5g t本章中更深入研究算子的结构，并将大多注意力放在复向量空间上．本章中的有些结论同时适用于实向量空间和复向量空间，于是我们并不总是假定 $\mathbf{F}=\mathbf{C}$．并且，内积对于本章内容并无用处，所以我们回到有限维向量空间这个一般的假定上来
即便是在有限维复向量空间上，一个算子也可能不具有足够的特征向量来形成向量空间的基．于是我们将考察一种名为广义特征向量的对象，它与上述问题联系紧密．我们将看到，对于有限维复向量空间上的每个算子，都存在由该算子的广义特征向量构成的该空间的基．届时，利用广义特征空间分解就能很好地描述有限维复向量空间上的任意算子
幂零算子是自乘若干次后会等于0 的算子，在上面这些问题的研究中，它们扮演着重要的角色．在我们证明有限维复向量空间上的每个可逆算子都有平方根，以及研究若当型的时候，幂零算子都是关键的工具
本章结尾处定义了迹，并证明了它的关键性质

8A 广义特征向量和幂零算子

算子的幂的零空间

> 8.1 递增的零空间序列
>
> $T\in\mathcal{L}(V)$，则
>
> $$\{0\}=\operatorname{null}T^0\subseteq\operatorname{null}T^1\subseteq\cdots\subseteq\operatorname{null}T^k\subseteq\operatorname{null}T^{k+1}\subseteq\cdots$$

证

设$k\in\mathbb{N}_0$，$\nu\in\operatorname{null}T^k$，则$T^k\nu=0$，故$T^{k+1}\nu=T(T^k\nu)=T(0)=0$，故$\nu\in\operatorname{null}T^{k+1}$，故$\operatorname{null}T^k\subseteq\operatorname{null}T^{k+1}$

下面结论表明若上述序列中存在相邻项相等，则这两项之后的所有项均相等
递减值域序列也有类似的结论（参习题6～8）

> 8.2 零空间序列的性质
>
> $T\in\mathcal{L}(V)$， $m\in\mathbb{N}_0$ 满足
>
> $$\operatorname{null}T^m=\operatorname{null}T^{m+1}$$
>
> 则
>
> $$\operatorname{null}T^m=\operatorname{null}T^{m+1}=\operatorname{null}T^{m+2}=\operatorname{null}T^{m+3}=\cdots$$

证

令 $k\in\mathbb{N}$
下面证

$$\operatorname{null}T^{m+k}=\operatorname{null}T^{m+k+1}$$

$\subseteq$)
由 8.1得
$\supseteq$)
设 $\nu\in\operatorname{null}T^{m+k+1}$，则$T^{m+1}(T^k\nu)=T^{m+k+1}\nu=0$，故

$$T^k\nu\in\operatorname{null}T^{m+1}=\operatorname{null}T^m$$

故$T^{m+k}\nu=T^m(T^k\nu)=0$，即$\nu\in\operatorname{null}T^{m+k}$

问：可能存在 m 使$\operatorname{null}T^m=\operatorname{null}T^{m+1}$成立吗?
下结论表明，至少在 m 大于等于$\dim V$ 时会成立

8.3 零空间停止增长

设 $T\in\mathcal{L}(V)$，则

$$\mathrm{null}T^{\dim V}=\mathrm{null}T^{\dim V+1}=\mathrm{null}T^{\dim V+2}=\cdots$$

证

只需证 $\mathrm{null}T^{\dim V}=\mathrm{null}T^{\dim V+1}$，这样由8.2可得证
（反证）设不成立，由8.1和8.2

$$\{0\}=\mathrm{null}T^0\subsetneq\mathrm{null}T^1\subsetneq\cdots\subsetneq\mathrm{null}T^{\dim V}\subsetneq\mathrm{null}T^{\dim V+1}$$

每个真包含号处维数至少增1
故$\dim\mathrm{null}T^{\dim V+1}\geq\dim V+1$，矛盾（因 V 的子空间的维数不大于 $\dim V$）

$V=\mathrm{null}T\oplus\mathrm{range}T$ 并不对每个 $T\in\mathcal{L}(V)$ 都成立。不过有下替代结论

8.4 V 是 null Tdim V 与 range Tdim V 的直和

$T\in\mathcal{L}(V)$，则

$$V=\mathrm{null}T^{\dim V}\oplus\mathrm{range}T^{\dim V}$$

证

令 $n=\dim V$
$(\cap=\{0\})$
设 $\nu\in(\mathrm{null}T^n)\cap(\mathrm{range}T^n)$，则$T^n\nu=0$ 且$\exists u\in V,\nu=T^nu$，故$T^n\nu=T^{2n}u$，故$T^{2n}u=0$，故$T^nu=0$（由 8.3，故$\nu=T^nu=0$，这证明了8.5
由8.5 null T^n+ range T^n 是直和（由1.46）且

$$\dim(\mathrm{null}T^n\oplus\mathrm{range}T^n)=\dim\mathrm{null}T^n+\dim\mathrm{range}T^n=\dim V$$

其中第一个等号由3.94，第二个等号由FTLM（3.21）
故$V=\mathrm{null}T^n\oplus\mathrm{range}T^n$

上述结论的加强版本见习题19

8.6 对于 $T\in\mathcal{L}(\mathbf{F}^3)$，$\mathbf{F}^3=\mathrm{null}T^3\oplus\mathrm{range}T^3$

$T\in\mathcal{L}(\mathbf{F}^3)$，$T(z_1,z_2,z_3)=(4z_2,0,5z_3)$
null $T=\{(z_1,0,0):z_1\in\mathbf{F}\}$
range $T=\{(z_1,0,z_3):z_1,z_3\in\mathbf{F}\}$
故null $T\cap$ range $T\neq\{0\}$，null T + range T 不是直和
同时注意到 $\mathrm{null}T+\mathrm{range}T\neq\mathbf{F}^3$，然而$T^3(z_1,z_2,z_3)=(0,0,125z_3)$，从而可知
因此正如 8.4 所述，$\mathbf{F}^3=\mathrm{null}T^3\oplus\mathrm{range}T^3$

广义特征向量

仅特征向量不足以很好地描述一些算子，引入广义特征向量：它将在描述算子的结构方面发挥重要作用。
为理解为何需要拓展特征向量，考虑将定义一算子的空间分解为若干不变子空间来描述该算子
$T\in\mathcal{L}(V)$，想通过找到一个好的直和分解$V=V_1\oplus\cdots\oplus V_n$（其中各 V_k 都是 V 的在 T 下不变的子空间）来描述 T，最简单的非零不变子空间是一维的。而当且仅当 V 具有由 T 的特征向量构成的基时，上述直和分解式中各 V_k 才全是 V 的在 T 下不变的一维子空间（见5.55）. 这又等价于 V 具有特征空间分解$V=E(\lambda_1,T)\oplus\cdots\oplus E(\lambda_m,T)$（其中 $\lambda_1,\ldots,\lambda_m$ 是 T 的所有互异特征值（见5.55））。由谱定理，若 V 是内积空间，形如8.7的分解对每个自伴算子（若

$\mathbf{F}=\mathbf{R}$）及每个正规算子（若 $\mathbf{F}=\mathbf{C}$ ）都成立，因这两类算子的特征向量足以形成 V 的基（见 7.29 和 7.31）
对更一般的算子形如8.7的分解不一定成立（即便在复向量空间上）：如5.57中所示算子的特征向量。现引入广义特征向量和广义特征空间解决这个问题。

