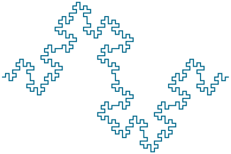
线性代数 B：特征值和特征向量

Tiao Lu
Peking University



October 13, 2019

第5章 本征值与本征向量

eigenvalue　　eigenvector

特征值　　特征向量

线性映射 $T \in \mathcal{L}(V, W)$ in chapter 3

chapter 5

→ codomain 和 domain 是同一个向量空间 V

→ $0 < \dim V < \infty$，即 V 是一个有限维的非零向量空间

Recall $T \in \mathcal{L}(V, V)$ 又称为线性算子，记作 $T \in \mathcal{L}(V)$

举例：恒等映射 $I \in \mathcal{L}(V)$. $Iv = v \quad v \in V$.

这个线性映射特别地简单，和 $1 \in \mathbb{F}$ 定义的数乘一样.

数乘和映射

设 V 是数域 F 上的一个非零向量空间，数乘是一个定义在 $F\times V$ 上的二元运算

$$F\times V\longrightarrow V$$

$$(k,v)\longmapsto kv$$

数乘和映射

设 V 是数域 F 上的一个非零向量空间，数乘是一个定义在 $F\times V$ 上的二元运算

$$F\times V\longrightarrow V$$

$$(k,v)\longmapsto kv$$

在取定数 $k\in F$ 后，上面的数乘就可以定义一个线性算子 $T_k\in\mathcal{L}(V)$，定义为

$$T_k v = v \qquad v\in V.$$

这是一类非常简单的线性算子，e.g. 零映射 $=T_0$

数乘和映射

设 V 是数域 F 上的一个非零向量空间，数乘是一个定义在 $F\times V$ 上的二元运算

$$F\times V\longrightarrow V$$

$$(k,v)\longmapsto kv$$

在取定数 $k\in F$ 后，上面的数乘就可以定义一个线性算子 $T_k\in\mathcal{L}(V)$，定义为

$$T_k v = kv \qquad v\in V.$$

这是一类非常简单的线性算子，e.g. 零映射 $=T_0$

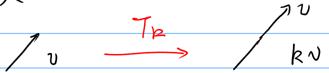
这一章的目的是把一个复杂的线性映射写成这样的简单算子之和。

特征值和特征向量

上面提到的简单的线性算子 $T_k \in \mathcal{L}(V)$

$$T_k v = kv \quad v \in V$$

的特点是 算子的作用没有改变 v 的方向，而只是改变了 v 的大小。



任给一个算子 $T \in \mathcal{L}(V)$，要求 T 作用在所有的 $v \in V$ 上都是将 v 作一个伸缩是太严格了，但是不是存在某些向量 v 使得 $Tv = kv$ 成立呢？
显然 $v=0$，$Tv=kv$ 总成立，但我们只对 $v \neq 0$ 时感兴趣。

如果存在 $k \in F$，和 $v \neq 0$ 使得 $Tv = kv$，那么称 k 是 T 的特征值(eigenvalue)，v 称为相应的特征向量

特征向量

设 $T \in \mathcal{L}(V)$. V 是 F 上的一个有限维的 nonzero vector space.

如果 $\lambda \in F$ 是 T 的特征值. $v \neq 0$ 是相应于 λ 的特征向量. 即 $Tv = \lambda v$. ①

那么 av, $a \neq 0$ 都是 T 的属于特征值 λ 的特征向量.

验证 $T(av) = a(Tv) = a(\lambda v) = (a\lambda)v = (\lambda a)v = \lambda(av)$

用 T 的齐性　用①式　用结合律　交换律　结合律.

T 的相应于 λ 的特征向量如果存在就不是惟一的, 记

$$U = \{u \in V : Tu = \lambda u\}$$

它是 T 的所有的属于 λ 的特征向量组成的集合. (多了一个元素 0)

特征子空间

命题：设 V is a finite dimensional nonzero vector space.

$T \in \mathcal{L}(V)$. λ 是 T 的一个特征值. 那么,

$$U = \{ u : Tu = \lambda u \}$$

是 V 的一个 subspace.

证明：→ It is easy to see that $0 \in U$.

→ 如果 $u_1, u_2 \in U$, 那么 $Tu_1 = \lambda u_1$ ① $Tu_2 = \lambda u_2$ ②

①+②, 可得. $Tu_1 + Tu_2 = \lambda u_1 + \lambda u_2$ 用 T 的加性 和 (V, F)

的分配律, 得 $T(u_1+u_2) = \lambda(u_1+u_2)$, 于是 $u_1+u_2 \in U$.

→ 齐性. 留给大家来验证.

这样就证明了 U is a subspace of V.

特征子空间

命题：设 V is a finite dimensional nonzero vector space.

$T \in \mathcal{L}(V)$，λ 是 T 的一个特征值，那么，

$U = \{u : Tu = \lambda u\}$ 称为 T 的一个相应于 λ 的特征子空间

是 V 的一个 subspace.

证明：→ It is easy to see that $0 \in U$.

→ 如果 $u_1, u_2 \in U$，那么 $Tu_1 = \lambda u_1$ ① $Tu_2 = \lambda u_2$ ②

①+②，可得 $Tu_1 + Tu_2 = \lambda u_1 + \lambda u_2$ 用 T 的加性和 $(V, \mathbb{F})$

的分配律，得 $T(u_1+u_2) = \lambda(u_1+u_2)$，于是 $u_1+u_2 \in U$.

→ 齐性，留给大家来验证.

这样就证明了 U is a subspace of V.

5.1 不变子空间 (invariant subspace)

$T \in \mathcal{L}(V)$ 的特征子空间

$$U = \{u : Tu = \lambda u\}$$

的一个特点是 T 把 U 中的元素还映射到 U 中

也就是说

$$u \in U \Rightarrow Tu \in U$$

5.1 不变子空间 (invariant subspace)

$T \in \mathcal{L}(V)$ 的特征子空间

$$U = \{u : Tu = \lambda u\}$$

的一个特点是 T 把 U 中的元素还映射到 U 中

也就是说

$$\boxed{u \in U \Rightarrow Tu \in U}$$

为此我们定义不变子空间.

