可对角化算子

1.对角矩阵

1.1 定义:对角矩阵和可对角化

对角矩阵是除对角线外其余元素均为0的矩阵.

若$T \in \mathcal{L}(V)$关于V的某个基具有对角矩阵,那么称T是**可对角化**的.

不难证明,如果$T \in \mathcal{L}(V)$关于V的一组基$v_1, \cdots, v_m$的矩阵是对角矩阵,那么对角线上的元素λ_k均为T的特征值,对应的特征向量就是v_k. 假定T具有特征值λ,那么自然地可以想到对应的特征向量构成一个向量空间.我们为此定义特征空间.

1.2 定义:特征空间

设$T \in \mathcal{L}(V)$且$\lambda \in \mathbb{F}$.T对应于λ的一组基**特征空间**$E(\lambda, T)$是定义如下的V的子空间.

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

因此,$E(\lambda, T)$是T对应λ的所有特征向量和$\mathbf{0}$的子空间.

由定义可知,λ是T的特征值当且仅当$E(\lambda, T) \neq \{\mathbf{0}\}$;亦可知$T|_{E(\lambda,T)}$就是将向量乘以$\lambda$的算子.

我们知道对应于不同特征值的特征向量是线性无关的.于是特征空间的和也应当是直和.

1.3 特征空间之和是直和

设$T \in \mathcal{L}(V)$且$\lambda_1, \cdots, \lambda_m$是$T$的互异特征值.那么

$$E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

是直和.此外若V是有限维的,那么

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leqslant \dim V$$

证明是简单的,在此不再赘述.

2.可对角化的条件

2.1 可对角化的等价条件

设V是有限维的,$T \in \mathcal{L}(V)$.令$\lambda_1, \cdots, \lambda_m$表示$T$的互异特征值,那么下面各命题等价.

(a) T是可对角化的.

(b) V有T的特征向量构成的基.

(c) $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$.

(d) $\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) = \dim V$.

Proof.

T关于V的某个基$v_1, \cdots, v_n$具有对角矩阵当且仅当对任意$1 \leqslant k \leqslant n$有$Tv_k = \lambda_k v_k$.于是**(a)**和**(b)**等价.

假定**(b)**成立,于是V中每个向量都可以表示为T的特征向量的线性组合,即

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$$

据**1.3**可知这与$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$等价,于是**(c)**成立.

假定**(c)**成立,于是

$$\dim V = \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$$

于是**(d)**成立.

假定**(d)**成立,为每个$E(\lambda_k, T)$选取一个基后组合,于是这构成了V中长度为$\dim V$的向量组$v_1, \cdots, v_m$.

又各v_k对应的特征值不同,于是它们线性无关.于是**(b)**成立.

下面的结论给出了算子特征值数目和能否对角化的关系.

2.2 特征值数目与对角化

设V是有限维的且$T \in \mathcal{L}(V)$有$\dim V$个互异特征值,那么T是可对角化的.

我们可以用**2.1**方便地证明之.需要说明的是上面只给出了能对角化的充分条件,而有些时候即使T的互异特征值数目少于$\dim V$,它仍可能是对角化的.事实上这是由于$E(\lambda, T)$的维数可能高于一.

我们可以用矩阵对角化方便的计算矩阵的高次幂.例如下面的例子.

2.3 矩阵高次幂的计算

定义$T \in \mathcal{L}(\mathbb{F}^3) : (x, y, z) \mapsto (2x + y, 5y + 3z, 8z)$.求$\mathcal{M}(T^{100})$(基向量选取标准基).

Solution.

T关于$\mathbb{F}^3$的标准基e_1,e_2,e_3的矩阵为

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

这是一个上三角矩阵,于是T具有特征值$2,5,8$.因而,T关于$\mathbb{F}^3$的某个基具有对角矩阵

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

为了求解这组基,令$T(x,y,z)=\lambda(x,y,z)$,然后分别令$\lambda=2,5,8$即可.解得这组基为

$$v_1=(1,0,0),v_2=(1,3,0),v_3=(1,6,6)$$

熟知对任意$k\in\mathbb{N}$有$T^k v_j=\lambda_j^k v_j$.

