$ 维线性空间 $\boldsymbol{V}$ 上的线性变换 $\mathscr{A}$ 关于基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n$ 的矩阵.

也就是$\mathscr{A}(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)=(\mathscr{A}(\varepsilon_1),\mathscr{A}(\varepsilon_2),\cdots,\mathscr{A}(\varepsilon_n))=(\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_n)A$.

(要证明矩阵A与B相似，只需要证明线性变换$\mathscr{A}$关于某个基下的矩阵是B即可.)

设$\eta_1,\eta_2,\cdots,\eta_r$是$\mathscr{A}V$的一组基，其原像分别为$\xi_1,\xi_2,\cdots,\xi_r$，即$\mathscr{A}(\xi_i)=\eta_i,(i=1,\cdots,r)$.

两端作线性变换得$\mathscr{A}^2(\xi_i)=\mathscr{A}(\eta_i),(i=1,\cdots,r)$.

因为$A^2=A$，所以$\mathscr{A}^2=\mathscr{A}$从而有$\mathscr{A}^2(\xi_i)=\mathscr{A}(\xi_i)=\eta_i=\mathscr{A}(\eta_i),(i=1,\cdots,r)$.这表明$\eta_1,\eta_2,\cdots,\eta_r$是$\mathscr{A}V$的一组基的原像.

设线性变换$\mathscr{A}$的核一组基为$\eta_{r+1},\eta_{r+2},\cdots,\eta_n$,即$\mathscr{A}(\eta_i)=0$，$i=r+1,\cdots,n.$

（下证线性变换$\mathscr{A}$关于基$\eta_1,\eta_2,\cdots,\eta_r$，$\eta_{r+1},\eta_{r+2},\cdots,\eta_n$的矩阵就是$B$）

由定理 11 知：$\eta_1,\eta_2,\cdots,\eta_r,\eta_{r+1},\eta_{r+2},\cdots,\eta_n$是$\boldsymbol{V}$的一组基.

$$A(\eta_1,\eta_2,\cdots,\eta_r,\eta_{r+1},\cdots,\eta_n)=(\eta_1,\eta_2,\cdots,\eta_r,0,\cdots,0)$$

$$=(\eta_1,\eta_2,\cdots,\eta_r,\eta_{r+1},\cdots,\eta_n)\begin{pmatrix}1&&&&&&&\\&1&&&&&&\\&&\ddots&&&&&\\&&&1&&&&\\&&&&0&&&\\&&&&&0&&\\&&&&&&\ddots&\\&&&&&&&0\end{pmatrix}$$

线性变换$\mathscr{A}$关于基$\eta_1,\eta_2,\cdots,\eta_r$，$\eta_{r+1},\eta_{r+2},\cdots,\eta_n$的矩阵就是$B$. 所以矩阵$A$与$B$相似.

五、小结和思考

1. 线性变换的值域、核、秩、零度等概念及其相关的性质；

2. 请叙述定理 10、定理 11 以及定理 11 的推论的内容并解释这 2 个定理和推理解决了哪些问题？

3. **思考**：线性变换的值域与核在该线性变换下得到的结果是不是在该线性变换的值域与核中？除此之外还有没有子空间满足这样性质.（留着问题下去思考）

六、作业（习题册）

七、教学反思

7.7 不变子空间

教学的时间：**年**月**日** **教学学时数:2 学时**

一、教学目标

1. 理解不变子空间的定义,会判定一个子空间是否是不变子空间；

2. 理解不变子空间与线性变换矩阵化简之间的关系；

3. 掌握将线性空间 V 按特征值分解成不变子空间的直和表达式.

二、 教学重点

1. 不变子空间的定义；

2. 不变子空间与线性变换矩阵化简之间的关系;

3. 线性空间 V 按特征值分解成不变子空间的直和表达式.

三、教学难点

1. 不变子空间的定义及其判断;

2. 不变子空间与线性变换矩阵化简之间的关系.