不变子空间：设 $T \in \mathcal{L}(V)$, U is a subspace of V,

我们称 U 是 T 的一个不变子空间.

如果. $Tu \in U$ for all $u \in U$.

不变子空间举例

例：$\{0\}$ 和 V 是 $T\in\mathcal{L}(V)$ 的两个不变子空间

trivial invariant subspace.

例 $T\in\mathcal{L}(V)$ 的特征子空间

$$U=\{u\in V:\ Tu=\lambda u\}$$

是 T 的一个不变子空间

例 设 λ 是 $T\in\mathcal{L}(V)$ 的一个特征值，u 是相应于 λ 的一个特征向量（这意味着 $u\neq 0,\ Tu=\lambda u$），那么

$\operatorname{span}(u)$ 是 T 的一个一维不变子空间.

为啥对不变子空间那么感兴趣.

$T \in \mathcal{L}(V)$, $\dim V = n$. n 可能是一个很大很大的数. 描述 T 的映射的矩阵有 10^{18} 个元素

例 $n = 10^9$

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

$a_{ij} \in \mathbb{R}$, 一个实数在计算机用一个 double float 存储, 用 2 个 bytes

2×10^{18} byte $= 2 \times 10^9$ G

generally, a PC has 16 G memory. to handle such a matrix is a challenge

如果 $U \subset V$ 是一个 m 维的子空间, e.g. $n = 100$, 那么我们就可以研究 T 限制在 U 上得到的算子 $\tilde{T} \in \mathcal{L}(U)$, $\tilde{T}u = Tu$, $u \in U$

问题：

如果U不是V的不变子空间，不是也可以研究T限制在U上的线性映射吗？

定的，

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

$$\tilde{T} u = Tu \quad u \in U$$

注意 Tu 可能不属于 U，此时 $\tilde{T}$ 的定义域和值域不再相同了 ①

而且 假设 $\dim U = 100$，$\dim V = 10^9$

$M(\tilde{T})$ 的元素有 100×10^9

是上页定义的 $M(\tilde{T})$ 的元素的 10^7 倍。②

我们不喜欢的两点…

不变子空间

设 $T \in \mathcal{L}(V)$. U is a invariant subspace of T,

定义 算子 $\tilde{T} \in \mathcal{L}(U)$ 如下

$$\tilde{T} u = T u \quad u \in U$$

称上面的算子 $\tilde{T}$ 为限制在 U 上的算子

记作 $T|_U$

举例: $T: \mathcal{P}_7(\mathbb{R}) \to \mathcal{P}_7(\mathbb{R})$ 定义为

$$(Tp)(x) = p'(x)$$

易知 $\mathcal{P}_4(\mathbb{R})$ 是 T 的一个不变子空间, $T|_{\mathcal{P}_4(\mathbb{R})} \in \mathcal{L}(\mathcal{P}_4(\mathbb{R}))$

Null T 和 Range T 都不变子空间

命题：设 $T \in \mathcal{L}(V)$，那么 Null T 和 Range T 都是 T 的不变子空间.

证明： Recall that we have proved both Null T and range T are subspaces of $\underline{V}$.

domain 和 codomain 都是 V

$$\text{Null}\, T = \{v : Tv = 0\}$$

Null T 和 Range T 都不变子空间

命题：设 $T \in \mathcal{L}(V)$，那么 Null T 和 Range T 都是 T 的不变子空间.

证明：Recall that we have proved both Null T and range T are subspaces of $\underline{V}$.

domain 和 codomain 都是 V

$$\text{Null}\,T = \{v : Tv = 0\}$$

显然 $0 \in \text{Null}\,T$。$\forall v \in \text{Null}\,T$，$Tv = 0 \in \text{Null}\,T$，这就验证了 Null T 是 T 的一个不变子空间.（Null T 可能是 T 的属于 0 的特征子空间，如果 $\dim \text{Null}\,T > 0$）

$\text{Range}\,T = \{v : \exists u \in V, \text{ s.t. } Tu = v\}$，$\forall v \in \text{Range}\,T$，$Tv \in \text{Range}\,T$

于是 Range T 是 T 的一个不变子空间.

T的一维不变子空间

前面我们举例说，如果 $Tu=\lambda u$，$u\neq 0$，

那么 $\mathrm{span}(u)$ 是 V 的一个一维不变子空间

逆向思维，如果 U 是 $T\in\mathcal{L}(V)$ 的一个一维子空间，

那么 $u\in U$，$u\neq 0$ 一定是 T 的一个特征向量吗？

T的一维不变子空间

前面我们举例说 如果 $Tu=\lambda u$, $u\neq 0$,

那么 $\mathrm{span}(u)$ 是 V 的一个一维不变子空间

逆向思维。如果 U 是 $T\in\mathcal{L}(V)$ 的一个一维子空间,

那么 $u\in U$, $u\neq 0$ 一定是 T 的一个特征向量吗?

Answer: Yes. U 是一维子空间, 那么存在 $e\in U$

使得 $U=\mathrm{span}(e)$

因为 U 是 T 的不变子空间, 所以 $Te\in\mathrm{span}(e)$。由子空间的定义

知 $\exists\, c\in F$, 使得 $Te=ce$. 即 e 是 T 一个属于特征值 c 的

特征向量. 易证任意 $u\in U$, $u\neq 0$ 都是 T 的属于特征值 c 的 eigenvector

不变子空间举例

$T \in \mathcal{L}(\mathbb{R}^4)$ 定义如下

$$Te_1 = e_2$$

$$Te_2 = -e_1$$

$$Te_3 = 0$$

$$Te_4 = 0$$

易知 $\operatorname{Null} T = \operatorname{span}(e_3, e_4)$ 是 T 的一个二维不变子空间，而且 e_3 和 e_4 是 T 的对应于特征值 0 的特征向量

不变子空间举例

$T \in \mathcal{L}(\mathbb{R}^4)$ 定义如下

$$Te_1 = e_2$$
$$Te_2 = -e_1$$
$$Te_3 = 0$$
$$Te_4 = 0$$

易知 $\operatorname{Null} T = \operatorname{span}(e_3, e_4)$ 是 T 的一个二维不变子空间，而且 e_3 和 e_4 是 T 的对应于特征值 0 的特征向量