8.8 广义特征向量

$T\in\mathcal{L}(V)$， λ 是 T 的特征值
称$\nu\in V$ 是 T 对应于 λ 的广义特征向量若 $\nu\neq 0$ 且$\exists k\in\mathbb{N}$，

$$(T-\lambda I)^k\nu=0$$

$\nu\in V\ \{0\}$ 是 T 对应于 λ 的广义特征向量 $\iff$ $(T-\lambda I)^{\dim V}\nu=0$
将 8.1 和 8.3 应用于算子 $T-\lambda I$ 即可得出这点
我们并未定义广义特征值，因为不会得到任何新东西（若 $(T-\lambda I)^k$ 对某个k不是单射，则$T-\lambda I$ 也不是单射，因此 λ 就是 T 的特征值）
我们知道复向量空间上的算子的特征向量可能不足以形成定义空间的基。下面结论说明，这些算子的广义特征向量总能形成定义空间的基

8.9 由广义特征向量构成的基

$\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$，则存在由 T 的广义特征向量构成的 V 的基

证

令 $n=\dim V$，对 n 强归纳
 $(n=1)$
成立，因为此时 V 中每个非零向量都是 T 的特征向量
 $(n>1)$
设欲证结论在 $\dim V$ 为更小的值时都成立．设 λ 是 T 的特征值．将8.4 应用于 $T-\lambda I$（正是这一步用上了 $\mathbf{F}=\mathbf{C}$ 这个前提条件，因为若 $\mathbf{F}=\mathbf{R}$ ，那么 T 可能不存在特征值）得

$$V=\operatorname{null}(T-\lambda I)^n\oplus\operatorname{range}(T-\lambda I)^n$$

若 $\operatorname{null}(T-\lambda I)^n=V$ ，那么 V 中每个非零向量都是 T 的广义特征向量，于是此时存在由 T 的广义特征向量构成的 V 的基
可设 $\operatorname{null}(T-\lambda I)^n\neq V$ ，故range $(T-\lambda I)^n\neq\{0\}$
同时 $\operatorname{null}(T-\lambda I)^n\neq\{0\}$，因为 λ 是 T 的一个特征值
有

$$0<\operatorname{dim}\operatorname{range}(T-\lambda I)^n<n$$

此外，range $(T-\lambda I)^n$ 在 T 下不变【由5.18，其中取 $p(z)=(z-\lambda)^n$ 】．令 $S\in\mathcal{L}(\operatorname{range}(T-\lambda I)^n)$ 等于限制在 range $(T-\lambda I)^n$ 上的算子 T ．将归纳假设应用于算子 s ，可得存在由 s 的广义特征向量构成的 range $(T-\lambda I)^n$ 的基，而这些向量当然也是 T 的广义特征向量．将range $(T-\lambda I)^n$ 的这个基与 $\operatorname{null}(T-\lambda I)^n$ 的基合并，就得到了由 T 的广义特征向量构成的 V 的基

若 $\mathbf{F}=\mathbf{R}$ 且 $\dim V>1$ ，那么 V 上有些算子的广义特征向量可构成 V 的基，其他算子则不具有该性质（与 $\mathbf{F}=\mathbf{C}$ 情形不同）．判定一算子是否具有该性质的充分必要条件见于习题11

8.10 $\mathbf{C}^3$ 上一算子的广义特征向量

$T\in\mathcal{L}(\mathbf{C}^3)$，$\forall(z_1,z_2,z_3)\in\mathbf{C}^3,T(z_1,z_2,z_3)=(4z_2,0,5z_3)$
由特征值的定义 T 的特征值是0和5
对应于特征值 0 的特征向量是形如 $(z_1,0,0)$ 的非零向量
对应于特征值 5 的特征向量是形如 $(0,0,z_3)$ 的非零向量
故该算子的特征向量不够张成该算子的定义空间 $\mathbf{C}^3$
知道 $T^3(z_1,z_2,z_3)=(0,0,125z_3)$，由 8.1 和 8.3， T 对应于特征值 0 的广义特征向量是形如 $(z_1,z_2,0)$ 的非零向量
同时 $(T-5I)^3(z_1,z_2,z_3)=(-125z_1+300z_2,-125z_2,0)$，故 T 对应于特征值 5 的广义特征向量是形如 $(0,0,z_3)$ 的非零向量
由上面几段的论述， $\mathbf{C}^3$ 的标准基中每个向量都是 T 的广义特征向量
正如8.9所言， $\mathbf{C}^3$ 的确具有由 T 的广义特征向量构成的基

若 ν 是 $T \in \mathcal{L}(V)$ 的特征向量，其对应的特征值 λ 由 $T\nu = \lambda\nu$ 唯一确定，且只对一个 $\boldsymbol{\lambda} \in \mathbf{F}$ 成立（因 $\nu \neq 0$）
若 ν 是 T 的广义特征向量，不知道是否仅有一个 $\boldsymbol{\lambda} \in \mathbf{F}$ 能满足方程 $(T - \lambda I)^{\dim V}\nu = 0$
下面结论表明，该问题的回答是肯定的

> **8.11 广义特征向量对应于唯一的特征值**
>
> $T \in \mathcal{L}(V)$，则T 的每个广义特征向量都仅对应于 T 的一个特征值

证

设 $\nu \in V$ 是 T 的同时对应于 T 的两特征值 α 和 λ 的广义特征向量
令 m 是满足$(T - \alpha I)^m\nu = 0$ 的最小正整数
令 $n = \dim V$，则

$$
\begin{aligned}
0 &= (T - \lambda I)^n \nu \\
&= \big((T - \alpha I) + (\alpha - \lambda)I\big)^n \nu \\
&= \sum_{k=0}^{n} b_k (\alpha - \lambda)^{n-k} (T - \alpha I)^k \nu
\end{aligned}
$$

其中 $b_0 = 1$，其余各二项式系数 b_k 的值无关紧要
将算子 $(T - \alpha I)^{m-1}$ 作用于上式两边，可得

$$0 = (\alpha - \lambda)^n (T - \alpha I)^{m-1} \nu$$

由 $(T - \alpha I)^{m-1}\nu \neq 0$，只能$(\alpha - \lambda)^n = 0$，即$\alpha = \lambda$

知道对应于不同特征值的特征向量是线性无关的（5.11），现证适用于广义特征向量的类似结论，思路大差不差

> **8.12 线性无关的广义特征向量**
>
> $T \in \mathcal{L}(V)$，则由对应于 T 的互异特征值的广义特征向量构成的每个向量组都线性无关

证

（反证）设结论不成立，则存在最小正整数 m 使对应于 T 的互异特征值 $\lambda_1, \ldots, \lambda_m$ 的广义特征向量 $\nu_1, \ldots, \nu_m$ 构成线性相关向量组 （注意 $m \geq 2$，因为由定义广义特征向量非零），故存在全不为0（因为 m 最小）的数 $a_1, \ldots, a_m \in \mathbf{F}$，

$$a_1\nu_1 + \cdots + a_m\nu_m = 0$$

令 $n = \dim V$，两边作用$(T - \lambda_m I)^n$

$$a_1(T - \lambda_m I)^n \nu_1 + \cdots + a_{m-1}(T - \lambda_m I)^n \nu_{m-1} = 0$$

设 $k \in \{1, \ldots, m-1\}$，则

$$(T - \lambda_m I)^n \nu_k \neq 0$$

否则 ν_k 就会成为 T 的同时对应于不同特征值 λ_k 和 λ_m 的广义特征向量，这会与 8.11 相矛盾
又有

$$(T - \lambda_k I)^n \big((T - \lambda_m I)^n \nu_k\big) = (T - \lambda_m I)^n \big((T - \lambda_k I)^n \nu_k\big) = 0$$

