定理 25.1　n 次複素正方行列 A は $\mathbf{C}^n$ の正規直交基底により「三角化」される．すなわち，固有多項式が

$$\det(A-\lambda E_n)=(\lambda_1-\lambda)^{m_1}(\lambda_2-\lambda)^{m_2}\cdots(\lambda_k-\lambda)^{m_k}$$

と表されるとき (E_n は n 次の単位行列)，適当な $\mathbf{C}^n$ 正規直交基底 $\{\,\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n\,\}$ をとり $U=(\ \mathbf{e}_1\ \mathbf{e}_2\ \cdots\ \mathbf{e}_n\ )$ と置くことにより

$$A=U\begin{pmatrix} \underbrace{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}_{m_1\text{ 個}} & & & & \Huge{*} \\ & \underbrace{\begin{matrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{matrix}}_{m_2\text{ 個}} & & & \\ & & \ddots & & \\ \Huge{O} & & & \underbrace{\begin{matrix}\lambda_k & & \\ & \ddots & \\ & & \lambda_k\end{matrix}}_{m_k\text{ 個}} \end{pmatrix}U^{-1}$$

という形に表すことができる．

(証明)　まず，固有値 λ_1 に属する固有ベクトル $\mathbf{p}_1$ を一つとり，$\mathbf{u}_1$ を含む $\mathbf{C}^n$ の正規直交基底 $\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_n\}$ をつくる．このとき

$$A(\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n)=(\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n)\begin{pmatrix}\lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

と表されるが，このとき

$$A-\lambda E_n=(\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n)\begin{pmatrix}\lambda_1-\lambda & * & \cdots & * \\ 0 & & & \\ \vdots & & A_1-\lambda E_{n-1} & \\ 0 & & & \end{pmatrix}(\mathbf{p}_1\ \mathbf{p}_2\ \cdots\ \mathbf{p}_n)^{-1}$$

から $\det(A-\lambda E_n)=(\lambda_1-\lambda)\det(A_1-\lambda E_{n-1})$ となり，従って A_1 の固有多項式は

$$\det(A_1-\lambda E_{n-1})=(\lambda_1-\lambda)^{m_1-1}(\lambda_2-\lambda)^{m_2}\cdots(\lambda_k-\lambda)^{m_k}$$

である．この手順を繰り返すことで正則行列 P により

$$A=P\begin{pmatrix}\lambda_1 & & * \\ & \ddots & \\ & & \lambda_k\end{pmatrix}P^{-1}$$

という形にできるが，さらに P の列ベクトルを Schmidt の方法で正規直交化することにより主張の形が得られることを見るのは容易である．

命題 25.2 n 次複素正方行列 A について，ある $N \in \mathbf{N}$ に対して $\mathrm{rank} A^{N+1} = \mathrm{rank} A^N$ となるならば任意の $k \in \mathbf{N}$ に対して $\mathrm{rank} A^{N+k} = \mathrm{rank} A_N$ が成り立つ．

(証明) $\mathrm{rank} A^{N+1} = \mathrm{rank} A^N = m$ とおく．まず

$$\mathrm{Im} f_{A^{N+2}} \subset \mathrm{Im} f_{A^{N+1}} = \mathrm{Im} f_{A^N}$$

ゆえ $\mathrm{rank} A^{N+2} \leq m$ であることに注意しよう．$\mathrm{rank} A^{N+2} \geq m$ を示せばよい．そこで，$A^N = (\ \mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n\ )$ と表したとき

$$A^{N+1} = (\ A\mathbf{a}_1\ A\mathbf{a}_2\ \cdots\ A\mathbf{a}_n\ )$$

であるが，$\mathrm{rank} A^{N+1} = m$ よりこの列ベクトルの中から m 本を選んで一次独立な組

$$\{\ A\mathbf{a}_{n_1}, A\mathbf{a}_{n_2}, \ldots, A\mathbf{a}_{n_m}\ \}$$

をつくることができる．このとき $\{\ \mathbf{a}_{n_1}, \mathbf{a}_{n_2}, \ldots, \mathbf{a}_{n_m}\ \}$ も一次独立なので，

$$\mathrm{Im} f_{A^{N+1}} = \langle\ A\mathbf{a}_{n_1}, A\mathbf{a}_{n_2}, \ldots, A\mathbf{a}_{n_m}\ \rangle$$

$$\mathrm{Im} f_{A^N} = \langle\ \mathbf{a}_{n_1}, \mathbf{a}_{n_2}, \ldots, \mathbf{a}_{n_m}\ \rangle$$

