arXiv:2601.18143v1 [math.RA] 26 Jan 2026

论矩阵的二维不变子空间

O. AL-RAISI 和 M. SHAHRYARI

摘要。我们提出一种统一方法，用于研究矩阵的二维不变子空间及其对应的超特征值。作为非交换代数中的一项新颖应用，我们展示了环 $\mathrm{Mat}_2(\mathbb{F})$ 上的矩阵的特征值与域 $\mathbb{F}$ 上的矩阵的二维不变子空间之间的联系。

1. 引言

关于各种矩阵对是否存在公共二维不变子空间的问题已得到详细研究；例如，在 [2],中，作者考察了两个旋转的公共二维不变子空间。在 [3] 和 [4], 中，也讨论了两个矩阵是否存在公共二维不变子空间（以及更高维不变子空间）的问题。本文提出一种统一方法来研究这一概念，以及矩阵的二维特征值（超特征值）这一概念。我们将引入一个四变量多项式，它在一定程度上发挥了特征多项式在二维不变子空间情形下的作用。随后，对于零特征或奇特征的域，我们将给出一个 2×2 矩阵成为矩阵 A 的超特征值的充要条件。我们将看到，这些结果与如下问题之间存在深刻联系：确定非交换环上自由模的线性映射的特征值，其中该环由域上元素构成的 2×2 矩阵组成。

这种联系可以描述如下：对于含幺交换环 R，研究元素属于 R 的矩阵的特征值，可以归结为求相应特征多项式的根。对于元素属于非交换环的矩阵，研究其特征值时则不能这样做。事实上，由于在这种情况下缺乏恰当的行列式定义，因此无法令人满意地定义元素属于非交换环的矩阵的特征多项式。我们

2020 数学主题分类。 主要分类 15A18；次要分类 15B33。关键词与短语。 非交换环上的矩阵、特征值、超特征值、不变子空间、特征多项式。

建立了这样一种联系：在环 $R = \mathrm{Mat}_2(\mathbb{F})$（由域 $\mathbb{F}$ 上所有 2×2 矩阵构成）上研究矩阵的特征值和特征向量，与研究元素属于 $\mathbb{F}$ 的矩阵的二维不变子空间之间存在联系，从而能够通过解决后一个问题来解决前一个问题。

2. 关于 R 的自由模

考虑 $\mathbb{F}$-空间 $\mathcal{U} = \mathbb{F}^{2n} = \mathbb{F}^n \oplus \mathbb{F}^n$，并定义 $R = \mathrm{Mat}_2(\mathbb{F})$ 在 $\mathcal{U}$ 上的作用如下：

$$\begin{bmatrix} p & q \\ r & s \end{bmatrix} \cdot \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} p\mathbf{u} + q\mathbf{v} \\ r\mathbf{u} + s\mathbf{v} \end{bmatrix}.$$

容易验证，$\mathcal{U}$ 是一个 R-模。注意，环 R 具有不变基数（IBN），因此自由 R-模的秩定义良好。

引理 2.1。设 $n = 2k$ 是一个偶数。那么 $\mathcal{U} \cong_R R^k$ 是秩为 k 的自由 R-模。

证明。设 $\{e_1, e_2, \ldots, e_n\}$ 是向量空间 $\mathbb{F}^n$ 的标准基。对于满足 $1 \leq i \leq k$ 的每个整数 i，定义

$$T_i = \begin{bmatrix} e_i \\ e_{k+i} \end{bmatrix}$$

并且我们证明 $\mathbf{B} = \{T_1, T_2, \ldots, T_k\}$ 是 $\mathcal{U}$ 的 R-基。设 $\begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix} \in \mathcal{U}$。我们有

$$\mathbf{u} = \sum_{i=1}^{k} (p_i e_i + q_i e_{k+i}),\ \mathbf{v} = \sum_{i=1}^{k} (r_i e_i + s_i e_{k+i}),$$

对于某些标量 $p_i, q_i, r_i, s_i \in \mathbb{F}$。这意味着

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \sum_{i=1}^{k} \begin{bmatrix} p_i & q_i \\ r_i & s_i \end{bmatrix} \cdot \begin{bmatrix} e_i \\ e_{k+i} \end{bmatrix} = \sum_{i=1}^{k} \begin{bmatrix} p_i & q_i \\ r_i & s_i \end{bmatrix} \cdot T_i$$

