Schur 分解

正方行列は上三角行列に変換できることを示します。
前半はブロック行列の積と行列式を求めています。
行列は大文字のローマ文字、スカラーは小文字のローマ文字かギリシャ文字、ベクトル ($n\times 1$ 行列) は太字にしています。

知らなくてもここでは問題ないですが、ついでなのでブロック行列の積と行列式を求めます。2×3 行列と 3×3 行列をブロックに分けて

$$\begin{pmatrix} e_{11} & e_{12} & f_{13} \\ e_{21} & e_{22} & f_{23} \end{pmatrix} = \begin{pmatrix} E & F \end{pmatrix} \ , \ \begin{pmatrix} a_{11} & a_{12} & c_{13} \\ a_{21} & a_{22} & c_{23} \\ d_{31} & d_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} A & C \\ D & B \end{pmatrix}$$

と分けたとき、積は

$$\begin{pmatrix} E & F \end{pmatrix}\begin{pmatrix} A & C \\ D & B \end{pmatrix} = \begin{pmatrix} EA+FD & EC+FB \end{pmatrix}$$

となります。実際に、成分を見ると

$$(1,1)\ e_{11}a_{11}+e_{12}a_{21}+f_{13}d_{31}\ ,\ (1,2)\ e_{11}a_{12}+e_{12}a_{22}+f_{13}d_{32}$$

$$(2,1)\ e_{21}a_{11}+e_{22}a_{21}+f_{23}d_{31}\ ,\ (2,2)\ e_{21}a_{12}+e_{22}a_{22}+f_{23}d_{32}$$

$$(1,3)\ e_{11}c_{13}+e_{12}c_{23}+f_{13}b_{33}\ ,\ (2,3)\ e_{21}c_{13}+e_{22}c_{23}+f_{23}b_{33}$$

上の 4 つは 2×2 行列 $EA+FD$、下の 2 つは 2×1 行列 $EC+FB$ です。成分を増やしてもブロック行列の積は、このようにブロックをスカラーとして扱った形になります。ただし、ブロックは行列なので、積が成立するのは積を取るブロックでの列と行の数が揃っている場合です。

簡単に示します。$m\times p$ 行列 $S,\ p\times n$ 行列 T をブロック行列として、その積を

$$M = ST = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}\begin{pmatrix} B_{11} & \cdots \\ B_{21} & \cdots \\ \vdots & \cdots \\ B_{r1} & \cdots \end{pmatrix}$$

とします。A_{ij}, B_{ij} は行列で、添え字は行列の区別であって成分ではないです。行列の成分は $(M)_{ij}$ か小文字で表記します。A_{1i} は $a\times c_i$ 行列、B_{j1} は $c_j\times b$ 行列とします。積を取るので A_{1i} の列の数と B_{i1} の行の数は一致させ、$c_1+\cdots+c_r=p$ です。M の $(1,1)$ 成分は、S の成分を s_{ij}、T の成分を t_{ij} として

$$(M)_{11} = s_{11}t_{11}+s_{12}t_{21}+\cdots+s_{1p}t_{p1}$$

これは A_{11}, B_{11} の成分で書けば

$$
\begin{aligned}
(M)_{11} &= (A_{11})_{11}(B_{11})_{11} + (A_{11})_{12}(B_{11})_{21} + \cdots + (A_{11})_{1c_1}(B_{11})_{c_1 1} + \sum_{k=c_1+1}^{p} s_{1k}t_{k1} \\
&= (A_{11}B_{11})_{11} + \sum_{k=c_1+1}^{p} s_{1k}t_{k1}
\end{aligned}
$$

A_{12} の列の数は c_2 なので、c_1+1 から c_1+c_2 までの $s_{1k}t_{k1}$ は

$$
(A_{12})_{11}(B_{21})_{11} + (A_{12})_{12}(B_{21})_{21} + \cdots + (A_{12})_{1c_2}(B_{21})_{c_2 1} = (A_{12}B_{21})_{11}
$$

