不变子空间

1.特征值

我们从算子开始本章的叙述.

1.1 定义:算子

称从一个向量空间到其本身的线性映射为**算子**.

当一个空间V较为复杂的时候,研究它的算子T时我们可能会想到将它分解为$V = V_1 \oplus \cdots \oplus V_m$,然后研究各$T|_{V_k}$.这通常会使得问题简单些.然而有时候我们会遇到一个问题:T可能并不是V_k上的算子,它可以将V_k中的元素映到其它子空间中.因此我们在研究问题时,总是选取这样的子空间,使得T也是这些子空间上的算子.我们做如下定义.

1.2 定义:不变子空间

设U是V的子空间.对于$T \in \mathcal{L}(V)$,若对任意$u \in U$均有$Tu \in U$,则称U在T下是**不变的**.

例如,对于$T \in \mathcal{L}$,可以验证$\text{null } T, \text{range } T, \{\mathbf{0}\}$和$V$本身均为在$T$下的不变子空间.
现在,我们先来研究最简单的非平凡不变子空间,即维数为1的不变子空间.设$v \neq \mathbf{0} \in V$,令$U = \text{span}(v)$.
那么U即V的一维子空间.我们假定U在T下不变,那么对于任意$u \in U$都有$Tu \in U$,即存在$\lambda \in F$使得$Tu = \lambda u$.
在下面的定义中,我们将反复地运用$Tv = \lambda v$这一条件.

1.3 定义:特征值

设$T \in \mathcal{L}(V)$.对于标量$\lambda \in \mathbb{F}$,若存在$v \neq \mathbf{0} \in V$使得$Tv = \lambda v$,则称λ为T的**特征值**.

下面的定理揭示了存在特征值的等价条件.

1.4 成为特征值的等价条件

设V是有限维的,$T \in \mathcal{L}(V)$且$\lambda \in \mathbb{F}$,那么如下几个命题等价.

(a) λ是T的特征值.

(b) $T - \lambda I$不是单射.

(c) $T - \lambda I$不是满射.

(d) $T - \lambda I$不可逆.

Proof.

(a)与**(b)**等价是因为$Tv=\lambda v \Leftrightarrow (T-\lambda I)(v)=\mathbf{0}$.

后面三者的等价性在线性映射的可逆一节已经说明.

1.5 定义:特征向量

设$T\in\mathcal{L}(V)$且$\lambda\in\mathbb{F}$是T的特征值.若向量$v\in V$满足$v\neq\mathbf{0}$且$Tv=\lambda v$,那么称v是T对应于λ的**特征向量**.

我们给出一个例子.设$T\in\mathcal{L}(\mathbb{F}^2):(w,z)\mapsto(-z,w)$.

若$\mathbb{F}=\mathbb{R}$,那么T有特征值$\lambda\in\mathbb{R}$即存在$w,z\in\mathbb{R}$使得

$$-z=\lambda w, w=\lambda z$$

即$\lambda^2=-1$.这方程在$\mathbb{R}$上无解,也即T没有特征值和特征向量.

从几何上理解,T即将向量(w,z)逆时针旋转$90°$.没有任何一个非零向量在这样的旋转操作后还能与原来共线的,于是T也就没有特征向量和特征值.

若$\mathbb{F}=\mathbb{C}$,那么同理可知$\lambda=\mathrm{i}$或$-\mathrm{i}$.

于是T有特征值i和$-\mathrm{i}$,对应的特征向量分别为$(w,-w\mathrm{i})$和$(w,w\mathrm{i})$,其中$w\neq0$.

上面的例子告诉我们,在本章的研究中数域$\mathbb{F}$的选取是重要的.

1.6 线性无关的特征向量

设$T\in\mathcal{L}(V)$,那么分别对应于T的不同特征值的特征向量构成的每个组都线性无关.

Proof.