因此，$\mathbf{B}$ 生成 $\mathcal{U}$。为了证明 $\mathbf{B}$ 线性无关，设

$$\sum_{i=1}^{k} \begin{bmatrix} p_i & q_i \\ r_i & s_i \end{bmatrix} \cdot T_i = \mathbf{0}.$$

那么，我们有

$$\sum_{i=1}^{k} (p_i e_i + q_i e_{k+i}) = \mathbf{0},\ \sum_{i=1}^{k} (r_i e_i + s_i e_{k+i}) = \mathbf{0},$$

这意味着所有的 p_i, q_i, r_i 和 s_i 均为零。 □

对于任意矩阵 $A \in \mathbf{Mat}_n(\mathbb{F})$，定义一个映射 $f_A : \mathcal{U} \to \mathcal{U}$，使得

$$f_A(\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}) = \begin{bmatrix} A\mathbf{u} \\ A\mathbf{v} \end{bmatrix}.$$

引理 2.2。映射 f_A 是一个 R-同态。如果 $n = 2k$ 是偶数，那么每个 R-同态 $f : \mathcal{U} \to \mathcal{U}$ 都具有 $f = f_A$ 的形式，其中 $A \in \mathrm{Mat}_n(\mathbb{F})$。

证明。f_A 是一个 R-同态这一点可以直接验证。现在，设 $f : \mathcal{U} \to \mathcal{U}$ 是任意一个 R-同态。为简便起见，我们考虑 $n = 4$ 的情形。为此，设 e_1，e_2，e_3，e_4 是 $\mathbb{F}^4$ 的标准基。正如我们刚才看到的，$\mathcal{U}$ 的一个 R-基是 $\mathbf{B} = \{T_1, T_2\}$，其中

$$T_1 = \begin{bmatrix} e_1 \\ e_3 \end{bmatrix},\ T_2 = \begin{bmatrix} e_2 \\ e_4 \end{bmatrix}.$$

令$\mathbf{X}$ 为由下列向量组成的集合

$$E_1 = \begin{bmatrix} e_1 \\ \mathbf{0} \end{bmatrix}, E_2 = \begin{bmatrix} e_2 \\ \mathbf{0} \end{bmatrix}, E_3 = \begin{bmatrix} e_3 \\ \mathbf{0} \end{bmatrix}, E_4 = \begin{bmatrix} e_4 \\ \mathbf{0} \end{bmatrix},$$

and

$$F_1 = \begin{bmatrix} \mathbf{0} \\ e_1 \end{bmatrix}, F_2 = \begin{bmatrix} \mathbf{0} \\ e_2 \end{bmatrix}, F_3 = \begin{bmatrix} \mathbf{0} \\ e_3 \end{bmatrix}, F_4 = \begin{bmatrix} \mathbf{0} \\ e_4 \end{bmatrix}.$$

那么，$\mathbf{X}$ 是 $\mathcal{U}$ 的一个 $\mathbb{F}$-基。此外，我们有以下显然的关系

$$T_1 = E_1 + F_3,\ T_2 = E_2 + F_4,$$

以及以下关系

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot T_1,\ E_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot T_2,\ E_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot T_1,\ E_4 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot T_2,$$

and

$$F_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot T_1,\ F_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \cdot T_2,\ F_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \cdot T_1,\ F_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \cdot T_2.$$

现在，假设 $[f]_{\mathbf{B}}$ 是 f 关于 R-基$\mathbf{B}$ 的矩阵表示，则

$$\begin{bmatrix} \begin{bmatrix} p_{1,1} & q_{1,1} \\ r_{1,1} & s_{1,1} \end{bmatrix} & \begin{bmatrix} p_{1,2} & q_{1,2} \\ r_{1,2} & s_{1,2} \end{bmatrix} \\ \\ \begin{bmatrix} p_{2,1} & q_{2,1} \\ r_{2,1} & s_{2,1} \end{bmatrix} & \begin{bmatrix} p_{2,2} & q_{2,2} \\ r_{2,2} & s_{2,2} \end{bmatrix} \end{bmatrix}.$$