同様に書き換えていけば

$$
\begin{aligned}
(M)_{11} &= \sum_{k=1}^{c_1}(A_{11})_{1k}(B_{11})_{k1} + \sum_{k=1}^{c_2}(A_{12})_{1k}(B_{21})_{k1} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{1k}(B_{r1})_{k1} \\
(M)_{12} &= \sum_{k=1}^{c_1}(A_{11})_{1k}(B_{11})_{k2} + \sum_{k=1}^{c_2}(A_{12})_{1k}(B_{21})_{k2} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{1k}(B_{r1})_{k2} \\
&\vdots \\
(M)_{1b} &= \sum_{k=1}^{c_1}(A_{11})_{1k}(B_{11})_{kb} + \sum_{k=1}^{c_2}(A_{12})_{1k}(B_{21})_{kb} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{1k}(B_{r1})_{kb} \\
(M)_{21} &= \sum_{k=1}^{c_1}(A_{11})_{2k}(B_{11})_{k1} + \sum_{k=1}^{c_2}(A_{12})_{2k}(B_{21})_{k1} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{2k}(B_{r1})_{k1} \\
&\vdots \\
(M)_{a1} &= \sum_{k=1}^{c_1}(A_{11})_{ak}(B_{11})_{k1} + \sum_{k=1}^{c_2}(A_{12})_{ak}(B_{21})_{k1} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{ak}(B_{r1})_{k1} \\
&\vdots \\
(M)_{ab} &= \sum_{k=1}^{c_1}(A_{11})_{ak}(B_{11})_{kb} + \sum_{k=1}^{c_2}(A_{12})_{ak}(B_{21})_{kb} + \cdots + \sum_{k=1}^{c_r}(A_{1r})_{ak}(B_{r1})_{kb}
\end{aligned}
$$

として、続いていきます。分かりやすくすれば

$$
\begin{aligned}
(M)_{11} =& (A_{11}B_{11})_{11} + (A_{12}B_{21})_{11} + \cdots + (A_{1r}B_{r1})_{11} \\
(M)_{12} =& (A_{11}B_{11})_{12} + (A_{12}B_{21})_{12} + \cdots + (A_{1r}B_{r1})_{12} \\
&\vdots \\
(M)_{1b} =& (A_{11}B_{11})_{1b} + (A_{12}B_{21})_{1b} + \cdots + (A_{1r}B_{r1})_{1b} \\
(M)_{21} =& (A_{11}B_{11})_{21} + (A_{12}B_{21})_{21} + \cdots + (A_{1r}B_{r1})_{21} \\
&\vdots \\
(M)_{a1} =& (A_{11}B_{11})_{a1} + (A_{12}B_{21})_{a1} + \cdots + (A_{1r}B_{r1})_{a1} \\
&\vdots \\
(M)_{ab} =& (A_{11}B_{11})_{ab} + (A_{12}B_{21})_{ab} + \cdots + (A_{1r}B_{r1})_{ab}
\end{aligned}
$$

なので、M の $(1,1)$ 成分から (a,b) 成分までは

$$
A_{11}B_{11} + A_{12}B_{21} + \cdots A_{1r}B_{r1}
$$

と一致します。M の他の成分も同様に構成していけるので、ブロック行列の積は各ブロックをスカラーとして扱うことで求められます。例えば、$m \times p$ 行列 $S,\ p \times n$ 行列 T をブロック行列として

$$
S = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix} ,\ T = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \\ B_{31} & B_{32} \end{pmatrix}
$$

としたとき、A_{1i}, A_{2i} が $m_1 \times c_i, m_2 \times c_i$ 行列、B_{i1}, B_{i2} が $c_i \times n_1, c_i \times n_2$ 行列なら $(m_1 + m_2 = m,\ n_1 + n_2 + n_3 = n)$、$ST$ は

$$
ST = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \end{pmatrix}
$$

となります。

ブロック行列の行列式を求めます。まず、4×4 行列 M を

$$
M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & c_{11} & c_{12} \\ a_{21} & a_{22} & c_{21} & c_{22} \\ 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & b_{21} & b_{22} \end{pmatrix}
$$

として、$A, B, C, 0$ に分けたときの行列式を求めます。行列 M の成分を m_{ij} とします。A の下側の成分は 0 なので、m_{ij} は $i \geq 3,\ j \leq 2$ のとき 0 です $(m_{31}, m_{32}, m_{41}, m_{42} = 0)$。そうすると、行列式は

$$\begin{aligned}\det M &= \sum_{i,j,k,l=1}^{4} \epsilon_{ijkl} m_{1i} m_{2j} m_{3k} m_{4l} \\ &= \epsilon_{1234} m_{11} m_{22} m_{33} m_{44} + \epsilon_{1243} m_{11} m_{22} m_{34} m_{43} \\ &\quad + \epsilon_{2143} m_{12} m_{21} m_{34} m_{43} + \epsilon_{2134} m_{12} m_{21} m_{33} m_{44}\end{aligned}$$