であり $\mathrm{Im} f_{A^{N+1}} = \mathrm{Im} f_{A^N}$ ゆえ

$$\langle\ A\mathbf{a}_{n_1}, A\mathbf{a}_{n_2}, \ldots, A\mathbf{a}_{n_m}\ \rangle = \langle\ \mathbf{a}_{n_1}, \mathbf{a}_{n_2}, \ldots, \mathbf{a}_{n_m}\ \rangle$$

とわかる．従って，基底の変換行列として

$$(\ A\mathbf{a}_{n_1}\ A\mathbf{a}_{n_2}\ \cdots\ A\mathbf{a}_{n_m}\ ) = (\ \mathbf{a}_{n_1}\ \mathbf{a}_{n_2}\ \cdots\ \mathbf{a}_{n_m})P$$

を満たす正則行列 P が存在し，このとき

$$\mathrm{rank}(\ A^2\mathbf{a}_{n_1}\ A^2\mathbf{a}_{n_2}\ \cdots\ A^2\mathbf{a}_{n_m}\ ) = \mathrm{rank}\big\{(\ A\mathbf{a}_{n_1}\ A\mathbf{a}_{n_2}\ \cdots\ A\mathbf{a}_{n_m}\ )P\big\} = m$$

となる．よって $\{\ A^2\mathbf{a}_{n_1}, A^2\mathbf{a}_{n_2}, \cdots, A^2\mathbf{a}_{n_m}\ \}$ は一次独立，しかも $A^2\mathbf{a}_{n_1}, A^2\mathbf{a}_{n_2}, \ldots, A^2\mathbf{a}_{n_m}$ は A^{N+2} の列ベクトルに含まれるのだから，このことは $\mathrm{rank} A^{N+2} \geq m$ であることを意味する．

定理 25.3　n 次複素正方行列 A の固有多項式が

$$\det(A-\lambda E_n) = (\lambda_1-\lambda)^{m_1}(\lambda_2-\lambda)^{m_2}\cdots(\lambda_k-\lambda)^{m_k}$$

と表されるとき

$$K_i = \{\,\mathbf{v}\in\mathbf{C}^n \mid (A-\lambda_i E)^{m_i}\mathbf{v}=\mathbf{0}\,\},\quad (i=1,2,\ldots,k)$$

とおくと，次が成り立つ：

(a) $\dim K_i = m_i,\quad (i=1,2,\ldots,k)$

(b) $\mathbf{v}\in K_i\backslash\{\mathbf{0}\}$ ならば $\displaystyle\prod_{j\neq i}(A-\lambda_j E)^{m_j}\mathbf{v}\neq\mathbf{0}$

(c) 各 K_i の基底を $\{\mathbf{a}^i_1,\ldots,\mathbf{a}^i_{m_i}\}$ とすると，これらをすべて合わせた $\displaystyle\bigcup_{i=1}^{k}\{\mathbf{a}^i_1,\ldots,\mathbf{a}^i_{m_i}\}$ は $\mathbf{C}^n$ の基底である．

(証明)

(a) 定理 25.1 より，A は適当な正則行列 U を用いて

$$A = U\begin{pmatrix} \underbrace{\begin{matrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_1\end{matrix}}_{m_1\text{ 個}} & & & * \\ & \underbrace{\begin{matrix}\lambda_2 & & \\ & \ddots & \\ & & \lambda_2\end{matrix}}_{m_2\text{ 個}} & & \\ & & \ddots & \\ O & & & \underbrace{\begin{matrix}\lambda_k & & \\ & \ddots & \\ & & \lambda_k\end{matrix}}_{m_k\text{ 個}} \end{pmatrix}U^{-1}$$

と表すことができる．このとき

$$A-\lambda_1E_n = U\begin{pmatrix} \boxed{\begin{matrix}0 & & * \\ & \ddots & \\ & & 0\end{matrix}} & & & & & * \\ & \lambda_2-\lambda_1 & & & & \\ & & \ddots & & & \\ & & & \lambda_2-\lambda_1 & & \\ & & & & \lambda_k-\lambda_1 & \\ & & & & & \ddots \\ O & & & & & & \lambda_k-\lambda_1 \end{pmatrix}U^{-1}$$

となる．点線で囲んだ部分に注目すると

$$\begin{pmatrix}0 & & * \\ & \ddots & \\ & & 0\end{pmatrix}^{m_1} = O_{m_1}\qquad (m_1 \text{ 次の零行列})$$

となることが容易にわかるから

$$
(A-\lambda_1 E_n)^{m_1} = U \begin{pmatrix} \boxed{O_{m_1}} & & & & & & * \\ & (\lambda_2-\lambda_1)^{m_1} & & & & & \\ & & \ddots & & & & \\ & & & (\lambda_2-\lambda_1)^{m_1} & & & \\ & & & & \ddots & & \\ O & & & & & (\lambda_k-\lambda_1)^{m_1} & \\ & & & & & \ddots & \\ & & & & & & (\lambda_k-\lambda_1)^{m_1} \end{pmatrix} U^{-1}
$$