那么，我们有

$$\begin{aligned} f(E_1) &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot f(T_1) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \left(\begin{bmatrix} p_{1,1} & q_{1,1} \\ r_{1,1} & s_{1,1} \end{bmatrix} \cdot T_1 + \begin{bmatrix} p_{2,1} & q_{2,1} \\ r_{2,1} & s_{2,1} \end{bmatrix} \cdot T_2 \right) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \left(\begin{bmatrix} p_{1,1}e_1 + q_{1,1}e_3 \\ r_{1,1}e_1 + s_{1,1}e_3 \end{bmatrix} + \begin{bmatrix} p_{2,1}e_2 + q_{2,1}e_4 \\ r_{2,1}e_2 + s_{2,1}e_4 \end{bmatrix} \right) \\ &= \begin{bmatrix} p_{1,1}e_1 + p_{2,1}e_2 + q_{1,1}e_3 + q_{2,1}e_4 \\ \mathbf{0} \end{bmatrix} \\ &= p_{1,1}E_1 + p_{2,1}E_2 + q_{1,1}E_3 + q_{2,1}E_4. \end{aligned}$$

因此， $[f]_{\mathbf{x}}$ 的第一列是 $[p_{1,1}\ p_{2,1}\ q_{1,1}\ q_{2,1}\ 0\ 0\ 0\ 0]^t$。类似地，第二列是 $[p_{1,2}\ p_{2,2}\ q_{1,2}\ q_{2,2}\ 0\ 0\ 0\ 0]^t$。由此， f 作为一个 $\mathbb{F}$-映射的矩阵表示具有如下形式

$$\begin{bmatrix} p_{1,1} & p_{1,2} & r_{1,1} & r_{1,2} & 0 & 0 & 0 & 0 \\ p_{2,1} & p_{2,2} & r_{2,1} & r_{2,2} & 0 & 0 & 0 & 0 \\ q_{1,1} & q_{1,2} & s_{1,1} & s_{1,2} & 0 & 0 & 0 & 0 \\ q_{2,1} & q_{2,2} & s_{2,1} & s_{2,2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{1,1} & p_{1,2} & r_{1,1} & r_{1,2} \\ 0 & 0 & 0 & 0 & p_{2,1} & p_{2,2} & r_{2,1} & r_{2,2} \\ 0 & 0 & 0 & 0 & q_{1,1} & q_{1,2} & s_{1,1} & s_{1,2} \\ 0 & 0 & 0 & 0 & q_{2,1} & q_{2,2} & s_{2,1} & s_{2,2} \end{bmatrix}.$$

因此，若我们考虑矩阵

$$A = \begin{bmatrix} p_{1,1} & p_{1,2} & r_{1,1} & r_{1,2} \\ p_{2,1} & p_{2,2} & r_{2,1} & r_{2,2} \\ q_{1,1} & q_{1,2} & s_{1,1} & s_{1,2} \\ q_{2,1} & q_{2,2} & s_{2,1} & s_{2,2} \end{bmatrix},$$

则 $f = f_A$。 □

容易验证，对于给定的矩阵 $A \in \mathsf{Mat}_n(\mathbb{F})$，我们还有

$$[f_A]_{\mathbf{x}} = \begin{bmatrix} \begin{bmatrix} a_{1,1} & a_{k+1,1} \\ a_{1,k+1} & a_{k+1,k+1} \end{bmatrix} & \cdots & \begin{bmatrix} a_{1,k} & a_{k+1,k} \\ a_{1,n} & a_{k+1,n} \end{bmatrix} \\ \vdots & \vdots & \vdots \\ \begin{bmatrix} a_{k,1} & a_{n,1} \\ a_{k,k+1} & a_{n,k+1} \end{bmatrix} & \cdots & \begin{bmatrix} a_{k,k} & a_{n,k} \\ a_{k,n} & a_{n,n} \end{bmatrix} \end{bmatrix}.$$

˜此矩阵今后记作 A。下一节将利用环 $\alpha : \mathsf{Mat}_k(R) \to \mathsf{Mat}_n(\mathbb{F})$ 的自然同一性。这一

