A new canonical form of square matrices

Daijun Tian

Department of Mathematics, Tianjin University, Tianjin, 300072, China

email: tiandaijun@tju.edu.cn

Keywords: invariant factor, elementary factor, canonical form, invariant subspace

Abstract. In this paper we give a new canonical form of square matrices over arbitrary number field. We also give several applications.

SOME FAMOUS CANONICAL FORM

A. Frobenius canonical form

Suppose $\mathbb{F}$ is arbitrary number field, $\boldsymbol{A} \in \mathbb{F}^{n\times n}$, $\boldsymbol{I}_n$ is the identity matrix of order n. $\lambda\boldsymbol{I}_n - \boldsymbol{A}$ has invariant factors group $\{d_1(\lambda), d_2(\lambda), \cdots, d_n(\lambda)\}$.

$\boldsymbol{A} \sim \boldsymbol{B} \Leftrightarrow \lambda\boldsymbol{I}_n - \boldsymbol{A},\ \lambda\boldsymbol{I}_n - \boldsymbol{B}$ have the same invariant factors group.

Suppose $d(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0 \in \mathbb{F}[\lambda](m \geqslant 1)$, then we call the matrix

$$\boldsymbol{F}_{d(\lambda)} = \begin{bmatrix} 0 & & & -a_0 \\ 1 & \ddots & & -a_1 \\ & \ddots & 0 & \vdots \\ & & 1 & -a_{m-1} \end{bmatrix}$$

is the **Frobenius block** with respect to $d(\lambda)$.

The invariant factors group of $\lambda\boldsymbol{I}_m - \boldsymbol{F}_{d(\lambda)}$ is $\{1, \cdots, 1, d(\lambda)\}$.

Suppose the degree $\geqslant 1$ invariant factors of $\lambda\boldsymbol{I}_n - \boldsymbol{A}$ are $d_i(\lambda), \cdots, d_n(\lambda)$, then

$$\boldsymbol{A} \sim \boldsymbol{F} = \mathrm{diag}(\boldsymbol{F}_{d_i(\lambda)}, \cdots, \boldsymbol{F}_{d_n(\lambda)}).$$

$\boldsymbol{F}$ is called the **Frobenius canonical form**[1] or **rational canonical form** of $\boldsymbol{A}$.

B. Jacobson canonical form

In $\mathbb{F}[\lambda]$, take the standard factorization of the degree $\geqslant 1$ invariant factors $d_i(\lambda), \cdots, d_n(\lambda)$ of $\lambda\boldsymbol{I}_n - \boldsymbol{A}$, then the greatest power of the irreducible factors are called the elementary factors.

$\boldsymbol{A} \sim \boldsymbol{B} \Leftrightarrow \lambda\boldsymbol{I}_n - \boldsymbol{A},\ \lambda\boldsymbol{I}_n - \boldsymbol{B}$ have the same elementary factors group.

For any $\boldsymbol{A}_1 \in \mathbb{F}^{n_1\times n_1}, \boldsymbol{A}_2 \in \mathbb{F}^{n_2\times n_2}$, the union of the elementary factors group of $\lambda\boldsymbol{I}_{n_1} - \boldsymbol{A}_1, \lambda\boldsymbol{I}_{n_2} - \boldsymbol{A}_2$ is just the elementary factors group of $\lambda\boldsymbol{I}_{n_1+n_2} - \mathrm{diag}(\boldsymbol{A}_1, \boldsymbol{A}_2)$.

Suppose $p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0 \in \mathbb{F}[\lambda]$ is irreducible, $\boldsymbol{E}_{1m} = \begin{bmatrix} \boldsymbol{0} & 1 \\ \boldsymbol{O} & \boldsymbol{0} \end{bmatrix}$ is the order m fundamental matrix, then we call the order km square matrix

$$\boldsymbol{J}_{p^k(\lambda)} = \begin{bmatrix} \boldsymbol{F}_{p(\lambda)} & & & \\ \boldsymbol{E}_{1m} & \ddots & & \\ & \ddots & \boldsymbol{F}_{p(\lambda)} & \\ & & \boldsymbol{E}_{1m} & \boldsymbol{F}_{p(\lambda)} \end{bmatrix}$$

is the **Jacobson block** with respect to $p^k(\lambda)$. $\lambda \boldsymbol{I}_{km} - \boldsymbol{J}_{p^k(\lambda)}$ has the unique elementary factor $p^k(\lambda)$. Over $\mathbb{F}$, $\boldsymbol{J}_{p^k(\lambda)}$ can not be similar to a quasi diagonal matrix consisting of two blocks of square matrices of smaller order.