$\operatorname{Range} T = \operatorname{span}(e_1, e_2)$ 是 T 的一个二维不变子空间

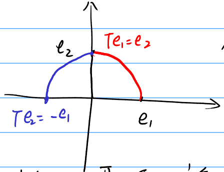
是否有特征值和特征向量？嗯。把 $\mathbb{R}^4$ 看成实数域 $\mathbb{R}$ 上的向量空间，$T \in \mathcal{L}(\mathbb{R}^4)$ 是没有特征值和特征向量的。

下一页仔细说。

$T\in\mathcal{L}(\mathbb{R}^2)$ 旋转变换

定义 $T\in\mathcal{L}(\mathbb{R}^2)$ 如下

$$Te_1 = e_2$$
$$Te_2 = -e_1$$



$$\mathcal{M}(T;(e_1,e_2),(e_1,e_2)) = \begin{array}{c|cc} & e_1 & e_2 \\ \hline e_1 & 0 & -1 \\ e_2 & 1 & 0 \end{array} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

这个线性变换表示逆时针旋转 $\frac{\pi}{2}$ 的旋转变换。

T的特征值和特征向量的定义说 存在 $\lambda\in\mathbb{R}$ 和 $u\in\mathbb{R}^2$, $u\neq 0$

使得 $Tu = \lambda u$

$\mathbb{R}^2$ 是数域 $\mathbb{R}$ 上的向量空间

这意味着 Tu 和 u 方向相同或者相反

而 T是逆时针旋转 $\frac{\pi}{2}$ 的旋转变换.

因此 Tu 和 u 不会共线。由此知 T 无特征值和特征向量.

$$T \in \mathcal{L}(\mathbb{C}^2)$$

define $T \in \mathcal{L}(\mathbb{C}^2)$ by

注意：$\mathbb{C}^2$ 是复数域 $\mathbb{C}$ 上的2维向量空间

$$Te_1 = e_2$$
$$Te_2 = -e_1$$

$$\mathcal{M}(T;(e_1,e_2),(e_1,e_2)) = \begin{array}{c|c} & \begin{array}{cc} e_1 & e_2 \end{array} \\ \begin{array}{c} e_1 \\ e_2 \end{array} & \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \end{array}$$

如果存在一个向量 $z = (z_1, z_2) \in \mathbb{C}^2$ 使得

$$Tz = \lambda z, \quad ①$$

而 $Tz = T(z_1e_1 + z_2e_2) = z_1Te_1 + z_2Te_2 = z_1e_2 - z_2e_1$ ②

$\lambda z = \lambda z_1 e_1 + \lambda z_2 e_2$ ③

把②③代入①，得 $z_1e_2 - z_2e_1 = \lambda z_1 e_1 + \lambda z_2 e_2$

由于 e_1 和 e_2 线性无关，知 $z_1 = \lambda z_2$ ④，$-z_2 = \lambda z_1$ ⑤

把⑤代入④，得 $z_1 = -\lambda^2 z_1$

设 $z_1 \neq 0$，得 $1 = -\lambda^2$，因此 $\lambda = \pm i$

这意味着，如果 λ 是 T 的特征值，一定有

$$\lambda = i \text{ or } -i$$

由④式 $z_1 = \lambda z_2$ 得

就看 $\lambda = i$，取 $z_2 = 1$，那么 $z_1 = \lambda z_2 = i \cdot 1 = i$

T 的对应 i 的特征向量是 $(z_1, z_2) = (i, 1)$

验证 $T(i,1) = T(ie_1 + e_2) = iTe_1 + Te_2 = ie_2 - e_1$

$$= (-1, i) = i(i,1)$$

☆ You write out a eigenvector of T associated with $\lambda = -i$.

上面的两个例子

$T \in \mathcal{L}(\mathbb{R}^2)$ 和 $T \in \mathcal{L}(\mathbb{C}^2)$

都表示旋转变换（在 $\mathbb{R}^2$ 上更直观），在标准基下的

矩阵都是 $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

但是 $T \in \mathcal{L}(\mathbb{R}^2)$ 无特征值和特征向量

$T \in \mathcal{L}(\mathbb{C}^2)$ 的特征值是 $i, -i$.

实数域和复数域 make a big difference.

$T\in\mathcal{L}(V)$ 的二维不变子空间

设 $U=\operatorname{span}(u_1,u_2)$ 是 $T\in\mathcal{L}(V)$ 的一个不变子空间。

令 $\tilde{T}=T|_U\in\mathcal{L}(U,U)$. 存 $a_{ij}\in F$ $i=1,2, j=1,2$.

使得. $\tilde{T}u_1=a_{11}u_1+a_{21}u_2$ $\quad \mathcal{M}(\tilde{T})=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$

$\tilde{T}u_2=a_{12}u_1+a_{22}u_2$

$\tilde{T}$ 有特征值和特征向量吗? 即是否能找到 $\lambda\in F$, $(x,y)\in F^2$

使得 $\tilde{T}(xu_1+yu_2)=\lambda(xu_1+yu_2)$ 成立呢?

$T\in\mathcal{L}(V)$ 的二维不变子空间

设 $U=\operatorname{span}(u_1,u_2)$ 是 $T\in\mathcal{L}(V)$ 的一个不变子空间。

令 $\tilde{T}=T|_U\in\mathcal{L}(U,U)$. 存 $a_{ij}\in F$ $i=1,2, j=1,2.$

使得. $\boxed{\begin{aligned}\tilde{T}u_1&=a_{11}u_1+a_{21}u_2\\ \tilde{T}u_2&=a_{12}u_1+a_{22}u_2\end{aligned}}$ $\mathcal{M}(\tilde{T})=\begin{bmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{bmatrix}$