下面我们以$T^{100}e_3$说明如何计算T^{100}中的元素.易知$e_3=\dfrac{1}{6}v_1-\dfrac{1}{3}v_2+\dfrac{1}{6}v_3$.于是

$$\begin{aligned}
T^{100}e_3 &= T^{100}\left(\frac{1}{6}v_1-\frac{1}{3}v_2+\frac{1}{6}v_3\right)\\
&= \frac{1}{6}T^{100}v_1-\frac{1}{3}T^{100}v_2+\frac{1}{6}T^{100}v_3\\
&= \frac{1}{6}\cdot 2^{100}v_1-\frac{1}{3}\cdot 5^{100}v_2+\frac{1}{6}\cdot 8^{100}v_3\\
&= \left(\frac{2^{100}-2\cdot 5^{100}+8^{100}}{6},8^{100}-5^{100},8^{100}\right)
\end{aligned}$$

我们将在之后的习题中用类似的方法推知斐波那契数列的通项公式.

现在,我们把目光回到最小多项式上.之前我们已经说过T是可对角化的当且仅当T的最小多项式是若干一次项的乘积.这个结论对复向量空间的算子总是成立.

现在我们要说明如何通过相似的方式得出T是否可对角化.

2.4 可对角化的充要条件

设V是有限维的,且$T\in\mathcal{L}(V)$,那么T是可对角化的当且仅当T的最小多项式等于$(z-\lambda_1)\cdots(z-\lambda_m)$,其中$\lambda_1,\cdots,\lambda_m$互异.也即,$T$的最小多项式的根的数目等于其次数而没有重根.

Proof.

$\Rightarrow$:设T是可对角化的,那么存在由T的特征向量$v_1,\cdots,v_n$构成的V的基.令$\lambda_1,\cdots,\lambda_m$是$V$的互异特征值.

那么对于任意$1 \leqslant j \leqslant n$,存在$1 \leqslant k \leqslant m$使得$Tv_j = \lambda_k v_j$,即$(T - \lambda_k I)v_j = \mathbf{0}$.于是

$$(T - \lambda_1 I) \cdots (T - \lambda_m I)v_j = \mathbf{0}$$

于是T的最小多项式为$(z - \lambda_1) \cdots (z - \lambda_m)$.

$\Leftarrow$:设T的最小多项式为$(z - \lambda_1) \cdots (z - \lambda_m)$,其中$\lambda_1, \cdots, \lambda_m \in \mathbb{F}$互异.于是

$$(T - \lambda_1 I) \cdots (T - \lambda_m I) = \mathbf{0}$$

我们将对m用归纳法以证明T是可对角化的.

$m = 1$时,$T - \lambda_1 I = \mathbf{0}$,这表明$T$是恒等算子的标量倍,于是$T$是可对角化的.

现在设$m > 1$且欲证结论在m更小时的所有情况均成立.记$U = \text{range } T - \lambda_m I$.

我们知道U在T下是不变的,于是$T|_U$是U的算子.对任意$u \in U$,总存在$v \in V$使得$(T - \lambda_m I)v = u$.由前面的定义可知

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = \mathbf{0}v = \mathbf{0}$$

于是$(z - \lambda_1) \cdots (z - \lambda_{m-1})$是$T|_U$的最小多项式的多项式倍.

由归纳假设,$T|_U$的对角矩阵存在,设对应的基为$u_1, \cdots, u_M$,各u_k均为T的特征向量.

设$u \in \text{range } (T - \lambda_m I) \cap \text{null } (T - \lambda_m I)$,那么$(T - \lambda_m I)u = \mathbf{0}$,即$Tu = \lambda_m u$.

于是

$$\mathbf{0} = (T - \lambda_1) \cdots (T - \lambda_{m-1} I)u = (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})u$$

由于$\lambda_1, \cdots, \lambda_m$互异,于是上式成立当且仅当$u = \mathbf{0}$.