となり、m_{ij} での i が 2 までででは j も 2 まで、i が 3 以上なら j も 3 以上の組み合わせになります。ϵ はレヴィ・チビタ記号で、$\epsilon_{1234} = +1$ です。変形すれば

$$\begin{aligned}\det M &= m_{11} m_{22} m_{33} m_{44} - m_{11} m_{22} m_{34} m_{43} + m_{12} m_{21} m_{34} m_{43} - m_{12} m_{21} m_{33} m_{44} \\ &= a_{11} a_{22} b_{11} b_{22} - a_{11} a_{22} b_{12} b_{21} + a_{12} a_{21} b_{12} b_{21} - a_{12} a_{21} b_{11} b_{22} \\ &= a_{11} a_{22} (b_{11} b_{22} - b_{12} b_{21}) - a_{12} a_{21} (b_{11} b_{22} - b_{12} b_{21}) \\ &= (a_{11} a_{22} - a_{12} a_{21})(b_{11} b_{22} - b_{12} b_{21}) \\ &= \det A \det B\end{aligned}$$

となり、A, B の行列式の積になります。この場合でのレヴィ・チビタ記号は、前の 2 つは $1, 2$、後ろの 2 つは $3, 4$ となっていて、$1, 2, 3, 4$ の並びに対してそれぞれ分離しているので

$$\epsilon_{1234} m_{11} m_{22} m_{33} m_{44} = \epsilon_{12} m_{11} m_{22} \epsilon_{34} m_{33} m_{44} = m_{11} m_{22} m_{33} m_{44}$$

$$\epsilon_{1243} m_{11} m_{22} m_{34} m_{43} = \epsilon_{12} m_{11} m_{22} \epsilon_{43} m_{34} m_{43} = -m_{11} m_{22} m_{34} m_{43}$$

と書き換えて、$\epsilon_{34} = +1, \epsilon_{43} = -1$ とできるようになっています。なので

$$\det M = \sum_{i,j=1}^{2} \epsilon_{ij} m_{1i} m_{2j} \sum_{k,l=3}^{4} \epsilon_{kl} m_{3k} m_{4l} = \det A \det B$$

と分かります。

これはそのまま一般化できて、$r \times r$ 行列 A、$s \times s$ 行列 B、$r \times s$ 行列 C によって

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}$$

となっている $n \times n$ 行列 M の行列式は

$$\det M = \det A \det B$$

となります。4×4 行列の手順を $n \times n$ 行列にして繰り返せば示せます。$n \times n$ 行列の行列式は

$$\det M = \sum_{k_1,k_2,\cdots,k_n=1}^{n} \epsilon_{k_1k_2\cdots k_n} m_{1k_1} m_{2k_2} \cdots m_{nk_n}$$

A の下側は 0 なので、この中で m_{ij} は $i \geq r+1$、$j \leq r$ なら 0 です。このため、m_{ij} での i が r までではj も r まで、i が $r+1$ 以上では j は $r+1$ 以上による組み合わせになり

$$\det M = \sum_{k_1,\cdots,k_r=1}^{r} \sum_{k_{r+1},\cdots,k_n=r+1}^{n} \epsilon(k_1k_2\cdots k_rk_{r+1}\cdots k_n) m_{1k_1} m_{2k_2} \cdots m_{rk_r} m_{r+1k_{r+1}} \cdots m_{nk_n}$$

レヴィ・チビタ記号の添え字が長くなるので括弧にしています。添え字の k_i は、k_r までは 1 から r、k_{r+1} からは $r+1$ から n なので、$\epsilon(k_1k_2\cdots k_rk_{r+1}\cdots k_n)$ を左から $1, 2, \ldots, n$ と並び変えたとき $k_1k_2\cdots k_r$ と $k_{r+1}\cdots k_n$ は混ざりません。このため、レヴィ・チビタ記号を

$$\epsilon(k_1k_2\cdots k_rk_{r+1}\cdots k_n) = \epsilon(k_1k_2\cdots k_r)\epsilon(k_{r+1}\cdots k_n)$$