$\tilde{T}$ 有特征值和特征向量吗？即是否能找到 $\lambda\in F,(x,y)\in F^2$

使得 $\tilde{T}(xu_1+yu_2)=\lambda(xu_1+yu_2)$ 成立呢？

$$x(\tilde{T}u_1)+y(\tilde{T}u_2)=(\lambda x)u_1+(\lambda y)u_2$$

$$x(a_{11}u_1+a_{21}u_2)+y(a_{12}u_1+a_{22}u_2)=(\lambda x)u_1+(\lambda y)u_2$$

$$x a_{11} + y a_{12} = \lambda x \qquad ①$$

$$x a_{21} + y a_{22} = \lambda y \qquad ②$$

$$\begin{cases} (\lambda - a_{11}) x - a_{12} y = 0 & ③ \\ -a_{21} x + (\lambda - a_{22}) y = 0 & ④ \end{cases}$$

$x=0, y=0$ 是 ③④ 方程组的解，但我们对非零解更感兴趣。 如果 $y \neq 0$.

$$③ \Rightarrow \frac{x}{y} = \frac{a_{12}}{\lambda - a_{11}}$$

$$④ \Rightarrow \frac{x}{y} = \frac{\lambda - a_{22}}{a_{21}}$$

$$\frac{a_{12}}{\lambda - a_{11}} = \frac{\lambda - a_{22}}{a_{21}} \Rightarrow (\lambda - a_{11})(\lambda - a_{22}) - a_{12} a_{21} = 0$$

$$\lambda^2 - (a_{11} + a_{22})x + a_{11}a_{22} - a_{12}a_{21} = 0$$

当数域 $F = \mathbb{C}$ 时，上式一定有解

$$\lambda = \frac{a_{11} + a_{22} \pm \sqrt{(a_{11}+a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}}{2}$$

有了 λ，再利用 ③ ④ 可解出非零解 (x, y)

从而知 $\tilde{T} = T|_U$ 有特征值和相应的特征向量.

$$\tilde{T}u = \lambda u. \quad u \neq 0$$

而由 $\tilde{T}$ 的定义知，$Tu = \lambda u$，$u \neq 0$ 已就说明

如果 $T = \mathcal{L}(V)$，而 V 是 $\mathbb{C}$ 上的向量空间 有一个二维的不变子空间，那么 T 至少有一个特征值。

特征值和特征向量

丘维声，简明线性代数，$P_{210}-P_{212}$

定义1（P_{210}，丘）设 V 是数域 F 上的一个线性空间，$T\in\mathcal{L}(V)$，如果 V 中存在一个非零向量 u 和一个数 $\lambda\in F$，使得

$$Tu=\lambda u.$$

则称 λ 是 T 的一个特征值，称 u 是 T 的属于 λ 的一个特征向量。

特征子空间（P_{212}，丘）：设 $T\in\mathcal{L}(V)$，$\lambda\in F$ 是 T 的一个特征值

$$V_\lambda \overset{\text{def}}{=} \{u\in V : Tu=\lambda u\}$$

称作 T 的属于特征值 λ 的特征子空间。丘书中指出 V_λ 中的全部非零向量，就是 A 的属于 λ 的全部特征向量。

eigenvalue and eigenvector

Sheldonc Axler，线性代数应该这样学Liner Algebra Done Right(中文)，version 2，杜现昆 / 马晶 翻译

Axler书中的特征值(eigenvalue)的定义和互书中一致.

特征值：设 $T \in \mathcal{L}(V)$，其中 V 是数域 F 上的 vector space，
如果存在非零向量 $u \in V$，和数 $\lambda \in F$，使得

$$Tu = \lambda u$$

那么称 λ 是 T 的一个 eigenvalue，u 是 T 的相应于 λ 的一个特征向量，

但 Axler 书中对特征向量又定义一次，设 λ 是 $T \in \mathcal{L}(V)$ 的特征值，如果 $u \in V$，使得 $Tu = \lambda u$，那么称 u 是 T 的属于 λ 的特征向量.

Axler 线性代数应该选择学的 eigenvector 定义 优缺点

优点：$V_\lambda = \{u \in V : Tu = \lambda u\}$ 是 T 的属于特征值 λ 的特征子空间

V_λ 中的所有向量都是 T 的属于 λ 的特征向量.
这很自然. 不用再费心地考虑 0 向量了.

缺点：我们对非零特征向量更感兴趣,
用五书中 eigenvector 的定义（也是大多数人采用的定义）
一提到特征向量就暗含着非零, 不用再加
"非零"这个定语了.

特征值和特征向量是一对多

一个特征值对应着无穷多个特征向量，那么会不会有一个向量同时是两个不同特征值的特征向量呢？

非零

特征值和特征向量是一对多

一个特征值对应着无穷多个特征向量，那么会不会有一个向量同时是两个不同特征值的特征向量呢？

非零

Answer: No. 反证法：设 $u \neq 0$, $u \in V$ 是 $T \in \mathcal{L}(V)$ 的特征值 λ_1, λ_2 ($\lambda_1 \neq \lambda_2$) 的特征向量，

即
$$Tu = \lambda_1 u \quad ①$$
$$Tu = \lambda_2 u \quad ②$$

① − ②，得
$$0 = (\lambda_1 - \lambda_2) u \quad ③$$

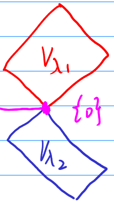
③式意味着，要么 $\lambda_1 - \lambda_2 = 0$，要么 $u = 0$. 这与 $u \neq 0$ 和 $\lambda_1 \neq \lambda_2$ 矛盾

$T\in\mathcal{L}(V)$ 的两个特征子空间 V_{λ_1}, V_{λ_2}

设 $\lambda_1\neq\lambda_2$, λ_1,λ_2 是 $T\in\mathcal{L}(V)$ 的特征值,

V_{λ_1} 和 V_{λ_2} 分别是 T 的属于 λ_1 和 λ_2 的特征子空间.

那么 $V_{\lambda_1}\cap V_{\lambda_2}=\{0\}$.



这从上一页的讨论很容易得到.

这也说明. 如果 $Tu_1=\lambda_1u_1$

$Tu_2=\lambda_2u_2$

$u_1\neq0$, $u_2\neq0$, $\lambda_1\neq\lambda_2$,

那么 u_1 和 u_2 是线性无关的.