四、教学过程

对于给定的 n 维线性空间 V , $\mathcal{A}\in L(V)$, 如何才能选到 V 的一个基, 使 $\mathcal{A}$ 关于这个基的矩阵具有尽可能简单的形式. 由于一个线性变换关于不同基的矩阵是相似的. 因而问题也可以这样提出:在一切彼此相似的 n 阶矩阵中, 如何选出一个形式尽可能简单的矩阵. 这一节介绍不变子空间的概念，来说明线性变换的矩阵的化简与线性变换的内在联系.

（一）不变子空间的定义

定义 7 设 A 是数域 P 上线性空间 V 的线性变换, W 是 V 的一个子空间. 若 W 中的向量在 $\mathcal{A}$ 下的像仍在 W 中，换句话说, 对于 W 中任一向量 ξ ，有 $\mathcal{A}(\xi)\in W$ ，就称 W 是 $\mathcal{A}$ 的不变子空间，简称 $\mathcal{A}$-子空间.

（二）不变子空间的例子

例 1 整个空间 V 和零子空间 $\{0\}$，对于每个线性变换 $\mathcal{A}$，都是 $\mathcal{A}$-子空间.

例 2 $\mathcal{A}$ 的值域 $\mathcal{A}V$ 与核 $\mathcal{A}^{-1}(0)$ 都是 $\mathcal{A}$-子空间.

例 3 若线性变换 $\mathcal{A}$ 与 $\mathcal{B}$ 是可交换的，则 $\mathcal{B}$ 的核 $\mathcal{B}^{-1}(0)$ 与值域 $\mathcal{B}V$ 都是 $\mathcal{A}$-子空间.

由例 3 知：$f(\mathscr{A})$的值域与核都是 $\mathscr{A}$-子空间，这是因为线性变换 $\mathscr{A}$ 的多项式 $f(\mathscr{A})$与线性变换 $\mathscr{A}$ 可交换的.

例 4 任何一个子空间都是数乘变换的不变子空间.

（三）属于特征值 λ_0 的一个特征子空间与一维不变子空间之间的联系

设W是一维 $\mathscr{A}$-子空间，ξ是W中任何一个非零向量，它构成W的基. 按 $\mathscr{A}$-子空间的定义，$\mathscr{A}\xi\in W$, 它必是ξ的一个倍数:

$$\mathscr{A}(\xi)=\lambda_0\xi.$$

这说明ξ是 $\mathscr{A}$ 的特征向量，而W即是由ξ生成的一维 $\mathscr{A}$-子空间.

反过来，设ξ是 $\mathscr{A}$ 属于特征值 λ_0 的一个特征向量，则ξ以及它任一倍数在 $\mathscr{A}$ 下的像是原像的 λ_0 倍，仍旧是ξ的一个倍数. 这说明ξ的倍数构成一个一维 $\mathscr{A}$-子空间.

问题： 1. $\mathscr{A}$ 的属于特征值 λ_0 的一个特征子空间V_{λ_0} 是不是 $\mathscr{A}$ 的一不变子空间?

2. **$\mathscr{A}$-子空间的和与交是不是 $\mathscr{A}$-子空间?**

（四）不变子空间上的线性变换与整个线性空间上的变换的区别与联系

设 $\mathscr{A}$ 是线性空间V的线性变换, W是 $\mathscr{A}$ 的不变子空间. 由于W中向量在 $\mathscr{A}$ 下的像仍在W中，这就使得有可能不必在整个空间V中来考虑 $\mathscr{A}$，而只在不变子空间W中考虑 $\mathscr{A}$，即把 $\mathscr{A}$ 看成是W的一个线性变换，称为 $\mathscr{A}$ 在不变子空间W上引起的变换. 为了区别起见，用符号 $\mathscr{A}|W$ 来表示它；但是在很多情况下，仍然用 $\mathscr{A}$ 来表示而不致引起混淆.

必须在概念上弄清楚 $\mathscr{A}$ 与 $\mathscr{A}|W$ 的异同: $\mathscr{A}$ 是V的线性变换, V中每个向量在 $\mathscr{A}$ 下都有确定的像；$\mathscr{A}|W$ 是不变子空间W上的线性变换，对于W中任一向量ξ，有

$$(\mathscr{A}|W)\xi=\mathscr{A}\xi.$$

但是对于V中不属于W的向量η来说，$(\mathscr{A}|W)\eta$是没有意义的.