综上 $(T - \lambda_m I)^n \nu_k$ 是 T 对应于特征值 λ_k 的广义特征向量，故

$$(T - \lambda_m I)^n \nu_1, \ldots, (T - \lambda_m I)^n \nu_{m-1}$$

是由 $m - 1$ 个对应于互异特征值的广义特征向量所构成的线性相关组【由式(8.13)】，与 m 最小矛盾

幂零算子

8.14 幂零

称一算子幂零若它的某幂为0

算子 $T \in \mathcal{L}(V)$ 幂零若 V 中每非零向量都是 T 对应于特征值 0 的广义特征向量

8.15 幂零算子

(a) 定义为$T(z_1, z_2, z_3, z_4) = (0, 0, z_1, z_2)$的算子 $T \in \mathcal{L}(\mathbf{F}^4)$ 幂零，因$T^2 = 0$
(b) 关于标准基的矩阵为

$$\begin{pmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{pmatrix}$$

的 $\mathbf{F}^3$ 上的算子是幂零的，因该矩阵的立方是零矩阵
(c) $\mathcal{P}_m(\mathbf{R})$ 上的微分算子幂零，因每个不超过 m 次的多项式的 $m+1$ 阶导为0

下结论说明当求一幂零算子的幂时，不用考虑幂次高于其定义空间的维数的幂。一个稍微强一些的结论见习题18

8.16 n 维空间上幂零算子的 n 次幂等于0

$T \in \mathcal{L}(V)$幂零$\Rightarrow T^{\dim V} = 0$

拉丁文词语“nil” 意为“无”或“零”，“potens”意为“幂”。故 “nilpotent” 的字面意思就是“幂零”

证

由幂零，$\exists k \in \mathbb{N}$，$T^k = 0$
故null $T^k = V$
由 8.1 和 8.3 $\operatorname{null} T^{\dim V} = V$
故 $T^{\dim V} = 0$

8.17 幂零算子的特征值

$T \in \mathcal{L}(V)$
(a) T 幂零，则 0 是 T 的特征值且 T 无其他的特征值
(b) 若 $\mathbf{F} = \mathbf{C}$ 且 0 是 T 的唯一特征值，则 T 幂零

证

(a) 设 T 幂零，故 $\exists m \in \mathbb{N}$，$T^m = 0$，故T 不是单射，故 0 是 T 的特征值
为证 T 无其他特征值，设 λ 是 T 的特征值，则 $\exists \nu \in V \ \{0\}$，$\lambda \nu = T \nu$
将 T 反复作用于上式两端得$\lambda^m \nu = T^m \nu = 0$，故 $\lambda = 0$
(b) 设 $\mathbf{F} = \mathbf{C}$ 且 0 是 T 的唯一特征值，由 5.27 (b) T 的最小多项式等于 $z^m \quad (m \in \mathbb{N})$，故 $T^m = 0$，故 T 幂零

习题23说明上述结论(b)中 $\mathbf{F} = \mathbf{C}$ 不可删去
给定 V 上一算子，想求其一基使该算子关于该基的矩阵尽可能简单，即矩阵中包含尽可能多的0。下结论说明若 T 幂零则可取 V 的一基使 T 关于该基的矩阵有超一半的元素为0
本章还将得到更好的结论

> 8.18 幂零算子的最小多项式和上三角矩阵
>
> $T \in \mathcal{L}(V)$，则下面各命题等价
> (a) T 幂零
> (b) T 的最小多项式为 z^m（$m \in \mathbb{N}$）
> (c) $\exists V$ 的一个基， T 关于该基的矩阵为主对角线全为0的上三角矩阵

证

设(a)成立，则 T 幂零，故存在正整数 n，$T^n = 0$
由5.29 z^n 是 T 的最小多项式的多项式倍，故 T 的最小多项式为 z^m （ m 为正整数）．(a) 蕴涵(b)
设(b)成立，则 T 的最小多项式是 z^m （其中 m 为正整数）．由此，根据5.27 (a) 可得，0（ z^m 的唯一零点）是 T 的唯一特征值；根据 5.44 还有，存在 V 的一个基，使得 T 关于该基具有上三角矩阵．进而，由5.41该矩阵中主对角线上元素全为0．(b)蕴涵 (c)
设(c)成立，由5.40 $T^{\dim V} = 0$，故 T 幂零，(c) 蕴涵(a)

8B 广义特征空间分解

广义特征空间

8.19 广义特征空间

设 $T \in \mathcal{L}(V)$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$． T 对应于 λ 的广义特征空间，记作 $G(\lambda, T)$，定义为

于是 $G(\lambda, T)$ 是由 T 对应于 λ 的广义特征向量以及0 向量所构成的集合
因为 T 的每个特征向量都是 T 的广义特征向量（在广义特征向量的定义中取 $k=1$ 即可），所以每个特征空间都包含于相对应的广义特征空间．换言之，若 $T \in \mathcal{L}(V)$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$，那么$E(\lambda, T) \subseteq G(\lambda, T)$ ·
下面结论表明，如果 $T \in \mathcal{L}(V)$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$，那么广义特征空间 $G(\lambda, T)$ 是 V 的一个子空间（因为 V 上每个线性映射的零空间都是 V 的子空间）

> 8.20 广义特征空间的描述
>
> $T \in \mathcal{L}(V)$ 且 $\boldsymbol{\lambda} \in \mathbf{F}$．那么 $G(\lambda, T) = \operatorname{null}(T - \lambda I)^{\dim V}$

证

设 $\nu \in \operatorname{null}(T - \lambda I)^{\dim V}$．由广义特征空间的定义知 $\nu \in G(\lambda, T)$
则 $G(\lambda, T) \supseteq \operatorname{null}(T - \lambda I)^{\dim V}$
设 $\nu \in G(\lambda, T)$，则存在正整数 k， $\nu \in \operatorname{null}(T - \lambda I)^k$
由 8.1 和 8.3 （将其中 T 用 $T - \lambda I$ 替代） $\nu \in \operatorname{null}(T - \lambda I)^{\dim V}$
故$G(\lambda, T) \subseteq \operatorname{null}(T - \lambda I)^{\dim V}$

8.21 $\mathbf{C}^3$ 上一算子的广义特征空间

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

$$T(z_1, z_2, z_3) = (4z_2, 0, 5z_3)$$

在例8.10 中，我们得知 T 的特征值是0 和5，且求出了与之对应的广义特征向量所构成的集合．将这些集合分别与 $\{0\}$ 取并集，我们有

$$G(0, T) = \{(z_1, z_2, 0) : z_1, z_2 \in \mathbf{C}\}$$

$$G(5,T)=\{(0,0,z_3):z_3\in\mathbf{C}\}$$

注意， $\mathbf{C}^3=G(0,T)\oplus G(5,T)$

在例 8.21 中，定义空间 $\mathbf{C}^3$ 是该例中算子 T 的广义特征空间的直和．接下来的结论就表明，这个性质是普遍成立的．具体而言，下面这个关键结论说明了，如果 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$ ，那么 V 是 T 的广义特征空间的直和，该直和中各项都在 T 下不变，并且作用于其上的 T 等于一个幂零算子加上恒等算子的标量倍．于是，借助下面结论，我们得以完成我们的目标：将 V 分解成不变子空间，且 T 在这些子空间上的性质为我们所知
我们将看到，以下证明过程综合运用了目前所学有关广义特征空间的知识，随后利用了这个结论：对于每个算子 $T\in\mathcal{L}(V)$ ，都存在由 T 的广义特征向量构成的 V 的基