という形になる．よって，$\lambda_2-\lambda_1 \neq 0, \ldots, \lambda_k-\lambda_1 \neq 0$ より $\mathrm{rank}(A-\lambda_1 E_n)^{m_1} = n-m_1$，従って $\mathrm{dim}K_1 = \mathrm{dimKer}f_{(A-\lambda_1 E_n)^{m_1}} = m_1$ とわかる．$K_2, \ldots, K_k$ についても同様.

(b) $i=1$ の場合を示す．(a) と同様に A を表すと

$$
A-\lambda_j E_n = U \begin{pmatrix} \underbrace{\begin{matrix} \lambda_1-\lambda_j & & \\ & \ddots & \\ & & \lambda_1-\lambda_j \end{matrix}}_{m_1\text{ 個}} & & & * \\ & \lambda_2-\lambda_j & & \\ & & \ddots & \\ O & & & \lambda_k-\lambda_j \end{pmatrix} U^{-1}
$$

という形になる．そこで，$d_j = \lambda_1-\lambda_j$ とおくと

$$
(A-\lambda_j E_n)^{m_j} = U \begin{pmatrix} {d_j}^{m_j} & & & & * \\ & \ddots & & & \\ & & {d_j}^{m_j} & & \\ & & & * & \\ O & & & & \ddots \\ & & & & * \end{pmatrix} U^{-1}
$$

から

$$
\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j} = U \begin{pmatrix} \prod_{j=2}^{k} {d_j}^{m_j} & & & & * \\ & \ddots & & & \\ & & \prod_{j=2}^{k} {d_j}^{m_j} & & \\ & & & * & \\ O & & & & \ddots \\ & & & & * \end{pmatrix} U^{-1}
$$

となる．ところで，(a) で見たように $(A-\lambda_1 E_n)^{m_1}$ は

$$
(A-\lambda_1 E_n)^{m_1} = U \begin{pmatrix} \boxed{O_{m_1}} & & & * \\ & (\lambda_2-\lambda_1)^{m_1} & & \\ & & \ddots & \\ O & & & (\lambda_k-\lambda_1)^{m_1} \end{pmatrix} U^{-1}
$$

という形になるから，$\mathbf{v} \in K_1 \backslash \{\mathbf{0}\}$，すなわち $(A-\lambda_1 E_n)^{m_1}\mathbf{v} = \mathbf{0}$ かつ $\mathbf{v} \neq \mathbf{0}$ ならば

$$
U^{-1}\mathbf{v} = \begin{pmatrix} c_1 \\ \vdots \\ c_{m_1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad (c_1, \ldots, c_{m_1}) \neq (0, \ldots, 0)
$$

と表される．従ってこのとき

$$\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v} = U \begin{pmatrix} \prod_{j=2}^{k} d_j{}^{m_j} & & & & \\ & \ddots & & * & \\ & & \prod_{j=2}^{k} d_j{}^{m_j} & & \\ & & & * & \\ & O & & & \ddots & \\ & & & & & * \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{m_1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

となるが，$d_j \neq 0 \ (j=2,\ldots,k)$ であるから

$$\begin{pmatrix} \prod_{j=2}^{k} d_j{}^{m_j} & & * \\ & \ddots & \\ O & & \prod_{j=2}^{k} d_j{}^{m_j} \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_{m_1} \end{pmatrix}$$

は零ベクトルではない．よって，U の正則性にも注意すると $\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v} \neq \mathbf{0}$ ということになる．

(c) 次のことを示せば十分であろう：

$\mathbf{v}_i \in K_i \backslash \{\mathbf{0}\} \ (i=1,2,\ldots,k)$ ならば $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ は一次独立である

そこで

$$t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k = \mathbf{0}, \quad t_1, t_2, \ldots, t_k \in C$$

とおく．このとき

$$\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}(t_1\mathbf{v}_1 + t_2\mathbf{v}_2 + \cdots + t_k\mathbf{v}_k) = \mathbf{0}$$

であるが

$$\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v}_2 = \cdots = \prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v}_k = \mathbf{0}$$

は明らかだから

$$t_1 \prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v}_1 = \mathbf{0}$$

さらに (b) より $\prod_{j=2}^{k}(A-\lambda_j E_n)^{m_j}\mathbf{v}_1 \neq \mathbf{0}$ であるから $t_1 = 0$．同様にして $t_2 = \cdots = t_k = 0$ となる．