例如，任一线性变换在它的核上引起的变换就是零变换，而在特征子空间V_{λ_0}上引起的变换是数乘变换 λ_0.

（五）子空间为 $\mathcal{A}$-子空间的条件

如果线性空间 V 的子空间 W 是由向量组 $\alpha_1,\alpha_2,\cdots,\alpha_s$ 生成的，即 $W=L(\alpha_1,\alpha_2,\cdots,\alpha_s)$，则 W 是 $\mathcal{A}$-子空间的充要条件为 $\mathcal{A}\alpha_1,\mathcal{A}\alpha_2,\cdots,\mathcal{A}\alpha_s$ 全属于 W.

（六）讨论不变子空间与线性变换矩阵化简之间的关系

1）设 $\mathcal{A}$ 是 n 维线性空间 V 的线性变换，W 是 V 的 $\mathcal{A}$-子空间. 在 W 中取一组基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k$，并且把它扩充成 V 的一组基

$$\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k,\varepsilon_{k+1},\cdots,\varepsilon_n. \tag{1}$$

那么，$\mathcal{A}$ 在这组基下的矩阵就具有下列形状

$$\begin{pmatrix} a_{11} & \cdots & a_{1k} & a_{1,k+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{k1} & & a_{kk} & a_{k,k+1} & \cdots & a_{kn} \\ 0 & \cdots & 0 & a_{k+1,k+1} & \cdots & a_{k+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & a_{n,k+1} & \cdots & a_{nn} \end{pmatrix} = \begin{pmatrix} A_1 & A_3 \\ O & A_2 \end{pmatrix}. \tag{2}$$

并且左上角的 k 级矩阵 A_1 就是 $\mathcal{A}|W$ 在的基 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k$ 下的矩阵.

反之，若 $\mathcal{A}$ 在基(1)下的矩阵是（2），则由 $\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_k$ 生成的子空间 W 是线性变换 $\mathcal{A}$ 的不变子空间.

2）设 V 分解成若干个 $\mathcal{A}$-子空间的直和：

$$V=W_1\oplus W_2\oplus\cdots\oplus W_s.$$

在每一个 $\mathcal{A}$-子空间 W_i 中取基

$$\varepsilon_{i1},\varepsilon_{i2},\cdots,\varepsilon_{in_i}\ (i=1,2,\cdots,s) \tag{3}$$

并把它们合并起来成为 V 的一组基 I. 则在这组基下，$\mathcal{A}$ 的矩阵具有准对角形状

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s \end{pmatrix} \tag{4}$$

其中 $A_i\,(i=1,2,\cdots,s)$ 就是 $\mathcal{A}|W$ 在基(3)下的矩阵.

反之，如果线性变换 $\mathcal{A}$ 在基 I 下的矩阵是准对角形（4），则由（3）生成的子空间 W_i 是 $\mathcal{A}$-子空间.

由此可知，**矩阵分解为准对角形与空间分解为不变子空间的直和是相当的.**

（七）线性空间的分解

下面应用哈密尔顿-凯莱定理将空间 V 按特征值分解成不变子空间的直和.

定理 12 设线性变换 A 的特征多项式为 $f(\lambda)$，它可分解成一次因式的乘积

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_s)^{r_s}$$

则 V 可分解成不变子空间的直和

$$V=V_1\oplus V_2\oplus\cdots\oplus V_s$$

其中

$$V_i=\left\{\xi\,|\,(A-\lambda_i\varepsilon)^{r_i}\xi=0,\xi\in V\right\}.$$

称 V_i 为线性变换 $\mathcal{A}$ 的属于特征值 λ_i 的根子空间.