同一性会去掉任意 $k \times k$ 矩阵中的内层括号

$$M = \begin{bmatrix} \begin{bmatrix} a_{1,1}^{(1,1)} & a_{1,2}^{(1,1)} \\ a_{2,1}^{(1,1)} & a_{2,2}^{(1,1)} \end{bmatrix} & \cdots & \begin{bmatrix} a_{1,1}^{(1,k)} & a_{1,2}^{(1,k)} \\ a_{2,1}^{(1,k)} & a_{2,2}^{(1,k)} \end{bmatrix} \\ \vdots & \cdots & \vdots \\ \begin{bmatrix} a_{1,1}^{(k,1)} & a_{1,2}^{(k,1)} \\ a_{2,1}^{(k,1)} & a_{2,2}^{(k,1)} \end{bmatrix} & \cdots & \begin{bmatrix} a_{1,1}^{(k,k)} & a_{1,2}^{(k,k)} \\ a_{2,1}^{(k,k)} & a_{2,2}^{(k,k)} \end{bmatrix} \end{bmatrix}$$

并返回 $n \times n$ 矩阵

$$\alpha(M) = \begin{bmatrix} a_{1,1}^{(1,1)} & a_{1,2}^{(1,1)} & \cdots & a_{1,1}^{(1,k)} & a_{1,2}^{(1,k)} \\ a_{2,1}^{(1,1)} & a_{2,2}^{(1,1)} & \cdots & a_{2,1}^{(1,k)} & a_{2,2}^{(1,k)} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{1,1}^{(k,1)} & a_{1,2}^{(k,1)} & \cdots & a_{1,1}^{(k,k)} & a_{1,2}^{(k,k)} \\ a_{2,1}^{(k,1)} & a_{2,2}^{(k,1)} & \cdots & a_{2,1}^{(k,k)} & a_{2,2}^{(k,k)} \end{bmatrix}.$$

因此，对于矩阵 $A \in \mathsf{Mat}_n(\mathbb{F})$，我们将相应的˜ ˆ矩阵 $\alpha(A)$记为 A。

3. 超特征值

对于矩阵 $A \in \mathsf{Mat}_n(\mathbb{F})$，普通特征向量确定了 A的 1 维不变子空间。我们通过将 A 的二维不变子空间定义为超特征向量，来推广这一思想。若超特征向量 W 不包含 A的任何特征向量，则称其为适当的。设 $W = \langle \mathbf{u},\mathbf{v} \rangle$ 是 A的超特征向量。那么

$$A\mathbf{u} = p\mathbf{u} + q\mathbf{v},\ A\mathbf{v} = r\mathbf{u} + s\mathbf{v}$$

对于某些标量 p，q，r和 s。我们称矩阵 $\Lambda = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ 为与 W对应的 A 的超特征值。容易看出，如果 $W = \langle \mathbf{u}', \mathbf{v}' \rangle$并且

$$A\mathbf{u}' = p'\mathbf{u}' + q'\mathbf{v}',\ A\mathbf{v}' = r'\mathbf{u}' + s'\mathbf{v}',$$

那么这些矩阵

$$\Lambda = \begin{bmatrix} p & q \\ r & s \end{bmatrix}, \Lambda' = \begin{bmatrix} p' & q' \\ r' & s' \end{bmatrix}$$

是相似的。因此，A的每个超特征向量都对应于 2×2 矩阵的一个相似类。超特征值 $\Lambda = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ 称为真，当且仅当相应的超特征向量是真超特征向量。可以看出，若 Λ 是真超特征值，则对应于 Λ 的每个超特征向量都是真超特征向量。

现在，设 $n = 2k$ 为偶数。正如我们在上一节所见，每个 R-同态 $f\colon \mathcal{U} \to \mathcal{U}$ 都具有形式 f_A，其中 $A \in \mathsf{Mat}_n(\mathbb{F})$。换言之，每个 $n \times n$矩阵 A（其元素属于域 $\mathbb{F}$）都对应

˜一个唯一的 $k\times k$矩阵 A，其元素属于 R。因此，我们立即得到以下结论。

命题 3.1. 矩阵 $\Lambda \in R$ 是 A 的超特征值，当且仅当 Λ是 A的特征值。

因此，我们得到以下结论。

推论 3.2. 矩阵 $\Lambda = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ 是 A 的超特征值，当且仅当

$$\det(\hat{A} - \Lambda \cdot I_k) = 0,$$

其中， $\Lambda \cdot I_k$ 表示矩阵

$$\begin{bmatrix} p & q & 0 & 0 & \cdots & 0 & 0 \\ r & s & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & p & q & \cdots & 0 & 0 \\ 0 & 0 & r & s & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & p & q \\ 0 & 0 & 0 & 0 & \cdots & r & s \end{bmatrix}.$$