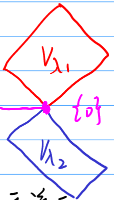
$T \in \mathcal{L}(V)$ 的两个特征子空间 V_{λ_1}, V_{λ_2}

设 $\lambda_1 \neq \lambda_2$, λ_1, λ_2 是 $T \in \mathcal{L}(V)$ 的特征值,

V_{λ_1} 和 V_{λ_2} 分别是 T 的属于 λ_1 和 λ_2 的特征子空间.

那么 $V_{\lambda_1} \cap V_{\lambda_2} = \{0\}$

这从上一页的讨论很容易得到.



这也说明. 如果 $Tu_1 = \lambda_1 u_1$

$Tu_2 = \lambda_2 u_2$

$u_1 \neq 0$, $u_2 \neq 0$, $\lambda_1 \neq \lambda_2$,

那么 u_1 和 u_2 是线性无关的.

属于两个不同特征值的特征向量是线性无关的

属于互不相同的特征值的特征向量构成一个线性无关组.

定理 5.6: 设 $T \in \mathcal{L}(V)$, $\lambda_1, \cdots, \lambda_m$ 是 T 的互不相同的特征值, $v_1, \cdots, v_m$ 是相应的~~非零~~特征向量, 则 $(v_1, \cdots, v_m)$ 线性无关.

用书的定义.

证明: 反证法. 设 $(v_1, \cdots, v_m)$ 线性相关, 那么必然存在某个 v_j 能由其余的向量线性表出

不妨设是 v_m. 那么存在 $a_i \in F$, $i = 1, \cdots, m-1$ 使得

$$v_m = a_1 v_1 + \cdots + a_{m-1} v_{m-1}$$

两边同时用 T 作用, $T v_m = T(a_1 v_1 + \cdots + a_{m-1} v_{m-1})$

由 $T v_i = \lambda_i v_i$, 得 $\lambda_m v_m = a_1 \lambda_1 v_1 + \cdots + a_{m-1} \lambda_{m-1} v_{m-1}$

属于互不相同的特征值的特征向量构成一个线性无关组.

定理 5.6：设 $T \in \mathcal{L}(V)$，$\lambda_1, \cdots, \lambda_m$ 是 T 的互不相同的特征值，$v_1, \cdots, v_m$ 是相应的~~非零~~特征向量，则 $(v_1, \cdots, v_m)$ 线性无关.

用书的定义.

证明：反证法，设 $(v_1, \cdots, v_m)$ 线性相关，那么必然存在某个 v_j 能由其余的向量线性表出

不妨设是 v_m，那么存在 $a_i \in F$，$i=1, \cdots, m-1$ 使得

$$v_m = a_1 v_1 + \cdots + a_{m-1} v_{m-1}$$

两边同时用 T 作用，$T v_m = T(a_1 v_1 + \cdots + a_{m-1} v_{m-1})$

用 $T v_i = \lambda_i v_i$，得 $\lambda_m v_m = a_1 \lambda_1 v_1 + \cdots + a_{m-1} \lambda_{m-1} v_{m-1}$ ①

两边同时乘 λ_m，得 $\lambda_m v_m = a_1 \lambda_m v_1 + \cdots + a_{m-1} \lambda_m v_{m-1}$ ②

①② 只能得出 $\lambda_m v_m$ 用 $(v_1, \cdots, v_{m-1})$ 线性表出是不唯一的，故 $(v_1, \cdots, v_{m-1})$ 是线性相关的，

进而可再用类似的方法得出

$(v_1, \cdots, v_{m-2})$ 不一定真的是前 m-2 个向量，只是向量的个数减少了一个。

是线性相关的 …… 一直到 (v_1, v_2) 是线性相关的，

矛盾 证毕。

①② 只能得出 $\lambda_m v_m$ 用 $(v_1, \cdots, v_{m-1})$ 线性表出是不唯一的，故 $(v_1, \cdots, v_{m-1})$ 是线性相关的，

进而可再用类似的方法得出

$(v_1, \cdots, v_{m-2})$ 不一定真的是前 m-2 个向量 只是向量的个数减少了一个.

是线性相关的 …… 一直到 (v_1, v_2) 是线性相关的，

矛盾 证毕.

一个简洁的证明见课本，用到了以前学习的一个定理：

$(v_1, \cdots, v_m)$ 线性相关，$v_1 \neq 0$，那么一定存在 $2 \leq k \leq m$，使得 $(v_1, \cdots, v_{k-1})$ 是线性无关的，且 $v_k \in \operatorname{span}(v_1, \cdots, v_{k-1})$.

$T \in \mathcal{L}(V)$ 的特征子空间的个数 $\leq \dim V$

设 $T \in \mathcal{L}(V)$ 的不同的特征值有 $\lambda_1, \dots, \lambda_m$,

那么存在 $v_1, \dots, v_m$ 是相应的~~非零~~特征向量,

由上面的定理知 $(v_1, \dots, v_m)$ 线性无关.

$(v_1, \dots, v_m)$ 是 V 中的线性无关组,故 $m \leq \dim V$.

5.9 推论: V 上的每个算子最多有 $\dim V$ 个互不相同的特征值(最多有 $\dim V$ 个不同的特征子空间)

$$T \in \mathcal{L}(\mathbb{C}^3)$$

$\mathbb{C}^3$ 的标准基 e_1, e_2, e_3

$$\begin{aligned} Te_1 &= a_{11}e_1 + a_{21}e_2 + a_{31}e_3 \\ Te_2 &= a_{12}e_1 + a_{22}e_2 + a_{32}e_3 \\ Te_3 &= a_{13}e_1 + a_{23}e_2 + a_{33}e_3 \end{aligned} \qquad M(T) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{22} & a_{23} & a_{33} \end{bmatrix}$$

$$T(x, y, z) = \lambda(x, y, z)$$

$$T(x, y, z) = T(xe_1 + ye_2 + ze_3) = xTe_1 + yTe_2 + zTe_3$$

$$= x(a_{11}e_1 + a_{21}e_2 + a_{31}e_3) + y(a_{12}e_1 + a_{22}e_2 + a_{32}e_3)$$