Suppose $\boldsymbol{A} \in \mathbb{F}^{n\times n}$, and $\lambda \boldsymbol{I}_n - \boldsymbol{A}$ has the elementary factors group $\{p_1^{k_1}(\lambda), p_2^{k_2}(\lambda), \cdots, p_s^{k_s}(\lambda)\}$, then

$$\boldsymbol{A} \sim \boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_{p_1^{k_1}(\lambda)}, \boldsymbol{J}_{p_2^{k_2}(\lambda)}, \cdots, \boldsymbol{J}_{p_s^{k_s}(\lambda)}).$$

$\boldsymbol{J}$ is called the **Jacobson canonical form**[2-4] of $\boldsymbol{A}$.

C. Jordan canonical form

When $\mathbb{F} = \mathbb{C}$, we only have linear irreducible factors. The matrix $\boldsymbol{J}_{(\lambda-\lambda_0)^{k_0}} = \boldsymbol{J}_{k_0}(\lambda_0)$ is also called the order k_0 **Jordan block** with respect to the eigenvalue λ_0.

Suppose $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, and $\lambda \boldsymbol{I}_n - \boldsymbol{A}$ has the elementary factors group $\{(\lambda-\lambda_1)^{k_1}, (\lambda-\lambda_2)^{k_2}, \cdots, (\lambda-\lambda_s)^{k_s}\}$, then

$$\boldsymbol{A} \sim \boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_{k_1}(\lambda_1), \boldsymbol{J}_{k_2}(\lambda_2), \cdots, \boldsymbol{J}_{k_s}(\lambda_s)).$$

$\boldsymbol{J}$ is called the **Jordan canonical form**[5] or **complex similar canonical form** of $\boldsymbol{A}$.

D. Real similar canonical form

Suppose $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, and $\lambda \boldsymbol{I}_n - \boldsymbol{A}$ has the elementary factors group $\{(\lambda-\lambda_1)^{k_1}, \cdots, (\lambda-\lambda_s)^{k_s}; (\lambda^2 + p_1\lambda + q_1)^{l_1}, \cdots, (\lambda^2 + p_t\lambda + q_t)^{l_t}\}$, then $\boldsymbol{A}$ has the **real similar canonical form**[5-6]

$$\operatorname{diag}(\boldsymbol{J}_{k_1}(\lambda_1), \cdots, \boldsymbol{J}_{k_s}(\lambda_s); \boldsymbol{T}_{(\lambda^2+p_1\lambda+q_1)^{l_1}}, \cdots, \boldsymbol{T}_{(\lambda^2+p_t\lambda+q_t)^{l_t}}),$$

where $\boldsymbol{J}_{k_i}(\lambda_i)$ is the order k_i Jordan block with respect to the eigenvalue λ_i ($i = 1, \cdots, s$),

$$\boldsymbol{T}_{(\lambda^2+p_j\lambda+q_j)^{l_j}} = \begin{bmatrix} \boldsymbol{F}_{\lambda^2+p_j\lambda+q_j} & & & \\ \boldsymbol{I}_2 & \ddots & & \\ & \ddots & \boldsymbol{F}_{\lambda^2+p_j\lambda+q_j} & \\ & & \boldsymbol{I}_2 & \boldsymbol{F}_{\lambda^2+p_j\lambda+q_j} \end{bmatrix},$$

$\lambda \boldsymbol{I}_{2l_j} - \boldsymbol{T}_{(\lambda^2+p_j\lambda+q_j)^{l_j}}$ has the unique elementary factor $(\lambda^2 + p_j\lambda + q_j)^{l_j}$ ($j = 1, \cdots, t$).