请注意，上述陈述仅在 n 为偶数时成立。对于一般情形，我们有以下结论。

命题 3.3. 设 $\Lambda = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$ 是 A 的一个超特征值。则有

$$\det((A - pI_n)(A - sI_n) - qrI_n) = 0.$$

证明。 设 $W = \langle \mathbf{u}, \mathbf{v} \rangle$ 是对应于 Λ 的一个超特征向量，使得

$$A\mathbf{u} = p\mathbf{u} + q\mathbf{v},\ A\mathbf{v} = r\mathbf{u} + s\mathbf{v}.$$

因此 $(A - pI_n)\mathbf{u} = q\mathbf{v}$，所以，

$$(A - sI_n)(A - pI_n)\mathbf{u} = q(A - sI_n)\mathbf{u} = qr\mathbf{u}.$$

因此，我们有

$$((A - pI_n)(A - sI_n) - qrI_n)\mathbf{u} = \mathbf{0};$$

由于 $\mathbf{u} \neq \mathbf{0}$，我们得到

$$\det((A - pI_n)(A - sI_n) - qrI_n) = 0.$$

□

注意，该命题的逆命题并不成立：设 p 是 A 的一个特征值，并假设 s 不是 A的特征值。那么显然

$(p,0,0,s)$ 满足方程 $\det((A-pI_n)(A-sI_n)-qrI_n)=0$，而矩阵 $\Lambda=\begin{bmatrix} p & 0 \\ 0 & s \end{bmatrix}$ 不是 A的超特征值。

下文中，$p_\Lambda(T)$ 表示 Λ 的特征多项式；也就是说，

$$p_\Lambda(T)=(T-p)(T-s)-qr=T^2-\mathrm{tr}(\Lambda)T+\det(\Lambda).$$

采用这一记号，上述命题表明，如果 Λ 是 A 的超特征值，那么 $\det(p_\Lambda(A))=0$。

定理 3.4. 设域 $\mathbb{F}$ 的特征不为 2。矩阵 $\Lambda=\begin{bmatrix} p & q \\ r & s \end{bmatrix}$ 是 A 的真超特征值，当且仅当 $\det(p_\Lambda(A))=0$ 且 $p_\Lambda(T)$ 不可约。

证明. 首先假设 Λ 是 A 的真超特征值。我们已经看到 $\det(p_\Lambda(A))=0$。假设 $p_\Lambda(T)$ 可约。那么 Λ 与某个三角矩阵相似，记为 $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$。令 $W=\langle \mathbf{u},\mathbf{v}\rangle$ 为对应于 Λ 的超特征向量。那么

$$A\mathbf{u}=p\mathbf{u}+q\mathbf{v},\ A\mathbf{v}=r\mathbf{u}+s\mathbf{v},$$

然而，还存在另外两个向量$\mathbf{u}'$，$\mathbf{v}'\in W$ 使得

$$A\mathbf{u}'=a\mathbf{u}'+b\mathbf{v}',\ A\mathbf{v}'=c\mathbf{v}'.$$

因此，$\mathbf{v}'$ 是 A的特征向量；这就产生了矛盾。

现在，假设 $\det(p_\Lambda(A))=0$ 且 $p_\Lambda(T)$不可约。我们证明存在一个与 Λ 对应的 A 的真超特征向量。由于 $p_\Lambda(A)$是奇异的，存在一个非零的$\mathbf{u}$ 使得 $p_\Lambda(A)\mathbf{u}=\mathbf{0}$，这意味着

$$(A-pI_n)(A-sI_n)\mathbf{u}=qr\mathbf{u}.$$

注意 $qr\neq 0$，否则 $p_\Lambda(T)=(T-p)(T-s)$，而它是可约的。令

$$\mathbf{v}=q^{-1}(A-pI_n)\mathbf{u}.$$

于是 $A\mathbf{u}=p\mathbf{u}+q\mathbf{v}$ 并且

$$\begin{aligned} A\mathbf{v} &= q^{-1}A(A-pI_n)\mathbf{u} \\ &= q^{-1}(A-sI_n+sI_n)(A-pI_n)\mathbf{u} \\ &= q^{-1}(qr\mathbf{u})+q^{-1}sq\mathbf{v} \\ &= r\mathbf{u}+s\mathbf{v}. \end{aligned}$$