$$+ z(a_{13}e_1 + a_{23}e_2 + a_{33}e_3)$$

$$= (xa_{11} + ya_{12} + za_{13})e_1 + \cdots$$

$$T \in \mathcal{L}(\mathbb{C}^3)$$

$\mathbb{C}^3$ 的标准基 e_1, e_2, e_3

$$\begin{aligned} Te_1 &= a_{11} e_1 + a_{21} e_2 + a_{31} e_3 \\ Te_2 &= a_{12} e_1 + a_{22} e_2 + a_{32} e_3 \\ Te_3 &= a_{13} e_1 + a_{23} e_2 + a_{33} e_3 \end{aligned} \qquad M(T) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{22} & a_{23} & a_{33} \end{bmatrix}$$

$$T(x, y, z) = \underline{\lambda}(\underline{x}, y, z)$$

$$T(x, y, z) = T(xe_1 + ye_2 + ze_3) = xTe_1 + yTe_2 + zTe_3$$

$$= x(a_{11} e_1 + a_{21} e_2 + a_{31} e_3) + y(a_{12} e_1 + a_{22} e_2 + a_{32} e_3)$$

$$+ z(a_{13} e_1 + a_{23} e_2 + a_{33} e_3)$$

$$= \underline{(xa_{11} + ya_{12} + za_{13})} e_1 + \cdots$$

$$T \in \mathcal{L}(\mathbb{C}^3)$$

$\mathbb{C}^3$ 的标准基 e_1, e_2, e_3

$$\begin{aligned} Te_1 &= a_{11}e_1 + a_{21}e_2 + a_{31}e_3 \\ Te_2 &= a_{12}e_1 + a_{22}e_2 + a_{32}e_3 \\ Te_3 &= a_{13}e_1 + a_{23}e_2 + a_{33}e_3 \end{aligned} \qquad \mathcal{M}(T) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{22} & a_{23} & a_{33} \end{bmatrix}$$

$$T(x, y, z) = \lambda(x, y, z)$$

$$T(x, y, z) = T(xe_1 + ye_2 + ze_3) = xTe_1 + yTe_2 + zTe_3$$

比较e_1的系数 $a_{11}x + a_{12}y + a_{13}z = \lambda x$ ①

比较e_2的系数, $a_{21}x + a_{22}y + a_{23}z = \lambda y$ ②

比较e_3的系数 $a_{31}x + a_{32}y + a_{33}z = \lambda z$ ③

$x=0,\ y=0,\ z=0$ 显然是 ①②③ 的解, 但

我们 are more interested in a nonzero solution.

① 和 ② 联立, 把 x, y 表示成 z 和 λ 的表达式, 然后代入

③, 得到一关于z和λ的表达式.

求 $T \in \mathcal{L}(\mathbb{C}^3)$ 的特征值

Not a easy task. 我们也不知道T是否有特征值

后面会学到

V 是复数域 $\mathbb{C}$ 上的 a vector space，$T \in \mathcal{L}(V)$ 都有一个特征值

为此我们需要学习算子多项式.

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

可以定义 $T^0 = I$，$T^1 = T$，$T^2 = TT, \dots, T^n$

都是 V 上的算子.

$a_i \in F$，$i = 0, 1, \dots, n$

那么可以定义 $p(T) = a_0 I + a_1 T + \dots + a_n T^n$

两个线性算子的乘积还是线性算子，这可以得出 $T^n \in \mathcal{L}(V)$

由线性算子 $\mathcal{L}(V)$ 是一个向量空间，知 $p(T) \in \mathcal{L}(V)$.

$T \in \mathcal{L}(\mathbb{R}^2)$ defined by

$Te_1 = e_2$

$Te_2 = -e_1$

$T^0 = I$, $T^1 = T$, $\mathcal{M}(T) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

T^2 是什么？ $T^2 e_1 = TTe_1 = Te_2 = -e_1$

$T^2 e_2 = TTe_2 = T(-e_1) = -e_2$

$\mathcal{M}(T^2) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

可以看出 -1 是 T^2 的特征值，e_1 和 e_2 都 T^2 的属于 -1 的特征向量.

还可以看出 $T^2 = -I$，$I \in \mathcal{L}(\mathbb{R}^2)$ 是恒等映射.

T Rotate by $\frac{\pi}{2}$ anticlockwise

T^2 Rotate by π anticlockwise

$T^2 = -I$

$T^4 = (-I)^2 = I$ 对吗？

这相当于 $T^4 = TTTT = (TT)(TT) = (-I)(-I) = I^2 = I$

结合律

对，我们就是使用了算子乘积的结合律

多项式分解定理

定理. $p \in \mathcal{P}_m(\mathbb{C})$ 是一个 m 次 $(m>0)$ 首一复系数多项式，那么 p 可以唯一地（除因式次序外）分解成

$$p(z) = (z-\lambda_1)\cdots(z-\lambda_m) \quad ①$$

其中 $\lambda_1, \dots, \lambda_m \in \mathbb{C}$.

上面的多项式的自变量是 z，如果换成 $T \in \mathcal{L}(V)$，上面的定理还成立吗？

$$p(T) = (T-\lambda_1 I)\cdots(T-\lambda_m I)$$

$$a_0 I + a_1 T + \cdots + a_{m-1} T^{m-1} + T^m = (T-\lambda_1 I)\cdots(T-\lambda_m I) \quad ②$$

分解定理实际上是给出 $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ 和系数

$a_0, a_1, \dots, a_{m-1} \in \mathbb{C}$

之间的关系. 多项式的变量是 z 还是 T 对此没有影响.
只要 $z^0(T^0), z^1(T^1), \dots, z^m(T^m)$ 有意义即可

因此 我们有下面的算子多项式的分解定理成立

定理: 设 $p \in \mathcal{P}_m(\mathbb{C})$ 是一个 m 次 $(m>0)$ 的复系数多项式,
$T \in \mathcal{L}(V)$, 那么 $p(T)$ 可以唯一地(除因子次序外)分解成

$$p(T) = (T-\lambda_1 I)\cdots(T-\lambda_m I)$$

其中 $\lambda_1, \dots, \lambda_m \in \mathbb{C}$.