と分離して書けます。そうすると

$$\det M = \sum_{k_1,\cdots,k_r=1}^{r} \epsilon(k_1k_2\cdots k_r) m_{1k_1} m_{2k_2} \cdots m_{rk_r} \sum_{k_{r+1},\cdots,k_n=r+1}^{n} \epsilon(k_{r+1}\cdots k_n) m_{r+1k_{r+1}} \cdots m_{nk_n}$$

左部分は A の行列式、右部分は B の行列式なので

$$\det M = \det A \det B$$

となります。同様にすることで

$$M = \begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$$

に対しても、$\det M = \det A \det B$ が示せます。

ただし、一般的には

$$M = \begin{pmatrix} A & C \\ D & B \end{pmatrix} \neq \det A \det B - \det C \det D$$

なので、勘違いしないように注意してください。M は、A が $m \times m$ 正則行列、B が $n \times n$ 行列なら、I_k を $k \times k$ 単位行列として

$$M = \begin{pmatrix} I_m & 0 \\ DA^{-1} & I_n \end{pmatrix} \begin{pmatrix} A & C \\ 0 & B - DA^{-1}C \end{pmatrix}$$

と書けるので、一般的に

$$\det M = \det A \det[B - DA^{-1}C]$$

となります。

また、$A = B = X,\ C = D = Y$ のとき、行列式は行もしくは列を加える基本操作で変わらないので

$$M = \begin{pmatrix} X & Y \\ Y & X \end{pmatrix} \Rightarrow \begin{pmatrix} X+Y & Y+X \\ Y & X \end{pmatrix} \Rightarrow \begin{pmatrix} X+Y & 0 \\ Y & X-Y \end{pmatrix}$$

と変形することで

$$\det M = \det[X+Y]\det[X-Y]$$

となります。

ブロック行列の話は終わりにして、必要になる単語の定義をします。行列が

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} , \begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{pmatrix}$$

となっているとき三角行列 (triangular matrix) と呼び、左を上三角行列 (upper triangular matrix)、下三角行列 (lower triangular matrix) と言います。上三角行列の成分 a_{ij} は $i > j$ では 0、下三角行列の成分 a_{ij} は $i < j$ では 0 です。

行列の相似を定義します。$n \times n$ 行列 A, B は、正則行列 P によって

$$B = P^{-1}AP$$

となっているとします。この変換を相似変換 (similar transformation) と呼び、A と B は相似 (similar) と呼ばれます。相似のとき同じ固有方程式になり、固有値は同じになります。B の固有値を λ とすれば、固有方程式は

$$\det[B - \lambda I] = 0$$

I は単位行列です。左辺を変形させると

$$\begin{aligned} \det[B - \lambda I] = \det[P^{-1}AP - \lambda I] &= \det[P^{-1}AP - \lambda P^{-1}IP] \\ &= \det[P^{-1}(A - \lambda I)P] \\ &= \det[P^{-1}]\det[A - \lambda I]\det[P] \\ &= \det[A - \lambda I]\det[P^{-1}P] \\ &= \det[A - \lambda I] \end{aligned}$$

このように A, B で固有方程式は同じになり、固有値も同じになります。$\lambda=0$ にすれば $\det B=\det A$ なので、相似なら行列式は同じです。トレースも、$\mathrm{tr}[XY]=\mathrm{tr}[YX]$ から、

$$\mathrm{tr}B=\mathrm{tr}[P^{-1}AP]=\mathrm{tr}[(P^{-1}A)P]=\mathrm{tr}[PP^{-1}A]=\mathrm{tr}A$$

となり、同じです。また、固有値 λ に対応する B の固有ベクトルを $\boldsymbol{x}$ とすると、$B\boldsymbol{x}=P^{-1}AP\boldsymbol{x}$ から

$$\begin{aligned}AP\boldsymbol{x}&=PB\boldsymbol{x}\\&=\lambda P\boldsymbol{x}\end{aligned}$$