为了证明 $W=\langle \mathbf{u},\mathbf{v}\rangle$ 是 Λ 的相应超特征向量，我们必须证明 $\mathbf{u}$ 和 $\mathbf{v}$ 构成一个线性无关组。反设

$\mathbf{v} = \lambda\mathbf{u}$，其中 Λ 是某个标量。那么 $A\mathbf{u} = p\mathbf{u} + q\mathbf{v} = (p + q\lambda)\mathbf{u}$，同理，$A\mathbf{v} = (p + q\lambda)\mathbf{v}$。因此，

$$r\mathbf{u} + s\mathbf{v} = (p + q\lambda)\mathbf{v},$$

这意味着

$$r + s\lambda = (p + q\lambda)\lambda.$$

因此，$q\lambda^2 + (p - s)\lambda - r = 0$，这意味着 λ 是二次方程 $qT^2 - (p - s)T - r = 0$ 的根。由于域的特征不为 2，判别式

$$(p - s)^2 + 4qr$$

是一个完全平方数。但这是 $p_\Lambda(T)$ 的判别式，这与 $p_\Lambda(T)$ 不可约的假设相矛盾。这表明$\mathbf{u}$ 和$\mathbf{v}$ 构成一个线性无关组，因此，W 是对应于 Λ 的超特征向量。接下来还需证明 W 是适当的。为此，设$\mathbf{v}' \in W$是 A 的一个特征向量，并设$\mathbf{u}' \in W$ 是任意满足 $W = \langle \mathbf{u}', \mathbf{v}' \rangle$的向量。那么，

$$A\mathbf{u}' = a\mathbf{u}' + b\mathbf{v}', \quad A\mathbf{v}' = c\mathbf{v}'$$

对于某些标量 $a, b,$ 和 c。因此，矩阵 $\Lambda' = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$ 也是 W 的一个对应超特征值。因此，Λ 和 Λ' 相似。这是不可能的，因为 Λ' 的特征多项式可约。□

我们以列出上述定理的一些推论来结束本节。再次假设域 $\mathbb{F}$ 的特征不为 2。

推论 3.5. 矩阵 Λ 是 A 的真超特征值，当且仅当其特征多项式 $p_\Lambda(T)$ 是 A 的特征多项式的一个不可约因子。

证明. 设 Λ 是与 $W = \langle \mathbf{u}, \mathbf{v} \rangle$ 对应的 A 的超特征值。将集合 $\{\mathbf{u}, \mathbf{v}\}$ 扩充为 $\mathbf{F}^n$ 的一个基，可知 A 相似于以下形式的矩阵

$$\begin{bmatrix} \Lambda^t & B \\ \mathbf{0} & C \end{bmatrix},$$

因此 $p_\Lambda(T)$ 是 A 的特征多项式 $p_A(T)$ 的一个因子。反过来，假设

$$p_A(T) = (T^2 + \lambda T + \mu)q(T),$$

其中二次多项式 $T^2 + \lambda T + \mu$ 不可约。由于 $p_A(A) = \mathbf{0}$，矩阵 $A^2 + \lambda A + \mu I_n$ 是奇异的，因此存在一个非零向量$\mathbf{u}$ 使得 $(A^2 + \lambda A + \mu I_n)\mathbf{u} = \mathbf{0}$。令$\mathbf{v} = A\mathbf{u} + \lambda\mathbf{u}$。于是我们有

$$A\mathbf{u} = -\lambda\mathbf{u} + \mathbf{v}, \; A\mathbf{v} = -\mu\mathbf{u}.$$

我们证明 $W = \langle \mathbf{u},\ \mathbf{v} \rangle$ 是与矩阵 $\Lambda = \begin{bmatrix} -\lambda & 1 \\ -\mu & 0 \end{bmatrix}$ 对应的 A 的真超特征向量。注意，$\dim(W) = 2$，否则必有$\mathbf{v} = \alpha\mathbf{u}$，其中某个标量为 α。但是，这样一来 $A\mathbf{u} = (\alpha - \lambda)\mathbf{u}$，这意味着

$$(\alpha - \lambda)^2 + \lambda(\alpha - \lambda) + \mu = 0$$

这意味着多项式 $T^2 + \lambda T + \mu$ 是可约的；这与假设矛盾。现在，$p_\Lambda(T) = T^2 + \lambda T + \mu$ 是不可约的，因此，根据前面的定理，Λ 是 A的真超特征值。□