When $l > 1$, $\boldsymbol{T}_{(\lambda^2+p\lambda+q)^l}$ is not a Jacobson block. So the Jacobson canonical form is not the generalization of the real similar canonical form.

A NEW CANONICAL FORM

Let $p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0 \in \mathbb{F}[\lambda]$ be irreducible, we introduce the order km square matrix

$$\boldsymbol{T}_{p^k(\lambda)} = \begin{bmatrix} \boldsymbol{F}_{p(\lambda)} & & & \\ \boldsymbol{I}_m & \ddots & & \\ & \ddots & \boldsymbol{F}_{p(\lambda)} & \\ & & \boldsymbol{I}_m & \boldsymbol{F}_{p(\lambda)} \end{bmatrix}$$

which has similar properties as the Jacobson block $\boldsymbol{J}_{p^k(\lambda)}$, and is more useful.

Theorem 1 $\lambda \boldsymbol{I}_{km} - \boldsymbol{T}_{p^k(\lambda)}$ has the unique elementary factor $p^k(\lambda)$.

Proof: Suppose $p(\lambda)$ has roots $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{C}$, then the order m square matrix $\boldsymbol{F}_{p(\lambda)}$ has m distinct complex eigenvalues, and there exists an invertible matrix $\boldsymbol{S} \in \mathbb{C}^{m\times m}$, such that

$$\boldsymbol{S}^{-1}\boldsymbol{F}_{p(\lambda)}\boldsymbol{S} = \operatorname{diag}(\lambda_1, \lambda_2, \cdots, \lambda_m)$$ [6].

Then

$$\lambda \boldsymbol{I}_{km} - \boldsymbol{T}_{p^k(\lambda)} = \begin{bmatrix} \lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)} & & & \\ -\boldsymbol{I}_m & \ddots & & \\ & \ddots & \lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)} & \\ & & -\boldsymbol{I}_m & \lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)} \end{bmatrix} \cong \begin{bmatrix} \boldsymbol{O} & & & [\lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)}]^k \\ -\boldsymbol{I}_m & \ddots & & \vdots \\ & \ddots & \boldsymbol{O} & [\lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)}]^2 \\ & & -\boldsymbol{I}_m & \lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)} \end{bmatrix}$$

$$\cong \begin{bmatrix} \boldsymbol{O} & [\lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)}]^k \\ -\boldsymbol{I}_{(k-1)m} & \boldsymbol{O} \end{bmatrix} \cong \begin{bmatrix} \boldsymbol{I}_{(k-1)m} & \boldsymbol{O} \\ \boldsymbol{O} & [\lambda\boldsymbol{I}_m - \boldsymbol{F}_{p(\lambda)}]^k \end{bmatrix} \cong \begin{bmatrix} \boldsymbol{I}_{(k-1)m} & \boldsymbol{O} \\ \boldsymbol{O} & [\lambda\boldsymbol{I}_m - \boldsymbol{S}^{-1}\boldsymbol{F}_{p(\lambda)}\boldsymbol{S}]^k \end{bmatrix}$$

$$= \operatorname{diag}\left(\boldsymbol{I}_{(k-1)m}, (\lambda-\lambda_1)^k, (\lambda-\lambda_2)^k, \cdots, (\lambda-\lambda_m)^k\right) \cong \operatorname{diag}\left(\boldsymbol{I}_{(k-1)m}, \boldsymbol{I}_{m-1}, (\lambda-\lambda_1)^k(\lambda-\lambda_2)^k\cdots(\lambda-\lambda_m)^k\right)$$

$$= \operatorname{diag}\left(\boldsymbol{I}_{km-1}, p^k(\lambda)\right).$$

Thus the unique elementarty factor of $\lambda\boldsymbol{I}_{km} - \boldsymbol{T}_{p^k(\lambda)}$ is $p^k(\lambda)$. □

Corollary 1 $\boldsymbol{J}_{p^k(\lambda)} \sim \boldsymbol{T}_{p^k(\lambda)}$. □