目的：证明 $T\in\mathcal{L}(V)$，V 是 $\mathbb{C}$ 上的 nonzero vector space，至少有一个特征值。

现状：直接去研究 $Tu=\lambda u$ 而求解问题太麻烦了。

$Tu=\lambda u$，$u\neq 0$ 意味着 $T-\lambda I$ 不是单射

$p(T) = a_0 I + \cdots + a_m T^m$ $(a_m\neq 0)$ 不是单射 如何证明呢？

如果能证明 $\underbrace{p(T) = c(T-\lambda_1 I)\cdots(T-\lambda_m I)}_{\text{利用分解定理知}}$，令 $p(T)$ 不是单射 知

存在某 λ_i，使得 $T-\lambda_i I$ 不是单射。

命题：设 V 是一个有限维的向量空间，$T, S \in \mathcal{L}(V)$。如果 T, S 有一个不是单射，那么 TS 和 ST 都不是单射。

证明：因为 S 不是单射，所以存在 $u \in V$，$u \neq 0$，$Su = 0$

因此 $(TS)u = T(Su) = T0 = 0$

即 $u \in \mathrm{Null}(TS)$，u 又不是 0，故 TS 不是单射

如果 S 是单射，那么 T 一定不单射，那么存在 $u \in V$，$u \neq 0$ 使得 $Tu = 0$。而 S 是单射，V 是有限维的，故 S 也是满射，那么存在 $v \in V$，使得 $Sv = u$。而且 u 和可推出 $v \neq 0$。于是 $TSv = Tu = 0$，v 又非零，这就证明了 TS 不是单射。

命题：设 V is a finite dimensional vector space, and $T, S \in \mathcal{L}(V)$. 如果 TS 不是单射，那么 T 和 S 中至少有一个不是单射.

if $\dim V < \infty$, $T \in \mathcal{L}(V)$ then 单 $\Leftrightarrow$ 满 $\Leftrightarrow$ 可逆

证明：反证法，不然 T 和 S 都是单射. 那么 T 和 S 都是可逆的，$T^{-1}, S^{-1} \in \mathcal{L}(V)$

$$(TS)(S^{-1}T^{-1}) = T(SS^{-1})T^{-1} = TIT^{-1} = TT^{-1} = I$$

$$(S^{-1}T^{-1})(TS) = S^{-1}(T^{-1}T)S = S^{-1}IS = S^{-1}S = I$$

结合律

这就验证了 TS 是可逆的，从而 TS 是单射. 这与 TS 不是单射矛盾. 故 T, S 中至少有一个不是单射

有限维复线性空间上的线性算子一定有特征值

V, Field $\mathbb{C}$, $n = \dim V < \infty$, $T \in \mathcal{L}(V)$

如果 $a_0 I + a_1 T + \cdots + a_m T^m$ $a_m \neq 0$, 不是单射.

$$a_0 I + a_1 T + \cdots + a_m T^m = c(T - \lambda_1 I) \cdots (T - \lambda_m I)$$

必有 $T - \lambda_i I$ 不单射. 则可证出 T 有特征值

有限维复线性空间上的线性算子一定有特征值

V, Field $\mathbb{C}$, $n = \dim V < \infty$, $T \in \mathcal{L}(V)$

如果 $a_0 I + a_1 T + \cdots + a_m T^m$ $a_m \neq 0$, 不是单射.

$$a_0 I + a_1 T + \cdots + a_m T^m = c(T - \lambda_1 I) \cdots (T - \lambda_m I)$$

必有 $T - \lambda_i I$ 不单射. 则可证出 T 有特征值

所有的问题就归结到证明存在 $a_0, \cdots, a_m$ 使得

$$a_0 I + \cdots + a_m T^m$$

不是单射. $(I, T, \cdots, T^m)$ 是 $\mathcal{L}(V)$ 中的一个向量组, $\mathcal{L}(V)$ 是有限维的, $\dim \mathcal{L}(V) = \dim V \cdot \dim V = n^2$.

m+1

只要 $(I, T, \dots, T^m)$ 的长度大于 n^2，则 $(I, \dots, T^m)$ 是线性相关的，故存在 $a_0, \dots, a_{n^2}$ 不全为 0，使得 $a_0 I + \dots + a_{n^2} T^{n^2} = 0$

自然不是单射.

而 $a_0 I + \dots + a_{n^2} T^{n^2}$ 的次数一定会大于 0，

(不然，$a_1 = a_2 = \dots = a_{n^2} = 0$，$a_0 I = 0 \Rightarrow a_0 = 0$，这与 $a_0, \dots, a_{n^2}$ 不全为 0 矛盾)。

课本上有另一种证明存在 $a_0, \dots, a_n$ 不全为 0，使得 $a_0 I + a_1 T + \dots + a_n T^n$ 不是单射

命题：设 $\dim V = n$，$T \in \mathcal{L}(V)$. 那么存在 $a_i \in F$, $i = 0, 1, \cdots, n$ 且 $a_1, \cdots, a_n$ 不全为 0，使得

$$a_0 I + a_1 T + \cdots + a_n T^n$$

不是单射.

证明：取 $v \in V \neq 0$. $(v, Tv, \cdots, T^n v)$ 是 V 中的一个向量组，且其长度为 $n+1$，是大于 $\dim V$.

故 $(v, Tv, \cdots, T^n v)$ 线性相关，

所以存在不全为 0 的数 $a_i \in F$, $i = 0, 1, \cdots, n$

使得 $a_0 v + a_1 T v + \cdots + a_n T^n v = 0$ 即 $\left(\sum_{i=0}^{n} a_i T^i\right) v = 0$

Recall $v \neq 0$. 这就证明了 $a_0 + \cdots + a_n T^n$ 不是单射.