となるので、$P\boldsymbol{x}$ は固有値 λ の A の固有ベクトルです。

相似変換によって、$n\times n$ 行列 A は対角成分が A の固有値となる三角行列にできることを示します。まずは A を 2×2 行列とします。A の固有値を λ_1, λ_2、その固有ベクトルを $\boldsymbol{x}_1, \boldsymbol{x}_2$ とします。$\boldsymbol{x}_1$ を 1 列目に持つ行列 R_1 を

$$R_1=(\boldsymbol{x}_1\ \boldsymbol{r}_2)=\begin{pmatrix}x_{11} & r_{12}\\ x_{21} & r_{22}\end{pmatrix}$$

とします。$\boldsymbol{x}_1, \boldsymbol{r}_i$ は 2×1 行列です。行列の積の規則 (ブロック行列の規則) から

$$\begin{aligned}AR_1=A(\boldsymbol{x}_1\ \boldsymbol{r}_2)&=\begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}\begin{pmatrix}x_{11} & r_{12}\\ x_{21} & r_{22}\end{pmatrix}\\&=\begin{pmatrix}a_{11}x_{11}+a_{12}x_{21} & a_{11}r_{12}+a_{12}r_{22}\\ a_{21}x_{11}+a_{22}x_{21} & a_{21}r_{12}+a_{22}r_{22}\end{pmatrix}\\&=\left(A\begin{pmatrix}x_{11}\\ x_{21}\end{pmatrix}\ A\begin{pmatrix}r_{12}\\ r_{22}\end{pmatrix}\right)\\&=(A\boldsymbol{x}_1\ A\boldsymbol{r}_2)\end{aligned}$$

これに、R_1^{-1} をかけると

$$R_1^{-1}AR_1=(R_1^{-1}A\boldsymbol{x}_1\ \ R_1^{-1}A\boldsymbol{r}_2)$$

1 列目の 2×1 行列 $R_1^{-1}A\boldsymbol{x}_1$ は $A\boldsymbol{x}_1=\lambda_1\boldsymbol{x}_1$ から

$$R_1^{-1}A\boldsymbol{x}_1=R_1^{-1}\lambda_1\boldsymbol{x}_1$$

$R_1^{-1}\boldsymbol{x}_1$ は $\boldsymbol{x}_1$ が R_1 の 1 列目なので

$$R_1^{-1}\boldsymbol{x}_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_{11} + \alpha_{12}x_{21} \\ \alpha_{21}x_{11} + \alpha_{22}x_{21} \end{pmatrix}$$

$$\Leftrightarrow \ R_1^{-1}R_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} x_{11} & r_{12} \\ x_{21} & r_{22} \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_{11} + \alpha_{12}x_{21} & \alpha_{11}r_{12} + \alpha_{12}r_{22} \\ \alpha_{21}x_{11} + \alpha_{22}x_{21} & \alpha_{21}r_{12} + \alpha_{22}r_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

と対応しているので、2×1 行列 $R_1^{-1}\boldsymbol{x}_1$ の 1 行目と 2 行目は

$$(R_1^{-1}\boldsymbol{x}_1)_{11} = 1 \ , \ (R_1^{-1}\boldsymbol{x}_1)_{21} = 0$$

これは、成分で書けば

$$(R_1^{-1}\boldsymbol{x}_1)_{11} = \sum_{k=1}^{2}(R_1^{-1})_{1k}(R_1)_{k1}$$

$$(R_1^{-1}\boldsymbol{x}_1)_{21} = \sum_{k=1}^{2}(R_1^{-1})_{2k}(R_1)_{k1}$$

$$(R_1^{-1}R_1)_{ij} = \sum_{k=1}^{2}(R_1^{-1})_{ik}(R_1)_{kj} = \delta_{ij} \tag{1}$$

となっているためです。δ_{ij} はクロネッカーデルタです。というわけで、

$$R_1^{-1}AR_1 = \lambda_1 R_1^{-1}\boldsymbol{x}_1 = \begin{pmatrix} \lambda_1 & \beta_1 \\ 0 & \beta_2 \end{pmatrix}$$

これの固有方程式を作るために

$$R_1^{-1}AR_1 - \lambda I_2 = \begin{pmatrix} \lambda_1 - \lambda & \beta_1 \\ 0 & \beta_2 - \lambda \end{pmatrix}$$