8.22 广义特征空间分解

$\mathbf{F}=\mathbf{C}$ ， $T\in\mathcal{L}(V)$ 。令 $\lambda_1,\ldots,\lambda_m$ 是 T 的所有互异特征值，则
(a) 对每个 $k=1,\ldots,m$ ， $G(\lambda_k,T)$ 在 T 下不变
(b) 对每个 $k=1,\ldots,m$ ， $(T-\lambda_kI)|_{G(\lambda_k,T)}$ 幂零
(c) $V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T)$

证

(a) 设 $k\in\{1,\ldots,m\}$ ，由8.20

$$G(\lambda_k,T)=\operatorname{null}(T-\lambda_kI)^{\dim V}$$

由 5.18 【其中取 $p(z)=(z-\lambda_k)^{\dim V}$ 】其在 T 下不变
(b) 设 $k\in\{1,\ldots,m\}$ ，若 $\nu\in G(\lambda_k,T)$ ，那么 $(T-\lambda_kI)^{\dim V}\nu=0$ （由8.20）．于是$\left((T-\lambda_kI)|_{G(\lambda_k,T)}\right)^{\dim V}=0$ ，因此 $(T-\lambda_kI)|_{G(\lambda_k,T)}$ 幂零
(c) 为了说明 $G(\lambda_1,T)+\cdots+G(\lambda_m,T)$ 是直和，设

$$\nu_1+\cdots+\nu_m=0$$

其中各 | ν_k 属于 $G(\lambda_k,T)$ ．因为 T 对应于互异特征值的广义特征向量线性无关（由 8.12），所以此式中各 ν_k 等于0．从而 $G(\lambda_1,T)+\cdots+G(\lambda_m,T)$ 是直和（由1.45）最后， V 中每个向量都可被写成 T 的广义特征向量的有限和的形式（由8.9）．于是

$$V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T)$$

$\mathbf{F}=\mathbf{R}$ 时的类似结论见于习题8

特征值的重数

若 V 是复向量空间且 $T\in\mathcal{L}(V)$ ，那么由广义特征空间分解（8.22）给出的 V 的分解式是个有力的工具．该分解式中各子空间的维数很重要，因而在下面定义中被赋予特殊的名称

8.23 重数

$T\in\mathcal{L}(V)$ ．定义 T 的特征值 λ 的重数为其对应的广义特征空间 $G(\lambda,T)$ 的维数换言之， T 的特征值 λ 的重数等于

$$\dim\operatorname{null}(T-\lambda I)^{\dim V}$$

上述第二点成立是因为 $G(\lambda,T)=\operatorname{null}(T-\lambda I)^{\dim V}$ （见 8.20）

8.24 一算子的各特征值的重数

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

$$T(z_1,z_2,z_3)=(6z_1+3z_2+4z_3,6z_2+2z_3,7z_3)$$

其关于标准基的矩阵

$$\begin{pmatrix}6&3&4\\0&6&2\\0&0&7\end{pmatrix}$$

由 5.41 可得，T 的特征值是矩阵对角线上的元素6 和7．你可自行验证，T 的广义特征空间是$G(6,T)=\operatorname{span}\big((1,0,0),(0,1,0)\big)$ 和
$G(7,T)=\operatorname{span}\big((10,2,1)\big)$
于是，特征值 6 的重数是 2，特征值 7 的重数是 1．由 8.22 所述广义特征空间分解可写出直和 $\mathbf{C}^3=G(6,T)\oplus G(7,T)$．如8.9 所言，$T$ 的广义特征向量 $(1,0,0),(0,1,0),(10,2,1)$ 构成 $\mathbf{C}^3$ 的一个基．$\mathbf{C}^3$ 中不存在由该算子的特征向量构成的基

上例中 T 的特征值的重数之和等于3，正是 T 的定义空间的维数。下面结论说明，这个性质对于有限维复向量空间上的每个算子都成立
在本例中，每个特征值的重数，都等于该特征值在算子的上三角矩阵对角线上出现的次数．我们将在 8.31 中看到，这个性质总是成立的

8.25 重数之和等于dim$V$

$\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$．那么 T 的所有特征值的重数之和等于 $\dim V$

证

由广义特征空间分解（8.22）与直和的维数公式（见3.94）得

有些书中会使用代数重数（algebraic multiplicity）和几何重数（geometric multiplicity）这两个术语．如果碰到这两个术语，你应明白代数重数与此处定义的重数是一样的，而几何重数是相应特征空间的维数
注意，按照上述定义，代数重数作为某个零空间的维数，同样有几何意义
此处给出的重数定义比涉及行列式的传统定义更加简洁，9.62 将说明这两种定义是等价的
若 V 是内积空间，$T\in\mathcal{L}(V)$ 是正规的，并且 λ 是 T 的一个特征值，那么将7A 节习题27 应用于正规算子 $T-\lambda I$ 上，即可见 λ 的代数重数等于 λ 的几何重数．由此顺便可得，$S\in\mathcal{L}(V,W)$（此处 V 和 W 都是有限维内积空间）的奇异值依赖于自伴算子 S^*S 的特征值的重数（代数重数或几何重数）
下面定义将有限维复向量空间上的每个算子都与一个首一多项式联系起来

8.26 特征多项式

$\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$．令 $\lambda_1,\ldots,\lambda_m$ 表示 T 的所有互异特征值，且其重数分别为$d_1,\ldots,d_m$．称多项式

$$(z-\lambda_1)^{d_1}\cdots(z-\lambda_m)^{d_m}$$

为 T 的特征多项式

8.27 一个算子的特征多项式

$T\in\mathcal{L}(\mathbf{C}^3)$ 定义如例8.24．因为 T 的特征值6 的重数是2，特征值7 的重数是1，所以我们可得 T 的特征多项式是 $(z-6)^2(z-7)$

8.28 特征多项式的次数和零点

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$．那么
(a) T 的特征多项式的次数是 $\dim V$
(b) T 的特征多项式的零点就是 T 的特征值

证

由关于重数之和的结论（8.25）可得(a)．由特征多项式的定义可得(b)

多数教材利用行列式定义特征多项式（由 9.62，它和此处的定义等价）．此处采用的处理方法要简洁许多，并可引出凯莱-哈密尔顿定理的下面这个漂亮的证明

8.29 凯莱-哈密尔顿定理

$\mathbf{F}=\mathbf{C}$ ， $T\in\mathcal{L}(V)$ ，且 q 是 T 的特征多项式，则 $q(T)=0$

证

设 $\lambda_1,\dots,\lambda_m$ 是 T 的所有互异特征值，并令 $d_k=\dim G(\lambda_k,T)$ ．我们知道，对每个 $k\in\{1,\dots,m\}$ ， $(T-\lambda_k I)|_{G(\lambda_k,T)}$ 都是幂零的．于是由8.16，我们有：对每个 $k\in\{1,\dots,m\}$

$$(T-\lambda_k I)^{d_k}|_{G(\lambda_k,T)}=0$$

广义特征空间分解（8.22）指出， V 中每个向量都是 $G(\lambda_1,T),\dots,G(\lambda_m,T)$ 中向量的
为证 $q(T)=0$ ，只需证对各 k 均有 $q(T)|_{G(\lambda_k,T)}=0$
固定 $k\in\{1,\dots,m\}$ ，

$$q(T)=(T-\lambda_1 I)^{d_1}\cdots(T-\lambda_m I)^{d_m}$$

上式右侧的算子都是可交换的，因此我们可将乘积项 $(T-\lambda_k I)^{d_k}$ 移至右侧表达式的最后一项因为 $(T-\lambda_k I)^{d_k}|_{G(\lambda_k,T)}$ 等于0，故 $q(T)|_{G(\lambda_k,T)}=0$