（定理 12 只对内容有一定的了解，不作证明的要求）

五、小结

1. 请叙述不变子空间的定义；
2. 特征向量与一维不变子空间的关系？；
3. $\mathcal{A}$ 在不变子空间上引起的变换与整个空间的线性变换的关系？

4. 子空间为 A -子空间的充要条件？

5. 不变子空间与矩阵化简之间的关系？

六、作业（习题册）274, 277

七、教学反思

§7.8 若尔当(Jordan)标准形介绍

教学的时间：**年**月**日　　　　　　　　　　教学学时数:2 学时**

一、教学目标

1. 了解若尔当块、若尔当形矩阵、线性变换和矩阵的若尔当标准形的概念；

2. 会求线性变换与矩阵的若尔当标准形.

三、 教学重点

1. 若尔当块、若尔当形矩阵、线性变换和矩阵的若尔当标准形的概念；

2. 求线性变换与矩阵的若尔当标准形.

三、教学难点

求线性变换与矩阵的若尔当标准形.

四、教学过程

由前面的讨论可知，并不是对于每一个线性变换都有一组基，使它在这组基下的矩阵成为对角形.下面先介绍一下，在适当选择的基下，一般的一个线性变换能化简成什么形状.

定义 8 形式为

$$J(\lambda,t)=\begin{pmatrix}\lambda & 0 & \cdots & 0 & 0 & 0\\ 1 & \lambda & \cdots & 0 & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots & \vdots\\ 0 & 0 & \cdots & 1 & \lambda & 0\\ 0 & 0 & \cdots & 0 & 1 & \lambda\end{pmatrix}_{t\times t}$$

的矩阵称为若尔当(Jordan)块，其中λ是复数.由若干个若尔当块组成的准对角矩阵称为若尔当形矩阵，其一般形状如

$$\begin{pmatrix}A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_s\end{pmatrix} \qquad (1)$$

其中

$$A_i=\begin{pmatrix}\lambda_i & & & & \\ 1 & \lambda_i & & & \\ & \ddots & \ddots & & \\ & & 1 & \lambda_i & \\ & & & 1 & \lambda_i\end{pmatrix}_{k_i\times k_i},$$

并且 $\lambda_1,\lambda_2,\cdots,\lambda_s$ 中有一些可以相等.

例如

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} i & 0 \\ 1 & i \end{pmatrix}$$

都是若尔当块，而

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{pmatrix}$$

是一个若尔当形矩阵.

一级若尔当块就是一级矩阵，因此若尔当形矩阵中包括对角矩阵.

在一个线性变换的若尔当标准形中，主对角线上的元素正是特征多项式的全部的根（重根按重数计算）.

定理 13 设 $\mathcal{A}$ 是复数域上 n 维线性空间 V 的一个线性变换，则在 V 中必定存在一组基，使 $\mathcal{A}$ 在这组基下的矩阵是若尔当形矩阵.而且这个若尔当形矩阵除去其中的若尔当块的排列顺序外，由线性变换 $\mathcal{A}$ 唯一决定，它称为线性变换 $\mathcal{A}$ 的矩阵的若尔当标准形.

推论 每个 n 级复矩阵 A 一定与一个若尔当形矩阵相似.且这个若尔当形矩阵除去其中的若尔当块的排列顺序外，由矩阵 A 唯一决定，它称为矩阵 A 的若尔当标准形.

例 求矩阵 $A=\begin{pmatrix} -3 & 3 & -2 \\ -7 & 6 & -3 \\ 1 & -1 & 2 \end{pmatrix}$ 的若尔当标准形.

解 $f(\lambda)=|\lambda E-A|=\lambda^3-5\lambda^2+8\lambda-4=(\lambda-2)^2(\lambda-1)$.

对于特征值 $\lambda=1$ 它是单根，因而它在矩阵 A 的若尔当标准形的主对角线上只出现一次.

对于特征值 $\lambda=2$ 它是 2 重根，先计算矩阵 $2E-A=\begin{pmatrix} 5 & -3 & 2 \\ 7 & -4 & 3 \\ -1 & 1 & 0 \end{pmatrix}$ 的秩.