命题：设 $\dim V = n$，$T \in \mathcal{L}(V)$，那么存在 $a_i \in F$，$i = 0, 1, \cdots, n$ $\boxed{\text{且 } a_1, \cdots, a_n \text{ 不全为 } 0}$，使得

$$a_0 I + a_1 T + \cdots + a_n T^n$$

不是单射。

证明：取 $v \in V \neq 0$。$(v, Tv, \cdots, T^n v)$ 是 V 中的一个向量组，且其长度为 $n+1$，是大于 $\dim V$。

故 $(v, Tv, \cdots, T^n v)$ 线性相关，

所以存在不全为 0 的数 $a_i \in F$，$i = 0, 1, \cdots, n$

使得 $\boxed{a_0 v + a_1 T v + \cdots + a_n T^n v = 0}$ 即 $\left(\sum_{i=0}^{n} a_i T^i\right) v = 0$

Recall $v \neq 0$，这就证明了 $a_0 + \cdots + a_n T^n$ 不是单射。

因为 $v \neq 0$
故 $a_1, \cdots, a_n$
不能全为 0
确定

命题：设 $\dim V = n$，$T \in \mathcal{L}(V)$，那么存在 $a_i \in F$，$i = 0, 1, \cdots, n$ $\boxed{\text{且 } a_1, \cdots, a_n \text{ 不全为 } 0}$，使得

$$a_0 I + a_1 T + \cdots + a_n T^n$$

不是单射。

证明：取 $v \in V \neq 0$。$(v, Tv, \cdots, T^n v)$ 是 V 中

的一个向量组，且其长度为 $n+1$，是大于 $\dim V$。

故 $(v, Tv, \cdots, T^n v)$ 线性相关，

所以 $\boxed{\text{存在不全为 } 0 \text{ 的数 } a_i \in F,\ i = 0, 1, \cdots, n}$

使得 $\boxed{a_0 v + a_1 T v + \cdots + a_n T^n v = 0}$ 即 $\left(\sum_{i=0}^{n} a_i T^i\right) v = 0$

Recall $v \neq 0$。这就证明了 $a_0 + \cdots + a_n T^n$ 不是单射。

因为 $v \neq 0$

故 $a_1, \cdots, a_n$ 不能全为 0

碰与 $\Leftarrow$ 存值

有限维复向量空间上的每个线性算子都有特征值

5.10 定理: 有限维复向量空间上的每个线性算子都有特征值

证明: 由上页命题知 存在 $a_0, \dots, a_n \in \mathbb{C}$ 使得

$$p(T) = a_0 I + \cdots + a_n T^n \quad ①$$

非单射, 且 $p(T)$ 的次数是大于 0 的. (注意 $p(T)$ 的次数可能小于 n) 因为 $a_1, \dots, a_n$ 不全为 0

由代数基本定理知 存在 $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ $(m \leq n)$, $c \in \mathbb{C}$ 使得

$$p(T) = c\,(T - \lambda_1 I) \cdots (T - \lambda_m I) \quad ②$$

由①②知, 必有某个 $T - \lambda_i I$ 非单射, 则 λ_i 是 T 的一个特征值. 证毕.

作业

1. 设 $T \in \mathcal{L}(V)$. 证明：若 V 的子空间 $U_1, \cdots, U_m$ 在 T 下都是不变的，则 $U_1 + \cdots + U_m$ 在 T 下也是不变的.

2. 设 $T \in \mathcal{L}(V)$. 证明 V 的任意一族在 T 下不变的子空间的交也是在 T 下不变的.

3. 证明或举反例：若 V 的子空间 U 在 V 的每个算子下都是不变的，则 $U = \{\mathbf{0}\}$ 或 $U = V$.

作业

作业3提示：

如果U是V的一个非平凡(即不是$\{0\}$, 又不是V)的子空间,

那么存在W(即不是$\{0\}$, 又不是V),使得

$V = U \oplus W$

设$U = span(u_1,\cdots,u_m), W = span(w_1,\cdots,w_n)$,

其中自然数$m > 1, n > 1$.

我们能否构造一个线性映射T,使得U 不是T 的不变子空间呢？

定义$Tu_1 = w_1, Tu_i = 0, i = 2,\cdots,m, Tw_i = 0, i = 1,\cdots,n$

验证一下，是否U是T的不变子空间?

4. 设 $S, T \in \mathcal{L}(V)$ 使得 $ST = TS$. 证明：对每个 $\lambda \in F$, $\operatorname{null}(T - \lambda I)$ 在 S 下都是不变的.

作业4提示：

$\operatorname{null}(T-\lambda I)=\{v\in V: Tv=\lambda v\}$

题目要证明

$\operatorname{null}(T-\lambda I)$是$S$ 的不变子空间，也就是说，

$\forall v\in$ $\operatorname{null}(T-\lambda I)$,仍然有$Sv\in$ $\operatorname{null}(T-\lambda I)$.

那我们就在在$v\in\operatorname{null}(T-\lambda I)$的条件下，验证

$(T-\lambda I)Sv=0$与否吧。

5. 定义 $T \in \mathcal{L}(\mathbf{F}^2)$ 如下

$$T(w, z) = (z, w).$$

求 T 的所有本征值和本征向量.

作业5提示：

和课上讲的一个例子很相似。 我们利用特征值和特征向量的定义来做。

先求特征值，设$\lambda \in F$是一个特征值，那么存在一个非零向量(a,b),使得

$T(a,b) = \lambda(a,b)$,

按照T的定义$T(a,b) = (b,a)$,也是我们有

$b = \lambda a, a = \lambda b$,

把后一个式子代入前一个，得到$b = \lambda^2 b$,

看看是是不是只能有$\lambda = -1, +1$,

然后再求特征向量。

6. 定义 $T \in \mathcal{L}(\mathbf{F}^3)$ 如下

$$T(z_1, z_2, z_3) = (2z_2, 0, 5z_3).$$

求 T 的所有本征值和本征向量.

作业6提示：

那么存在一个非零向量(x_1, x_2, x_3), 使得

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

再和T的定义相对照，我们可以得到

$\lambda x_1 = 2x_2$,

$\lambda x_2 = 0$,

$\lambda x_3 = 5x_3$

要注意特征向量不是唯一的

（如果v是T的特征向量，那么$av, a \neq 0$都是T的特征向量）。

7. 设 n 是正整数, 并设 $T \in \mathcal{L}(\mathbf{F}^n)$ 定义如下

$$T(x_1, \cdots, x_n) = (x_1 + \cdots + x_n, \cdots, x_1 + \cdots + x_n);$$

也就是说, 算子 T (关于标准基) 的矩阵的元素全是 1. 求 T 的所有本征值和本征向量.

作业7提示：

可以先尝试一下$n=1, n=2, n=3$,的情形，然后看看规律

就可以做出来。