下面结论表明，如果一算子 $T\in\mathcal{L}(V)$ 的最小多项式的次数是 $\dim V$ （大多数情况下都是如此——见5.24 之后的几段话），那么 T 的特征多项式就等于 T 的最小多项式

8.30 特征多项式是最小多项式的多项式倍

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$ ．那么 T 的特征多项式是 T 的最小多项式的多项式倍

8.30 特征多项式是最小多项式的多项式倍

由凯莱-哈密尔顿定理（8.29）和5.29 立刻可得上述结论现在，我们可以证明，例8.24 体现出的结论对于有限维复向量空间上的所有算子都成立

8.31 特征值的重数等于其在对角线上出现的次数

$\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$ ．设 $\nu_1,\dots,\nu_n$ 是 V 的一个基且使得 $M(T,(\nu_1,\dots,\nu_n))$ 为上三角矩阵．那么 T 的每个特征值 λ 在 $M(T,(\nu_1,\dots,\nu_n))$ 对角线上出现的次数，就等于 λ 作为 T 的特征值的重数

证

令 $A=\mathcal{M}\big(T,(\nu_1,\dots,\nu_n)\big)$ ．于是 A 是上三角矩阵．令 $\lambda_1,\dots,\lambda_n$ 表示 A 对角线上的各元素．于是对每个 $k\in\{1,\dots,n\}$ $T\nu_k=u_k+\lambda_k\nu_k$
其中 $u_k\in\operatorname{span}(\nu_1,\dots,\nu_{k-1})$ ．因此若 $k\in\{1,\dots,n\}$ 且 $\lambda_k\neq 0$ ，那么 $T\nu_k$ 不是 $T\nu_1,\dots,T\nu_{k-1}$ 的线性组合．现由线性相关性引理（2.19）可得由这些 $T\nu_k$ （满足 $\lambda_k\neq 0$ ）所构成的组是线性无关的
令 d 表示各下标 $k\in\{1,\dots,n\}$ 中使得 $\lambda_k=0$ 的个数．由上段结论可知
因为 $n=\dim V=\dim\operatorname{null}T+\dim\operatorname{range}T$ ，所以上述不等式表明$\dim\operatorname{null}T\le d$
算子 T^n 关于基 $\nu_1,\dots,\nu_n$ 的矩阵是上三角矩阵 A^n ，且其对角线上的元素是 $\lambda_1^n,\dots,\lambda_n^n$ 【见5C 节习题 2 (b)】．因为 $\lambda_k^n=0$ 当且仅当 $\lambda_k=0$ ，所以 A^n 的对角线上出现 0 的次数就等于 d 于是利用式(8.32)（将其中的 T 替换为 T^n），我们就有
$\dim\operatorname{null}T^n\le d$
对于 T 的一个特征值 λ ，令 m_λ 表示 λ 作为 T 的特征值的重数，并令 d_λ 表示 λ 出现在 A 的对角线上的次数．将式(8.33) 中的 T 替换成 $T-\lambda I$ ，我们就会发现

$$m_\lambda\le d_\lambda$$

对 T 的每个特征值 λ 都成立． T 的所有特征值 λ 的重数 m_λ 之和等于 n ，也就是 V 的维数（由8.25）． T 的所有特征值 λ 的出现次数 d_λ 之和也等于 n ，因为 A 的对角线上共有 n 个数

因此，将式(8.34) 的两端对 T 的所有特征值 λ 求和会得到一个等式．因此，式(8.34) 对于T 的每个特征值 λ 也一定是个等式．于是 λ 作为 T 的特征值的重数，等于 λ 在 A 的对角线上出现的次数

分块对角矩阵

为了以矩阵形式解读我们的结论，我们通过推广对角矩阵的概念，作出如下定义

8.35 分块对角矩阵

一个分块对角矩阵是形如

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

的方阵，其中 $A_1, \ldots, A_m$ 是排列在对角线上的方阵，且矩阵其他各元素都等于0

（若各矩阵 A_k 都是 1×1 矩阵，那么我们其实就得到了对角矩阵）

8.36 一个分块对角矩阵

形如

$$A = \begin{pmatrix} A_1 & & 0 \\ & A_2 & \\ 0 & & A_3 \end{pmatrix},$$

5×5 的分块对角矩阵，其中

$$A_1 = (4), \quad A_2 = \begin{pmatrix} 2 & -3 \\ 0 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 7 \\ 0 & 1 \end{pmatrix}$$

这里，我们将 5×5 矩阵分解出一块一块的内部矩阵，以便说明我们是怎么将其看成分块对角矩阵的
注意到，在上例中， A_1, A_2, A_3 都是上三角矩阵，且各自对角线上元素都相等．下面结论表明，有限维复向量空间中每个算子关于适当的基都具有该形式的矩阵．注意，这个结论给出的矩阵比一般的上三角矩阵具有更多的零

8.37 由上三角块构成的分块对角矩阵

$\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$． 令 $\lambda_1, \ldots, \lambda_m$ 是 T 的所有互异特征值，它们的重数分别为$d_1, \ldots, d_m$． 那么存在 V 的一个基，使得 T 关于该基具有形如

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

的分块对角矩阵，其中各 A_k 是形如

$$A_k = \begin{pmatrix} \lambda_k & & * \\ & \ddots & \\ 0 & & \lambda_k \end{pmatrix}$$

的 $d_k \times d_k$ 上三角矩阵

证

每个 $(T-\lambda_k I)|_{G(\lambda_k,T)}$ 都是幂零的（见 8.22）．对各 k，选取 $G(\lambda_k,T)$（它是维数为 $_k$ 的向量空间）的一个基，使得 $(T-\lambda_k I)|_{G(\lambda_k,T)}$ 具有形如 8.18 (c) 的矩阵．而 $T|_{G(\lambda_k,T)} = (T-\lambda_k I)|_{G(\lambda_k,T)} + \lambda_k I|_{G(\lambda_k,T)}$，于是 $T|_{G(\lambda_k,T)}$ 关于该基的矩阵就具有上面所示 A_k 的形式

广义特征空间分解（8.22）表明，将上面选取的各 $G(\lambda_k,T)$ 的基合并起来，就得到 V 的一个基． T 关于该基的矩阵就具有我们期望的形式

8.38 由广义特征向量得出分块对角矩阵

$T \in \mathcal{L}(\mathbf{C}^3)$， $T(z_1,z_2,z_3) = (6z_1+3z_2+4z_3, 6z_2+2z_3, 7z_3)$． T（关于标准基）的矩阵是

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix},$$

这是上三角矩阵，但不是8.37 所给出的形式
在例8.24 中我们看到， T 的特征值是6 和7，且有

$$G(6,T) = \operatorname{span}\big((1,0,0),(0,1,0)\big)$$

我们还知道由 T 的广义特征向量所构成的 $\mathbf{C}^3$ 的基是

$$(1,0,0),(0,1,0),(10,2,1)$$

T 关于该基的矩阵是

$$\begin{matrix} 6 & 3 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{matrix}$$

该矩阵具有 8.37 所给出的分块对角形式

8C 广义特征空间分解的推论

算子的平方根

回忆一下，算子 $T \in \mathcal{L}(V)$ 的平方根是满足 $R^2 = T$ 的算子 $R \in \mathcal{L}(V)$（见7.36）．每个复数都有平方根，但复向量空间上并非每个算子都有平方根．例如，$\mathbf{C}^3$ 上定义为 $T(z_1,z_2,z_3) = (z_2,z_3,0)$ 的算子没有平方根，本节习题1 就要求你证明这一点．很快我们将看到，这个无平方根的算子不可逆并非偶然．我们首先证明恒等算子加上每个幂零算子都有平方根