由$2E-A=\begin{pmatrix}5 & -3 & 2\\ 7 & -4 & 3\\ -1 & 1 & 0\end{pmatrix}\to\begin{pmatrix}-1 & 1 & 0\\ 0 & 3 & 3\\ 0 & 2 & 2\end{pmatrix}\to\begin{pmatrix}-1 & 1 & 0\\ 0 & 1 & 1\\ 0 & 0 & 0\end{pmatrix}$知：$R(2E-A)=2$.从而主对角线元素为 2 的若尔当块的总数为 n-R(2E-A)=1.(主对角线上的元素等于 2，共有 2 个 2，这样的若尔当块).

综上述得，A 的若尔当标准形为$\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 1 & 2\end{pmatrix}$.

五、小结

1. 叙述若尔当块、若尔当形矩阵、线性变换和矩阵的若尔当标准形的概念以及定理 13 及其推论；

2. 如何求线性变换与矩阵的若尔当标准形.

六、作业

七、教学反思

§7.9 最小多项式

教学的时间：**年**月**日** **教学学时数:2 学时**

一、教学目标

1. 了解最小多项式的概念及其性质；

2. 掌握一个矩阵相似于一个对角阵与它的最小多项式的关系.

四、 教学重点

1. 最小多项式的概念及其性质；

2. 矩阵相似于一个对角阵与它的最小多项式的关系

三、教学难点

矩阵相似于一个对角阵与它的最小多项式的关系

四、教学过程

根据哈密尔顿—凯莱定理，任给数域P上一个n级矩阵A，总可以找到数域P上一个多项式$f(x)$，使$f(A)=0$.如果多项式$f(x)$使$f(A)=0$，就称$f(x)$以A为根.当然，以为A根的多项式是很多的，其中次数最低的首项系数为 1 的以A为根的多项式称为A的最小多项式.这一节讨论应用最小多项式来判断一个矩阵能否对角化的问题.

引理 1 矩阵A的最小多项式是唯一的.

引理 2 设$g(x)$是矩阵A的最小多项式，那么$f(x)$以A为根的充要条件是$g(x)$整除$f(x)$.

由此可知，矩阵A的最小多项式是A的特征多项式的一个因式.

例 1 数量矩阵kE的最小多项式为$x-k$，特别地，单位矩阵的最小多项式为$x-1$，零矩阵的最小多项式为x.另一方面，若A的最小多项式是 1 次多项式，则A一定是数量矩阵.

例 2 设

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

求A的最小多项式.

例 3 设 $A=\begin{pmatrix}1&1&&\\&1&&\\&&1&\\&&&2\end{pmatrix}, B=\begin{pmatrix}1&1&&\\&1&&\\&&2&\\&&&2\end{pmatrix}$. A 与 B 的最小多项式都等于 $(x-1)^2(x-2)$，但是它们的特征多项式不同，因此 A 和 B 不是相似的.

引理 3 设 A 是一个准对角矩阵

$$A=\begin{pmatrix}A_1&\\&A_2\end{pmatrix},$$

并设 A_1 的最小多项式为 $g_1(x)$， A_2 的最小多项式为 $g_2(x)$,那么 A 的最小多项式为 $g_1(x), g_2(x)$ 的最小公倍式 $[g_1(x), g_2(x)]$.

这个结论可以推广到 A 为若干个矩阵组成的准对角矩阵的情形.即：如果

$$A=\begin{pmatrix}A_1&&&\\&A_2&&\\&&\ddots&\\&&&A_s\end{pmatrix},$$

A_i 的最小多项式为 $g_i(x), i=1,2,\cdots,s$，那么 A 的最小多项式为 $[g_1(x), g_2(x),\cdots,g_s(x)]$

引理 4 k 级若尔当块

$$J=\begin{pmatrix}a&&&\\1&a&&\\&\ddots&\ddots&\\&&1&a\end{pmatrix}$$

的最小多项式为 $(x-a)^k$.

定理 15 数域 P 上 n 级矩阵 A 与对角矩阵相似的充要条件为 A 的最小多项式是 P 上互素的一次因式的乘积.

推论 复数矩阵 A 与对角矩阵相似的充要条件是 A 的最小多项式没有重根.

五、小结

1.叙述最小多项式的概念及其性质以及定理 14

六、作业

七、教学反思