1 列目での余因子展開から (ブロック行列の行列式から)

$$\det[R_1^{-1}AR_1 - \lambda I_2] = (\lambda_1 - \lambda)\det[\beta_2 - \lambda]$$

$\beta_2 - \lambda$ は行列ではないですが後のために行列式のままにしています。このため、A の固有方程式とは

$$\det[A - \lambda I_2] = \det[R_1^{-1}AR_1 - \lambda I_2] = \det[\beta_2 - \lambda] = 0$$

となり、β_2 は λ_1 でない A の固有値になる必要があり、$\beta_2 = \lambda_2$ です。よって、$R_1^{-1}AR_1$ は対角成分が A の固有値の上三角行列となります。

同様のことが 3×3 行列でも行えます。実際に、A, R_1 は 3×3 行列とすれば

$$R_1^{-1}AR_1 = \begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}$$

C は 1×2 行列、B は 2×2 行列になるので、2×2 正則行列 R_2 による

$$R'_2 = \begin{pmatrix} 1 & 0 \\ 0 & R_2 \end{pmatrix}\ ,\ R'^{-1}_2 = \begin{pmatrix} 1 & 0 \\ 0 & R_2^{-1} \end{pmatrix}$$

によって

$$\begin{aligned} R'^{-1}_2 R_1^{-1}AR_1R'_2 &= \begin{pmatrix} 1 & 0 \\ 0 & R_2^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & R_2 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & R_2^{-1} \end{pmatrix}\begin{pmatrix} \lambda_1 & CR_2 \\ 0 & BR_2 \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 & CR_2 \\ 0 & R_2^{-1}BR_2 \end{pmatrix} \end{aligned}$$

このため、$R_2^{-1}BR_2$ で $R_1^{-1}AR_1$ と同様にすることで

$$R'^{-1}_2 R_1^{-1}AR_1R'_2 = \begin{pmatrix} \lambda_1 & \beta_1 & \beta_2 \\ 0 & \lambda_2 & \beta_3 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

このように行列の成分が増えても同じ手順の繰り返しになるので、$n\times n$ 行列で成立します。

簡単に言っておきます。行列の積の規則は変更されないので、単純に $n\times n$ 行列に拡張するだけです。A の固有値を λ_i、その固有ベクトルを $\boldsymbol{x}_i$ $(i=1,2,\dots,n)$ とします。$n\times 1$ 行列 $\boldsymbol{x}_1, \boldsymbol{r}_i$ によって $n\times n$ 正則行列 R_1 を

$$R_1 = (\boldsymbol{x}_1\ \boldsymbol{r}_2\ \dots\ \boldsymbol{r}_n)\ ,\ AR_1 = (A\boldsymbol{x}_1\ A\boldsymbol{r}_2\ \dots\ A\boldsymbol{r}_n)$$

$R_1^{-1}AR_1$ は、(1) での k の範囲を n に変更するだけなので

$$R_1^{-1}AR_1 = \begin{pmatrix} \lambda_1 & C \\ 0 & B \end{pmatrix}$$

B は $(n-1)\times(n-1)$ 行列、C は $1\times(n-1)$ 行列です。$R_1^{-1}AR_1 - \lambda I_n$ の行列式は

$$\det[R_1^{-1}AR_1 - \lambda I_n] = (\lambda_1 - \lambda)\det[B - \lambda I_{n-1}]$$

I_n は $n\times n$ 単位行列です。これから、B は固有値 $\lambda_2, \lambda_3, \dots, \lambda_n$ を持ちます。

次に、B の固有値 λ_2 に対応する固有ベクトル $\boldsymbol{x}_2$ を 1 列目に持つ R_2 を作って、$R_2^{-1}BR_2$ を同様に求めれば

$$R_2^{-1}BR_2 = \begin{pmatrix} \lambda_2 & C' \\ 0 & B' \end{pmatrix}$$

となり、同じことの繰り返しになります。そして、変換行列 R は

$$R = R_1 \begin{pmatrix} I_1 & 0 \\ 0 & R_2 \end{pmatrix} \begin{pmatrix} I_2 & 0 \\ 0 & R_3 \end{pmatrix} \cdots \begin{pmatrix} I_{n-2} & 0 \\ 0 & R_{n-1} \end{pmatrix}$$