8.39 恒等算子加上幂零算子有平方根

$T \in \mathcal{L}(V)$ 是幂零的．那么 $I+T$ 有平方根

证

考虑函数 $\sqrt{1+x}$ 的泰勒级数

$$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots$$

我们并不需要求出系数的确切表达式，也无需担心无限和是否收敛，因为我们仅仅是将该式用作引子
因为 $a_1 = \frac{1}{2}$，所以上述公式表明，当 x 很小时， $1+\frac{x}{2}$ 可很好地近似 $\sqrt{1+x}$
因为 T 是幂零的，所以对某正整数 ∇m 有 $T^m = 0$．假设我们将式 (8.40) 中的 x 换成 T，1 换成 I．那么式子右端的无限和就变为有限和（因为对所有 $k \geq m$ 有 $T^k = 0$）．于是我们猜想$I+T$ 具有形如

$$I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}$$

的平方根．作出猜想后，我们尝试找出 $a_1, a_2, \ldots, a_{m-1}$ 使得上述算子的平方等于 $I+T$．则

$$(I + a_1 T + a_2 T^2 + a_3 T^3 + \cdots + a_{m-1} T^{m-1})^2$$

$$= I + 2a_1T + (2a_2 + a_1^2)T^2 + (2a_3 + 2a_1a_2)T^3 + \cdots$$

我们想让上式右侧等于 $I+T$．因此取 a_1 使得 $2a_1 = 1$ （于是 $a_1 = \frac{1}{2}$ ）．接下来，取 a_2 使得$2a_2 + a_1^2 = 0$ （于是 $a_2 = -\frac{1}{8}$ ）．然后选取 a_3 使得上式中 T^3 前系数为 0（于是 $a_3 = \frac{1}{16}$ ）．对每个$k = 4, \ldots, m-1$ 都按此方式求解下去，即在每步都求出使得上式右侧 T^k 项的系数等于0 的 a_k 事实上我们不关心各 a_k 的确切值，而只需要知道存在一组 a_k 能使 $I + a_1T + \cdots + a_{m-1}T^{m-1}$ 等于 $I+T$ 的平方根

上述引理在实向量空间和复向量空间上都成立．然而，下面的结论则只在复向量空间上成立．例如，一维实向量空间 $\mathbf{R}$ 上"与 $^{-1}$ 相乘"这个算子就没有平方根

找不到"images/c216303f55bfa161af1c5efaa5d12a4e1d72b23318d9c714913573f1185cd387.jpg"。

为证明下一条结果，我们需要知道每个复数 $z \in \mathbf{C}$ 都有属于 C 的平方根为此，将 z 写成

$$z = r(\cos\theta + \mathrm{i}\sin\theta)$$

用极坐标来表达一个复数
其中 $\boldsymbol{r}$ 是复平面上从原点到 z 点的线段长度， θ 是该线段与横轴正方向的夹角．那么

$$\sqrt{r}\left(\cos\frac{\theta}{2} + \mathrm{i}\sin\frac{\theta}{2}\right)$$

就是 z 的平方根．你可通过证明上面这个复数的平方等于 z 来验证这点

8.41 C 上，可逆算子具有平方根

设 V 是复向量空间， $T \in \mathcal{L}(V)$ 是可逆的．那么 T 有平方根

证

令 $\lambda_1, \ldots, \lambda_m$ 是 T 的所有互异特征值．对每个 k 都存在幂零算子 $T_k \in \mathcal{L}(G(\lambda_k, T))$ 使得$T|_{G(\lambda_k,T)} = \lambda_k I + T_k$ 【见 8.22 (b)】．因为 T 是可逆的，所以各 λ_k 均不等于 0，因此我们可对每个 k 写出

$$T|_{G(\lambda_k,T)} = \lambda_k\left(I + \frac{T_k}{\lambda_k}\right)$$

因为 T_k/λ_k 是幂零的，所以 $I + T_k/\lambda_k$ 具有平方根（由 8.39）．将复数 λ_k 的平方根和 $I + T_k/\lambda_k$ 的平方根相乘，我们就可得 $T|_{G(\lambda_k,T)}$ 的平方根 R_k
由广义特征空间分解（8.22），向量 $\nu \in V$ 可被唯一写成

$$\nu = u_1 + \cdots + u_m$$

其中各 u_k 属于 $G(\lambda_k, T)$．利用这个分解式，定义算子 $R \in \mathcal{L}(V)$ 为

$$R\nu = R_1u_1 + \cdots + R_mu_m$$

你应自行验证，这个算子 R 是 T 的平方根．证毕

可效仿本小节中使用的技巧证明：
如果 V 是复向量空间且 $T \in \mathcal{L}(V)$ 可逆，那么T 具有任意正整数次方根

若当型

我们知道，如果 V 是复向量空间，那么对每个 $T \in \mathcal{L}(V)$，都存在 V 的一个基，使得 T 关于该基有上三角矩阵（见 8.37），这已经是个很好的结论．在本小节中我们会得到更好的结论—存在 V 的一个基，使得 T 关于该基的矩阵除了对角线及紧挨在对角线正上方的元素之外，其他各元素都是0
先看两个有关幂零算子的例子

8.42 具有很好的矩阵的幂零算子

$\mathbf{C}^4$ 上定义算子

$$T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3)$$

那么 $T^4 = 0$，从而 T 是幂零的．若 $\boldsymbol{\nu} = (1, 0, 0, 0)$，那么 $T^3\nu, T^2\nu, T\nu, \nu$ 是 $\mathbf{C}^4$ 的基． T 关于该基的矩阵是

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

8.43 具有稍复杂点的矩阵的幂零算子

T 是 $\mathbf{C}^6$ 上定义为

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0)$$

的算子．那么 $T^3 = 0$，从而 T 是幂零的．该算子不像上例中的幂零算子那样容易处理，因为并不存在向量 $\nu \in \mathbf{C}^6$ 使得 $T^5\nu, T^4\nu, T^3\nu, T^2\nu, T\nu, \nu$ 是 $\mathbf{C}^6$ 的一个基．然而，如果我们取 $\nu_1 = (1, 0, 0, 0, 0, 0)$, $\nu_2 = (0, 0, 0, 1, 0, 0), \nu_3 = (0, 0, 0, 0, 0, 1)$，那么 $T^2\nu_1, T\nu_1, \nu_1, T\nu_2, \nu_2, \nu_3$ 是 $\mathbf{C}^6$ 的一个基． T 关于该基的矩阵是

$$\begin{pmatrix} \begin{matrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & (0) \end{pmatrix} & \end{pmatrix}.$$

我们将上面这个 6×6 矩阵分解出«一块一块的内部矩阵，以¬便说明我们是怎么将其看成分块对角矩阵的：它包含一个 3×3 的块（其中紧挨在对角线正上方的元素是 1，其余元素均为 0)，一个 2×2 的块（其中紧挨在对角线正上方的元素是1，其余元素均为0)，以及一个包含0 的1×1 的块
接下来的目标是证每个幂零算子 $T \in \mathcal{L}(V)$ 都和上例中的算子具有类似性质具体而言，存在有限多个向量 $\nu_1, \ldots, \nu_n \in V$，使得 V 有由形如 $T^j\nu_k$ 的向量构成的基，其中各k 从1 取至 n 而 j 从使得 $T^{m_k}\nu_k \neq 0$ 的最大非负整数 m_k 取至0． T 关于该基的矩阵就类似于上例中的矩阵．更具体地说， T 关于该基有分块对角矩阵，其中每个块除了紧挨在对角线正上方的元素，其余各元素都等于0
在下面定义的各 A_k 中，对角线上的元素都是 T 的特征值 λ_k，紧挨在对角线正上方的元素都是 1，其余各元素都是 0（要理解为何每个 λ_k 都是 T 的特征值，见 5.41)．各 λ_k 不一定互异；同时， A_k 也可能是仅包含 T 的特征值的 1×1 矩阵 (λ_k))．若各 λ_k 都是0，那么下面定义就与上段所述幂零算子的性质相合（回忆一下，若 T 是幂零的，那么0 是 T 唯一的特征值）