因此，为了确定自由 R-模上的（真）特征值和 R-自同态，我们必须先确定相应的矩阵 A。然后，利用上述结果，我们可以求出 A的所有真超特征值。

推论 3.6. 设 $R =$ 矩阵$_2(\mathbb{F})$ 为特征不等于2的域 $\mathbb{F}$ 上所有 2×2 矩阵构成的环。设 $f : R^k \to R^k$ 为任意的 R-同态。假设 A 是 f 对应的矩阵，并考虑集合 $\{p_1(T) \ldots, p_m(T)\}$，它由特征多项式 $p_A(T)$ 的所有二次不可约因子构成。若对所有 i，均有 $p_i(T) = T^2 + \lambda_i T + \mu_i$，则这些矩阵

$$\Lambda_i = \begin{bmatrix} -\lambda_i & 1 \\ -\mu_i & 0 \end{bmatrix},\ (1 \leq i \leq m)$$

都是 f 的两两不相似真特征值。

R 同态 f 的两两不相似真特征值的个数至多为 k。证明如下。设 A 是 f 对应的矩阵，并假设对每个 $1 \leq i \leq m$，$W_i = \langle \mathbf{u}_i,\ \mathbf{v}_i \rangle$ 都是 A 的真超特征向量，对应的超特征值为 Λ_i。接下来证明和 $\sum_{i=1}^m W_i$ 是直和。众所周知，每个特征多项式 $p_{\Lambda_i}(T)$ 都是 $p_A(T)$ 的不可约因子。因此，我们有

$$p_A(T) = p_1(T)^{\alpha_1} \cdots p_m(T)^{\alpha_m} \times \text{other factors},$$

其中某些正整数为 α_1、.....、α_m。令 $V_i =$ ker$(p_i(A)^{\alpha_i})$。根据主分解定理，和 $\sum_{i=1}^m V_i$ 是直和。另一方面，直接检验可得

$$p_i(A)\mathbf{u}_i = p_i(A)\mathbf{v}_i = \mathbf{0},$$

因此 $W_i \subseteq V_i$。这表明和 $\sum_{i=1}^m W_i$ 也是直和，因此，A 只有有限多个两两不相似的真超特征值。

4. 一个例子

假设 $\mathbb{F} = \mathbb{R}$ 是实数域。令 $R = \mathrm{Mat}_2(\mathbb{R})$ 并且$\mathcal{U} = \mathbb{R}^4 \oplus \mathbb{R}^4$。映射

$$(x_1, x_2, x_3, x_4, y_1, y_2, y_3, y_4) \mapsto (-x_4, x_1, x_2, x_3, -y_4, y_1, y_2, y_3)$$

是一个 R-同态。相应的矩阵为

$$A = \begin{bmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

A 的特征多项式为 $T^4 + 1$，它没有根；然而，我们证明了 A 具有两个（不相似的）真超特征值

$$\begin{bmatrix} 0 & -1 \\ 1 & \sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & -\sqrt{2} \end{bmatrix}.$$

我们有

$$(A - pI_4)(A - sI_4) - qrI_4 = \begin{bmatrix} d & 0 & -1 & t \\ -t & d & 0 & -1 \\ 1 & -t & d & 0 \\ 0 & 1 & -t & d \end{bmatrix},$$

其中 $d = ps - qr$ 和 $t = p + s$。上述矩阵的行列式为$t^4 - 4dt^2 + (d^2 + 1)^2$。因此，为了求出超特征值，我们必须求解方程 $t^4 - 4dt^2 + (d^2 + 1)^2 = 0$。这是关于 t^2 的二次方程，其判别式为

$$16d^2 - 4(d^2 + 1)^2 = -4(d^2 - 1)^2.$$

我们必有 $-4(d^2 - 1)^2 \geq 0$，因此唯一的可能情况是 $d = \pm 1$。情形 $d = -1$ 会导出多项式 $t^4 + 4t^2 + 4 = 0$，该多项式没有根。因此，$d = 1$，从而 $t = \pm\sqrt{2}$。这意味着，对于任意真超特征值 Λ，都有 $\det(\Lambda) = 1$且 $\mathrm{tr}(\Lambda) = \pm\sqrt{2}$。因此，特征多项式为 $p_\Lambda(T) = T^2 \pm \sqrt{2}T + 1$，它确实是不可约的。因此，$\Lambda$ 与其中一个矩阵具有相同的特征多项式和最小多项式

$$\begin{bmatrix} 0 & -1 \\ 1 & \sqrt{2} \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & -\sqrt{2} \end{bmatrix}.$$