假定欲证明的结论错误,那么存在最小的$m\in\mathbb{N}^*$使得T对应于其互异的特征值$\lambda_1,\cdots,\lambda_m$的特征向量$v_1,\cdots,v_m$线性相关.于是存在$a_1,\cdots,a_m\in\mathbb{F}$,其中各数均非零(根据$m$的最小性),使得

$$a_1v_1+\cdots+a_mv_m=\mathbf{0}$$

将$T-\lambda_mI$作用于上式的两端可知

$$a_1\left(\lambda-\lambda_m\right)v_1+\cdots+a_{m-1}\left(\lambda_{m-1}-\lambda_m\right)=\mathbf{0}$$

由于$a_1,\cdots,a_{m-1}$均非零,而$\lambda_1,\cdots,\lambda_m$各不相同,于是上式中各$v_k$的系数均非零.

这表明$v_1,\cdots,v_{m-1}$线性相关,而这与线性相关组最小长度为m不符.于是命题得证.

我们可以根据上面的证明得出下面的结论.

1.7 算子的特征值数目

设V是有限维的,那么V上的算子至多有$\dim V$个互异特征值.

这是容易说明的,因为互异的特征值对应了相同数目的线性无关向量,这向量组的长度必然不大于V的维数.

2.将多项式作用于算子

算子的理论比一般的线性映射更丰富的原因在于算子可以自乘为幂.本小节将给出算子的幂和将多项式作用于算子的定义.

2.1 定义:算子的幂

设$T \in \mathcal{L}(V)$,$m \in \mathbb{N}^*$.

(a) 定义$T^m \in \mathcal{L}(V)$为$T^m = \underbrace{T \cdots T}_{m个T}$.

(b) 定义T^0为V上的恒等算子I.

(c) 若T是可逆的,记T的逆为T^{-1},定义$T^{-m} \in \mathcal{L}(V)$为$T^{-m} = (T^{-1})^m$.

于是线性映射的整数幂次有了定义.我们自然地想到多项式的定义,并把它们联系起来.

2.2 定义:算子的多项式

设$T \in \mathcal{L}(V)$且$p \in \mathcal{P}(\mathbb{F})$是由下式给出的多项式:对于任意$z \in \mathbb{F}$有

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

那么$p(T)$是V上的算子,定义为

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m$$

算子的多项式自然可以进行相应的加法和乘法.现在我们定义其乘积.

2.3 定义:多项式的乘积

对于$p, q \in\in \mathcal{P}(\mathbb{F})$,那么$p$与$q$的乘积$pq$是按$\forall z \in \mathbb{F}, (pq)(z) = p(z)q(z)$定义的多项式.

以及,自然地,我们有如下性质.

2.4 乘积的性质

设$p, q \in\in \mathcal{P}(\mathbb{F})$且$T \in \mathcal{L}(V)$.那么$(pq)(T) = p(T)q(T) = q(T)p(T)$.

证明这一点也是容易的,只需展开即可.

之前我们已经看到了不变子空间的例子,例如range T和null T均在T下不变.我们接下来将看到这两个空间在T的任意多项式下均不变.

2.5 $p(T)$**的值域和零空间在**T**下不变**

设$\mathcal{L}(V)$和$p \in \mathcal{P}(\mathbb{F})$,那么null $p(T)$和range $p(T)$在T下不变.

Proof.

设$u \in \text{null } p(T)$,那么$(p(T))(u) = \mathbf{0}$.于是

$$(p(T))(Tu) = (p(T)T)(u) = (Tp(T))(u) = T((p(T))(u)) = T(\mathbf{0}) = \mathbf{0}$$

从而$Tu \in \text{null } p(T)$,即null $p(T)$在T下不变.

设$u \in \text{range } p(T)$,于是存在$v \in V$使得$u = (p(T))(v)$.于是

$$Tu = T((p(T))(v)) = (p(T))(Tv)$$

因此$Tu \in \text{range } p(T)$,即range $p(T)$在T下不变.