Corollary 2 Over $\mathbb{F}$, $\boldsymbol{T}_{p^k(\lambda)}$ can not be similar to a quasi diagonal matrix consisting of two blocks of square matrices of smaller order. □

Theorem 2 Suppose $\boldsymbol{A} \in \mathbb{F}^{n\times n}$, and $\lambda\boldsymbol{I}_n - \boldsymbol{A}$ has the elementary factors group $\{p_1^{k_1}(\lambda), p_2^{k_2}(\lambda), \cdots, p_s^{k_s}(\lambda)\}$, then $\boldsymbol{A} \sim \boldsymbol{T} = \operatorname{diag}(\boldsymbol{T}_{p_1^{k_1}(\lambda)}, \boldsymbol{T}_{p_2^{k_2}(\lambda)}, \cdots, \boldsymbol{T}_{p_s^{k_s}(\lambda)})$. □

$\boldsymbol{T}$ is just the best similar canonical form of $\boldsymbol{A}$. When $\mathbb{F} = \mathbb{C}$, $\boldsymbol{T}$ reduces to the complex similar canonical form of $\boldsymbol{A}$. When $\mathbb{F} = \mathbb{R}$, $\boldsymbol{T}$ reduces to the real similar canonical form of $\boldsymbol{A}$. The new similar canonical form is a common generalization of both complex and real similar canonical form.

INVARIANT SUBSPACE

Ler V be a n dimensional vector space over $\mathbb{F}$, and $\mathsf{A} \in \operatorname{End}V$ has matrix representation

$$\boldsymbol{J} = \operatorname{diag}(\boldsymbol{J}_{p_1^{k_1}(\lambda)}, \boldsymbol{J}_{p_2^{k_2}(\lambda)}, \cdots, \boldsymbol{J}_{p_s^{k_s}(\lambda)}) \text{ or } \boldsymbol{T} = \operatorname{diag}(\boldsymbol{T}_{p_1^{k_1}(\lambda)}, \boldsymbol{T}_{p_2^{k_2}(\lambda)}, \cdots, \boldsymbol{T}_{p_s^{k_s}(\lambda)}),$$

then V is decomposed as the direct sum of s A – invariant subspaces. These invariant subspaces are called the **Jacobson subspaces** of A, they can not be decomposed as the direct sum of two smaller dimensional invariant subspaces.

The invariant subspace with respect to the invariant factor is called the **Frobenius subspace** of A, it can be decomposed as the direct sum of the Jacobson subspaces.

Theorem 3 Ler V be a n dimensional vector space over $\mathbb{F}$, $\mathsf{A} \in \operatorname{End}V$, and

$$p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0$$

is an irreducible factor of $f_{\mathsf{A}}(\lambda) = \det(\lambda\mathsf{E} - \mathsf{A}) \in \mathbb{F}[\lambda]$, then A must have a m dimensional invariant subspace W, and $\mathsf{A}|_W$ has matrix representation $\boldsymbol{F}_{p(\lambda)}$.

Proof: From theorem 2, A has matrix representation $\boldsymbol{T} = \operatorname{diag}(\cdots, \boldsymbol{T}_{p^k(\lambda)})$ under certain basis of V. Denote by W the subspace generated by the last m vectors in this basis, then W is a m

dimensional $\mathcal{A}$ –invariant subspace, and $\mathcal{A}|_W$ has matrix representation $\boldsymbol{F}_{p(\lambda)}$. □

Remark W is a minimal nonzero $\mathcal{A}$ –invariant subspace. In fact, $p(\lambda)$ is the characteristic polynomial of $\mathcal{A}|_W$. Assume $U \subset W$ is a smaller dimensional $\mathcal{A}$ –invariant subspace, $q(\lambda)$ is the characteristic polynomial of $\mathcal{A}|_U$, then $q(\lambda) | p(\lambda)$. This contradicts the fact that $p(\lambda)$ is irreducible.