として作れます。I_1 は 1 で、R_1 は $n\times n$ 行列、R_2 は $(n-1)\times(n-1)$ 行列となっています。このようにして $n\times n$ 行列で成立します。

変換行列 R を制限できます。そのために、R_1 を作るときに A の固有ベクトルとしかしませんでしたが、これに制限を加えます。まず、固有ベクトル $\boldsymbol{x}_1$ を含めて線形独立なベクトルの組 $(\boldsymbol{x}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$ を用意します。この組からグラム・シュミットの正規直交化法によって、$(\boldsymbol{u}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)$ として正規直交系を作れます。実数に制限していないので、ベクトルの大きさ $|\boldsymbol{v}|$ は

$$|\boldsymbol{v}| = \boldsymbol{v}^\dagger \cdot \boldsymbol{v} = (\boldsymbol{v}^t)^* \cdot \boldsymbol{v}$$

内積は行列計算に合わせています。「$\dagger$」はエルミート共役、「$*$」は複素共役、t は転置です。直交性は

$$|\boldsymbol{u}_1| = 1\ ,\ |\boldsymbol{v}_j| = 1$$

$$\boldsymbol{u}_1^\dagger \cdot \boldsymbol{v}_j = 0\ ,\ \boldsymbol{v}_j^\dagger \cdot \boldsymbol{v}_k = 0 \quad (j \neq k)$$

となります。

$\boldsymbol{u}_1$ は A の固有ベクトルなので、$A\boldsymbol{u}_1 = \lambda_1 \boldsymbol{u}_1$ として固有値 λ_1 を与えます。このときの R_1 は $\boldsymbol{u}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を列として

$$R_1 = (\boldsymbol{u}_1\ \boldsymbol{v}_2\ \ldots\ \boldsymbol{v}_n)$$

R_1 の逆行列は、転置は行と列を入れ替えるために R_1 の転置は $\boldsymbol{u}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ を行とする行列になることと直交性から

$$R_1^{-1} = \begin{pmatrix} \boldsymbol{u}_1^\dagger \\ \boldsymbol{v}_2^\dagger \\ \vdots \\ \boldsymbol{v}_n^\dagger \end{pmatrix} = R_1^\dagger$$

と分かり、R_1 はユニタリー行列です。

R_2 がユニタリー行列のとき R_2' は

$$R_2' = \begin{pmatrix} I_1 & 0 \\ 0 & R_2 \end{pmatrix}$$

これのエルミート共役を取ると、転置は行と列の入れ替えなので、R_2 部分だけで転置は起こり

$$R_2'^{\dagger} = \begin{pmatrix} I_1 & 0 \\ 0 & R_2^{\dagger} \end{pmatrix}$$

そうすると、

$$R_2'^{\dagger} R_2' = \begin{pmatrix} I_1 & 0 \\ 0 & R_2^{\dagger} \end{pmatrix} \begin{pmatrix} I_1 & 0 \\ 0 & R_2 \end{pmatrix} = \begin{pmatrix} I_1 & 0 \\ 0 & R_2^{\dagger} R_2 \end{pmatrix} = \begin{pmatrix} I_1 & 0 \\ 0 & I_{n-1} \end{pmatrix}$$

となるので、単位行列をくっつけた R_2' もユニタリー行列です。そして、U, V をユニタリー行列とすれば

$$(UV)^{\dagger} UV = V^{\dagger} U^{\dagger} UV = V^{-1} U^{-1} UV = I$$

から、ユニタリー行列同士の積はユニタリー行列と分かります。なので、R_1, R_2' の積もユニタリー行列です。このようにユニタリー行列になるだけで、話は同じです。

というわけで、$n \times n$ 行列 A はユニタリー行列 U の相似変換 $U^{-1}AU = U^{\dagger}AU$ によって、$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$ $(i = 1, 2, \ldots, n)$ の固有値を対角成分に持つ上三角行列にできます (相似変換で上三角行列にできるユニタリー行列が存在する)。これを Schur 分解 (Schur decomposition) や Schur 三角化 (Schur triangulation) と言い、これによる上三角行列は Schur 標準形と呼ばれます。分解は行列を行列同士の積の形にすることを指します。また、行列が実数に制限され、固有値が全て実数なら複素共役が外れるので、ユニタリー行列でなく直交行列になります。