8.44 若当基

$T \in \mathcal{L}(V)$．称 V 的一个基是若当基，如果 T 关于该基具有分块对角矩阵

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

其中各 A_k 是形如

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}$$

的上三角阵

有限维复向量空间上每个算子都有若当基

8.45 每个幂零算子都有若当基

$T \in \mathcal{L}(V)$ 是幂零的，则V 中有一个基是 T 的若当基

证

对 $\dim V$ 用归纳法
$\dim V = 1$ 时欲证结论成立（因为在此情形下，唯一的幂零算子是 0 算子）．设 $\dim V > 1$ 且欲证结论对于所有维数更小的向量空间都成立
令 m 是使得 $T^m = 0$ 的最小正整数，于是存在 $u \in V$ 使得 $T^{m-1}u \neq 0$．令

$$U = \operatorname{span}(u, Tu, \ldots, T^{m-1}u)$$

$u, Tu, \ldots, T^{m-1}u$ 是线性无关的（见8A 节习题2）．若 $U = V$，那么将该组反过来写，就得到了T 的一个若当基，证明完成．于是我们可设 $U \neq V$
注意到 U 在 T 下不变．由归纳假设知 U 有一个基是 $T|_U$ 的若当基．我们的证明策略是，找到 V 的子空间 W，它同样在 T 下不变且满足 $V = U \oplus W$．再利用归纳假设可知，W 有一个基是 $T|_W$ 的若当基．将 $T|_U$ 和 $T|_W$ 的若当基合并就得到了 T 的若当基
令 $\varphi \in V'$ 满足 $\varphi(T^{m-1}u) \neq 0$．令
那么 W 是 V 的在 T 下不变的子空间【这是因为如果 $\nu \in W$，那么对 $k = 0, \ldots, m-1$ 有$\varphi\left(T^k(T\nu)\right) = 0$，其中 $k = m-1$ 的情形成立是由于 $T^m = 0$】．下面我们证明 $V = U \oplus W$
为了说明 $U + W$ 是直和，设 $\nu \in U \cap W$ 且 $\nu \neq 0$．因为 $\nu \in U$，所以存在 $c_0, \ldots, c_{m-1} \in \mathbf{F}$ 使

$$\nu = c_0 u + c_1 Tu + \cdots + c_{m-1} T^{m-1} u$$

令 j 是使得 $c_j \neq 0$ 的最小下标．将 T^{m-j-1} 作用于上式两侧得

$$T^{m-j-1}\nu = c_j T^{m-1} u$$

这里利用了等式 $T^m = 0$
现将 φ 作用于上式两端得

$$\varphi(T^{m-j-1}\nu) = c_j \varphi(T^{m-1}u) \neq 0$$

上式表明 $\nu \notin W$．因此我们证明了 $U \cap W = \{0\}$，这就表明 $U + W$ 是直和（见 1.46）
为了证明 $U \oplus W = V$，定义 $S : V \to \mathbf{F}^m$ 为

$$S\nu = \left(\varphi(\nu), \varphi(T\nu), \ldots, \varphi(T^{m-1}\nu)\right)$$

于是 null $S = W$，故

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m$$

其中第二个等号源于线性映射基本定理（3.21）．利用上述不等式，可得

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

从而 $U \oplus W = V$ （由2.39），这就完成了证明

现在，利用广义特征空间分解可将上述结论推广至非幂零的算子．该推广只有在复向量空间上才可进行

8.46 若当型

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$．那么 V 有一个基是 T 的若当基

证

令 $\lambda_1, \ldots, \lambda_m$ 是 T 的所有互异特征值．由广义特征空间分解

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

且各 $(T - \lambda_k I)|_{G(\lambda_k, T)}$ 都是幂零的（见 8.22）．于是 8.45 就表明，每个 $G(\lambda_k, T)$ 中都有某基是$(T - \lambda_k I)|_{G(\lambda_k, T)}$ 的若当基．将这些基合并，即可得到 V 的一个基，且它是 T 的若当基

8D 迹

从定义方阵的迹开启本节，先讨论性质，再利用方阵的迹定义算子的迹

8.47 矩阵的迹

各元素均属于$\mathbf{F}$的方阵的迹:=主对角线元素和

8.48 迹

$$A=\begin{pmatrix}3&-1&-2\\3&2&-3\\1&2&0\end{pmatrix}$$

tr $A=3+2+0=5$

矩阵乘法不满足交换律．但是，下面结论表明，交换乘积项的顺序不影响乘积的迹

8.49 AB 的迹等于 BA 的迹

A 是 $m\times n$ 矩阵， B 是 $n\times m$ 矩阵，则

$$\operatorname{tr}(AB)=\operatorname{tr}(BA)$$

证

设

$$A=\begin{pmatrix}A_{1,1}&\cdots&A_{1,n}\\\vdots&&\vdots\\A_{m,1}&\cdots&A_{m,n}\end{pmatrix},\quad B=\begin{pmatrix}B_{1,1}&\cdots&B_{1,m}\\\vdots&&\vdots\\B_{n,1}&\cdots&B_{n,m}\end{pmatrix}$$

$m\times m$ 矩阵 AB 对角线上的第 j 个元素等于 $\sum_{k=1}^{n}A_{j,k}B_{k,j}$，故

$$\begin{aligned}\operatorname{tr}(AB)&=\sum_{j=1}^{m}\sum_{k=1}^{n}A_{j,k}B_{k,j}\\&=\sum_{k=1}^{n}\sum_{j=1}^{m}B_{k,j}A_{j,k}\\&=\operatorname{tr}(BA)\end{aligned}$$

想定义算子 $T\in\mathcal{L}(V)$ 的迹为 T 关于 V 的某个基的矩阵的迹，然而这定义不应依赖于基的选取。下面的结论说明这是可做到的

8.50 算子的矩阵的迹不依赖于基的选取

$T\in\mathcal{L}(V)$．设 $u_1,\ldots,u_n$ 和 $\nu_1,\ldots,\nu_n$ 是 V 的基．那么

$$\operatorname{tr}M\big(T,(u_1,\ldots,u_n)\big)=\operatorname{tr}M\big(T,(\nu_1,\ldots,\nu_n)\big)$$

证

令 $A=\mathcal{M}\big(T,(u_1,\ldots,u_n)\big)$， $B=\mathcal{M}\big(T,(\nu_1,\ldots,\nu_n)\big)$．换基公式告诉我们，存在可逆的 $n\times n$ 矩阵 c 使得 $A=C^{-1}BC$ （见3.84)．于是

$$\begin{aligned}\operatorname{tr}A&=\operatorname{tr}\big((C^{-1}B)C\big)\\&=\operatorname{tr}\big(C(C^{-1}B)\big)\\&=\operatorname{tr}\big((CC^{-1})B\big)\\&=\operatorname{tr}B\end{aligned}$$

其中第二行源于8.49

因为8.50，所以下面定义是合理的

8.51 算子的迹

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

$$\operatorname{tr} T = \operatorname{tr}\ M\left(T, (\nu_1, \ldots, \nu_n)\right)$$