Theorem 4 Let $p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0 \in \mathbb{F}[\lambda]$ be irreducible, $\mathcal{A} \in \mathrm{End}V$ has matrix representation $\boldsymbol{J}_{p^k(\lambda)}$ or $\boldsymbol{T}_{p^k(\lambda)}$, then $\mathcal{A}$ has exactly $k+1$ invariant subspaces.

Proof: Suppose W is an $\mathcal{A}$ –invariant subspace, then

$$f_{\mathcal{A}}(\lambda) = \det(\lambda\mathcal{E} - \mathcal{A}) = p^k(\lambda), \quad f_{\mathcal{A}|_W}(\lambda) \mid f_{\mathcal{A}}(\lambda), \quad f_{\mathcal{A}|_W}(\lambda) = \det(\lambda\mathcal{E}|_W - \mathcal{A}|_W) = p^i(\lambda), \quad i = 0, 1, \cdots, k.$$

Denote by $W_i = \mathrm{Ker}\, p^i(\mathcal{A})$, then

$$\{\boldsymbol{0}\} = W_0 < W_1 < \cdots < W_k = V, \quad p^i(\mathcal{A})W = \{\boldsymbol{0}\} \Rightarrow W \subset W_i, \quad \dim W = im, \quad i = 0, 1, \cdots, k.$$

From the proof of theorem 1 we see that both $\boldsymbol{J}_{p^k(\lambda)}, \boldsymbol{T}_{p^k(\lambda)}$ are similar to

$$\boldsymbol{J} = \mathrm{diag}(\boldsymbol{J}_k(\lambda_1), \boldsymbol{J}_k(\lambda_2), \cdots, \boldsymbol{J}_k(\lambda_m))$$

over $\mathbb{C}$, where $\lambda_1, \lambda_2, \cdots, \lambda_m \in \mathbb{C}$ are all the roots of $p(\lambda)$. Hence

$$\dim W_i = km - r[p^i(\boldsymbol{J})] = im = \dim W, \quad W = W_i, \quad i = 0, 1, \cdots, k.$$

Therefore $\mathcal{A}$ has $k+1$ invariant subspaces: $W_0 = \{\boldsymbol{0}\}, W_1, \cdots, W_k = V$. □

By theorem 2 and 4 we have the following result.

Theorem 5 Ler V be a n dimensional vector space over $\mathbb{F}$, the characteristic polynomial of $\mathcal{A} \in \mathrm{End}V$ has the standard factorization $f_{\mathcal{A}}(\lambda) = \det(\lambda\mathcal{E} - \mathcal{A}) = p_1^{k_1}(\lambda)p_2^{k_2}(\lambda)\cdots p_s^{k_s}(\lambda)$, then

$$\{\det(\lambda\mathcal{E}|_W - \mathcal{A}|_W) : W < V, \ \mathcal{A}W \subset W\} = \{p_1^{l_1}(\lambda)p_2^{l_2}(\lambda)\cdots p_s^{l_s}(\lambda) : \ 0 \leqslant l_i \leqslant k_i\},$$

$$\{\dim W : W < V, \ \mathcal{A}W \subset W\} = \left\{\sum_{i=1}^{s} l_i \deg p_i(\lambda) : \ 0 \leqslant l_i \leqslant k_i\right\}. \quad □$$

References

[1] Roger A. Horn, Charles R. Johnson. Matrix analysis. England : Cambridge University Press, 1985.

[2] Jacobson N. Lectures in Abstract Algebra , vol2 : linear algebra. New York : Springer-Verlag, 1953.

[3] Krull W. Über Begleitmatrizen und Elementarteilertheorie(PhD thesis, 1921). Gesammelte Abhandlungen collected papers. Berlin : de Gruyter, 1999. vol1, 55-95.

[4] Boxun Tu, Chenghao Xu, Fen Wang. Advanced algebra. Shanghai : Shanghai scientific & technical Publishers, 1987(in Chinese).

[5] Daoji Meng. Advanced algebra and analytic geometry(2nd edition). Beijing : Science Press, 2006(in Chinese).

[6] Daijun Tian. New compilation of linear algebra. Tianjin : Tianjin University Press, 2012(in Chinese).