这意味着 Λ 与上述某个矩阵相似。现在，我们来确定相应的超特征向量。首先考虑 $\Lambda = \begin{bmatrix} 0 & -1 \\ 1 & \sqrt{2} \end{bmatrix}$。我们必须求解以下方程组

$$\begin{bmatrix} A\mathbf{u} \\ A\mathbf{v} \end{bmatrix} = \Lambda \cdot \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix},$$

其中$\mathbf{u}$ 和 $\mathbf{v}$ 是向量

$$\mathbf{u}=\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix},\ \mathbf{v}=\begin{bmatrix}y_1\\y_2\\y_3\\y_4\end{bmatrix}.$$

由此得到线性方程

$$\begin{array}{rrrrrrrrl}
 & & x_4 & -y_1 & & & & = & 0\\
-x_1 & & & & -y_2 & & & = & 0\\
 & -x_2 & & & & -y_3 & & = & 0\\
 & & -x_3 & & & & -y_4 & = & 0\\
x_1 & & & +\sqrt{2}y_1 & & & +y_4 & = & 0\\
 & x_2 & & -y_1 & +\sqrt{2}y_2 & & & = & 0\\
 & & x_3 & & -y_2 & +\sqrt{2}y_3 & & = & 0\\
 & & & x_4 & & -y_3 & +\sqrt{2}y_4 & = & 0.
\end{array}$$

因此，我们有

$$\mathbf{u}=\begin{bmatrix}0\\1\\\sqrt{2}\\1\end{bmatrix},\ \mathbf{v}=\begin{bmatrix}1\\0\\-1\\-\sqrt{2}\end{bmatrix},$$

因此，相应的超特征向量为

$$W=\langle\begin{bmatrix}0\\1\\\sqrt{2}\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\\-\sqrt{2}\end{bmatrix}\rangle.$$

同理，$\Lambda=\begin{bmatrix}0&-1\\1&-\sqrt{2}\end{bmatrix}$ 的相应超特征向量为

$$W=\langle\begin{bmatrix}0\\1\\-\sqrt{2}\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\\\sqrt{2}\end{bmatrix}\rangle.$$

因此，关于基

$$\begin{bmatrix}0\\1\\\sqrt{2}\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\\-\sqrt{2}\end{bmatrix},\begin{bmatrix}0\\1\\-\sqrt{2}\\1\end{bmatrix},\begin{bmatrix}1\\0\\-1\\\sqrt{2}\end{bmatrix}$$

$\mathbb{R}^4$的上述映射的矩阵将具有分块形式

$$\begin{bmatrix}0&-1&0&0\\1&\sqrt{2}&0&0\\0&0&0&-1\\0&0&1&-\sqrt{2}\end{bmatrix}.$$

参考文献

[1] Dennis S. Bernstein。《矩阵数学：理论、事实与公式》（第二版）。普林斯顿大学出版社，2009年。[2] Ernest Dieterich。旋转对的二维不变子空间的存在性与构造。《数学论坛》，第114卷，第203-211页，2009年。[3] Hakim Ikramov。关于共对角化矩阵。《线性代数及其应用》，第424卷，第456-465页，2007年。[4] Alan George、Hakim Ikramov。两个矩阵的公共不变子空间。《线性代数及其应用》，第287卷，第171-179页，1999年。[5] Steven Roman。《高等线性代数》（第三版）。施普林格，2008年。[6] Joseph J. Rotman。《高等现代代数》（第一版）。普伦蒂斯·霍尔，2003年。

数学系，理学院，苏丹卡布斯大学，马斯喀特，阿曼电子邮箱地址：omartalibmiran@gmail.com

数学系，理学院，苏丹卡布斯大学，马斯喀特，阿曼电子邮箱地址：m.ghalehlar@squ.edu.om