于是$U + \text{null } (T - \lambda_m I)$是直和,且据线性映射基本定理有$\dim U + \dim \text{null } (T - \lambda_m I) = \dim V$.

因此$U \oplus \text{null } (T - \lambda_m I) = V$.而$\text{null } (T - \lambda_m I)$中的所有向量都是对应于$\lambda_m$的特征向量,取其基为$v_1, \cdots, v_N$.

于是$u_1, \cdots, u_M, v_1, \cdots, v_N$是$V$的一组基,且各基向量均为$T$的特征向量.$T$关于这基的矩阵即为对角矩阵.

我们可以将可对角化算子限制于不变子空间上.

2.5 作用于不变子空间上的可对角化算子

设$T \in \mathcal{L}(V)$可对角化,U是V在T下的不变子空间,那么$T|_U$是U上的可对角化算子.

Proof.

由**2.4**,T的最小多项式应具有$(z - \lambda_1) \cdots (z - \lambda_m)$的形式,其中$\lambda_1, \cdots, \lambda_m$互异.该多项式是$T|_U$的最小多项式的多项式倍,因而$T|_U$的最小多项式也具有类似的形式,因而$T|_U$是可对角化的.

3.格什戈林圆盘定理

3.1 定义:格什戈林圆盘

设$T \in \mathcal{L}(V)$且$v_1, \cdots, v_n$是V的一个基.令$A = \mathcal{M}(T, (v_1, \cdots, v_n))$.$T$关于该基的**格什戈林圆盘**是形如

$$\left\{ z \in \mathbb{F} : |z - A_{j,j}| \leqslant \sum_{k \in \{1, \cdots, n\} \setminus \{j\}} |A_{j,k}| \right\}$$

的集合,其中$1 \leqslant j \leqslant n$.

因为j有n种取值,所以T相对应的也有n个格什戈林圆盘.

当$\mathbb{F} = \mathbb{C}$时,每个圆盘是$\mathbb{C}$中以$A_{j,j}$为圆心的圆盘,半径为这行上除了$A_{j,j}$的所有元素的绝对值之和.

当$\mathbb{F} = \mathbb{R}$是,每个圆盘是$\mathbb{R}$中以$A_{j,j}$为中心的闭区间,闭区间的半径即上面提到的半径.

特别地,当非对角线元素均为0时,每个圆盘都退化为一个点,这点也是T的特征值(这是由于对角矩阵的性质).而格什戈林圆盘定理告诉我们,T的每个特征值都包含于T的格什戈林圆盘中.

3.2 格什戈林圆盘定理

设$T \in \mathcal{L}(V)$且$v_1, \cdots, v_n$是V的一个基,那么T的每个特征值都包含于T关于$v_1, \cdots, v_n$的某个格什戈林圆盘中.

Proof.

设$\lambda \in \mathbb{F}$是T的一个特征值,令w是与之对应的特征向量.于是存在$a_1, \cdots, a_n \in \mathbb{F}$使得$w = a_1 v_1 + \cdots + a_n v_n$.

令$A = \mathcal{M}(T, (v_1, \cdots, v_n))$,将$T$作用于上式两侧有

$$\lambda w = \sum_{k=1}^{n} a_k T v_k = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_k A_{j,k} \right) v_j$$

于是我们有$\lambda a_j = \sum\limits_{k=1}^{n} a_k A_{j,k}$对任意$1 \leqslant j \leqslant n$成立.取$j \in \{1, \cdots, n\}$使得$|c_j| = \max\limits_{1 \leqslant i \leqslant n} |c_i|$.变形可得

$$|\lambda - A_{j,j}| = \left| \sum_{k \in \{1, \cdots, n\} \setminus \{j\}} A_{j,k} \frac{a_k}{a_j} \right| \leqslant \sum_{k \in \{1, \cdots, n\} \setminus \{j\}} |A_{j,k}|$$

于是λ在关于$v_1, \cdots, v_n$的第j个格什戈林圆盘中.