其中 $\nu_1, \ldots, \nu_n$ 是 V 的任一基

设 $T \in \mathcal{L}(V)$ 且 λ 是 T 的一个特征值．回忆 λ 的重数为广义特征空间 $G(\lambda, T)$ 的维数（见8.23）
我们证明了该重数等于dim null $(T-\lambda I)^{\dim V}$ （见8.20）
此外，若 V 是复向量空间，那么 T 的所有特征值的重数之和等于 $\dim V$ （见 8.25）
下面结论中“各特征值出现次数等于其重数”的特征值之和意为如果 $\lambda_1, \ldots, \lambda_m$ 是T 的所有互异特征值，且重数分别为 $d_1, \ldots, d_m$ ，那么求和式为

$$d_1\lambda_1 + \cdots + d_m\lambda_m$$

如果你更喜欢将所有特征值按一组未必互异的数（其中各特征值出现的次数等于其重数）来处理，那么可将所有特征值记为 $\lambda_1, \ldots, \lambda_n$ （其中 n 等于 $\dim V$ ），这样求和式可写为

$$\lambda_1 + \cdots + \lambda_n$$

8.52 复向量空间迹=特征值之和

设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ ．那么 $\operatorname{tr} T$ 等于 T 的特征值之和，其中各特征值出现次数等于其重数

证

存在 V 的一个基使得 T 关于该基有上三角矩阵，且对角线上各元素都是 T 的特征值，并且各特征值出现次数等于其重数（见8.37）．于是由算子的迹的定义以及8.50 （该结论让我们得以自由选取基），可得tr$T$ 等于 T 的特征值之和，其中各特征值出现次数等于其重数

8.53 $\mathbf{C}^3$ 上一个算子的迹

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

$$T(z_1, z_2, z_3) = (3z_1 - z_2 - 2z_3, 3z_1 + 2z_2 - 3z_3, z_1 + 2z_2)$$

关于 $\mathbf{C}^3$ 的标准基的矩阵

$$\begin{pmatrix} 3 & -1 & -2 \\ 3 & 2 & -3 \\ 1 & 2 & 0 \end{pmatrix}$$

求该矩阵对角线上各元素的和，我们可见tr $T = 5$
T 的特征值是 $1, 2+3\mathrm{i}, 2-3\mathrm{i}$ ，它们的重数都是1（你可自行验证）．这些特征值之和（其中各特征值出现次数等于其重数）是 $1 + (2+3\mathrm{i}) + (2-3\mathrm{i})$ ，等于5，正如 8.52 所言
迹和特征多项式具有紧密的联系．假设 $\mathbf{F} = \mathbf{C}$ 且 $T \in \mathcal{L}(V)$ ，且 $\lambda_1, \ldots, \lambda_n$ 是 T 的特征值（各特征值出现的次数等于其重数）．那么由定义（见8.26）， T 的特征多项式等于

$$(z-\lambda_1)\cdots(z-\lambda_n)$$

将上述多项式展开，可将 T 的特征多项式写成

$$z^n - (\lambda_1 + \cdots + \lambda_n)z^{n-1} + \cdots + (-1)^n(\lambda_1 \cdots \lambda_n)$$

由上述表达式立刻可得出下面结论．该结论也见于9.65，但9.65 无需假设 $\mathbf{F} = \mathbf{C}$

8.54 迹与特征多项式

设 $\mathbf{F}=\mathbf{C}$ 且 $T\in\mathcal{L}(V)$ ．令 $n=\dim V$ ．那么 $\mathrm{tr}T$ 等于 T 的特征多项式中 z^{n-1} 项的系数的相反数

下面结论给出了内积空间上算子的迹的一个很漂亮的公式

8.55 内积空间上的迹

设 V 是内积空间， $T\in\mathcal{L}(V)$ ， $e_1,\dots,e_n$ 是 V 的规范正交基．则

$$\mathrm{tr}T=\langle Te_1,e_1\rangle+\cdots+\langle Te_n,e_n\rangle$$

证

注意到 $\mathcal{M}(T,(e_1,\dots,e_n))$ 第 k 行第 k 列中的元素等于 $\langle Te_k,e_k\rangle$ 【利用 6.30 (a)，取 $\nu=Te_k$ 】，即可得欲证的公式

方阵的迹的代数性质可迁移至算子的迹，如下结论所示

8.56 迹是线性的

函数 $\mathrm{tr}:\mathcal{L}(V)\to\mathbf{F}$ 是 $\mathcal{L}(V)$ 上的线性泛函，且使

$$\mathrm{tr}(ST)=\mathrm{tr}(TS)$$

对所有 $S,T\in\mathcal{L}(V)$ 都成立

证

取定 V 的一个基．本证明中所有算子的矩阵都是关于该基的．设 $S,T\in\mathcal{L}(V)$
若 $\boldsymbol{\lambda}\in\mathbf{F}$ ，那么

$$\mathrm{tr}(\lambda T)=\mathrm{tr}\mathcal{M}(\lambda T)=\mathrm{tr}\big(\lambda\mathcal{M}(T)\big)=\lambda\mathrm{tr}\mathcal{M}(T)=\lambda\mathrm{tr}T,$$

其中第一个和最后一个等号源于算子的迹的定义，第二个等号来自3.38，而第三个等号由方阵的迹的定义可得
同时

$$\mathrm{tr}(S+T)=\mathrm{tr}\mathcal{M}(S+T)=\mathrm{tr}\big(\mathcal{M}(S)+\mathcal{M}(T)\big)=\mathrm{tr}\mathcal{M}(S)+\mathrm{tr}\mathcal{M}(T)=\mathrm{tr}S+\mathrm{tr}T$$

其中第一个和最后一个等号源于算子的迹的定义，第二个等号来自3.35，而第三个等号由方阵的迹的定义可得．上述两段说明了 $\mathrm{tr}:\mathcal{L}(V)\to\mathbf{F}$ 是 $\mathcal{L}(V)$ 上的线性泛函
此外

$$\mathrm{tr}(ST)=\mathrm{tr}\mathcal{M}(ST)=\mathrm{tr}\big(\mathcal{M}(S)\mathcal{M}(T)\big)=\mathrm{tr}\big(\mathcal{M}(T)\mathcal{M}(S)\big)=\mathrm{tr}\mathcal{M}(TS)=\mathrm{tr}(TS)$$

其中第二个和第四个等号源于3.43，关键的第三个等号则来自8.49

等式 $\mathrm{tr}(ST)=\mathrm{tr}(TS)$ 和 $\mathrm{tr}\,I=\dim V$ 在 $\mathcal{L}(V)$ 上的线性泛函中唯一刻画了迹— 见习题10
等式 $\mathrm{tr}(ST)=\mathrm{tr}(TS)$ 引出了下面结论，该结论在无限维向量空间上不成立（见本节习题13）．然而，对 s,T 和 V 附加限制条件，即
下面结论的表述中并不涉及迹，但是这个简短的证明用到了迹．数学中，发生这种情况时，往往是有个好的定义做好了铺垫
可将下面结论推广至无限维向量空间，推广后的结论在量子理论中有重要应用

8.57 恒等算子不等于 ST 与 TS 之差

不存在使得 $ST-TS=I$ 成立的算子 $S,T\in\mathcal{L}(V)$

证

设 $S, T \in \mathcal{L}(V)$ ．那么

$$\operatorname{tr}(ST - TS) = \operatorname{tr}(ST) - \operatorname{tr}(TS) = 0,$$

其中两个等号都源自 8.56．而 I 的迹等于 $\dim V$ ，肯定不是 0．因为 $ST - TS$ 和 I 有不同的迹，所以它们不可能相等