

Research Article

Christopher S. Withers and Saralees Nadarajah*

The Jordan form of a triangular matrix

<https://doi.org/10.1515/spma-2025-0042>

received April 10, 2025; accepted September 25, 2025

Abstract: We show how to obtain analytic expressions for the Jordan form of an upper triangular matrix, including that of the standardizing matrix. For example, the Jordan form of a triangular banded matrix (a banded matrix is a matrix where nonzero entries are confined to a band along the main diagonal) is given in terms of Bell polynomials.

Keywords: Bell polynomials, Jordan form, triangular matrix

MSC 2020: 11C99

1 Introduction

This short note gives original results on the Jordan form for triangular matrices. To the best of our knowledge, there has been little or no work on this important area. Triangular matrices arise in many theoretical and applied subjects. Some recent applications have included: co-channel interference rejection in MIMO systems [3]; Cholesky-GARCH models with applications to finance [2]; underwater target detection [5].

The Jordan form of a square matrix is a fundamental tool for solving a variety of matrix problems, such as the behavior of large powers of a matrix [6], solving of Fredholm integral equations with nonsymmetric kernels [7], and many classes of linear and nonlinear matrix recurrence formulas. It is defined in terms of

$$J_m(\lambda) = \lambda I_m + U_m,$$

where I_m is the $m \times m$ identity matrix, and U_m is the shift matrix, the $m \times m$ matrix of zeros except for ones on its first superdiagonal:

$$(U_m)_{i,j} = \delta_{i+1,j} \quad \text{for } 1 \leq i, j \leq m,$$

and $\delta_{j,k}$ is the Kronecker delta, equal to 1 or 0 for $j = k$ or $j \neq k$. For example,

$$U_1 = 0, \quad U_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad U_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad U_4 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

So,

$$U^m = 0 \in \mathbb{C}^{m \times m}, \quad (U_m^k)_{i,j} = \delta_{i+k,j}$$

for $1 \leq i, j \leq m$, and $0 \leq k < m$, so that U_m^k has zeros on its first $k - 1$ superdiagonals. LU^k is L with its columns moved right by k columns. That is, for $L = (l_1, \dots, l_r) \in \mathbb{C}^{r \times r}$,

$$LU^k = (0_r, 0_r, \dots, 0_r, l_1, l_2, \dots, l_{r-k}), \quad k = 1, 2, \dots, r - 1.$$

* **Corresponding author: Saralees Nadarajah**, Department of Mathematics, University of Manchester, Manchester M13 9PL, United Kingdom, e-mail: mbbssn2@manchester.ac.uk

Christopher S. Withers: Callaghan Innovation, Lower Hutt 5040, New Zealand, e-mail: kit.withers@gmail.com

The Jordan form of a matrix $\mathbf{A} \in \mathbb{C}^{r \times r}$ is the decomposition

$$\mathbf{A} = \mathbf{L}\mathbf{\Lambda}\mathbf{R}^*, \quad (1)$$

where

$$\begin{aligned} \mathbf{L}\mathbf{R}^* &= \mathbf{I}_r, \quad \mathbf{L}, \mathbf{R}^* \in \mathbb{C}^{r \times r}, \\ \mathbf{\Lambda} &= \text{diag}(\mathbf{\Lambda}_1, \dots, \mathbf{\Lambda}_s), \quad \mathbf{\Lambda}_j = \mathbf{J}_{m_j}(\lambda_j), \end{aligned}$$

where λ_j are the eigenvalues of $\mathbf{A}$ repeated m_j times, for $j = 1, \dots, s$. Also

$$\mathbf{A}\mathbf{L} = \mathbf{L}\mathbf{\Lambda}, \quad \mathbf{\Lambda} = \text{diag}(\mathbf{\Lambda}_1, \dots, \mathbf{\Lambda}_s). \quad (2)$$

So, writing $\mathbf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_s)$ and $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_s)$, where $\mathbf{L}_i, \mathbf{R}_i \in \mathbb{C}^{r \times m_i}$, gives

$$\mathbf{A} = \mathbf{L}\mathbf{\Lambda}\mathbf{R}^* = \sum_{i=1}^s \mathbf{L}_i \mathbf{\Lambda}_i \mathbf{R}_i^*, \quad (3)$$

$$\begin{aligned} \mathbf{I}_r &= \sum_{i=1}^s \mathbf{L}_i \mathbf{R}_i^*, \quad \mathbf{I}_r = \mathbf{R}^* \mathbf{L} = (\mathbf{R}_i^* \mathbf{L}_j), \quad \mathbf{R}_i^* \mathbf{L}_j = \delta_{i,j} \mathbf{I}_{m_i}, \\ \mathbf{A}\mathbf{L}_i &= \mathbf{L}_i \mathbf{\Lambda}_i, \quad (\mathbf{A} - \lambda_i \mathbf{I}_r) \mathbf{L}_i = \mathbf{L}_i \mathbf{U}_{m_i}, \quad i = 1, \dots, s. \end{aligned} \quad (4)$$

Writing

$$M_0 = 0, \quad M_i = \sum_{j=1}^i m_j, \quad \mathbf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_r), \quad \mathbf{L}_i = (\mathbf{L}_{M_{i-1}+1}, \dots, \mathbf{L}_{M_i}),$$

the k th column in (4) is

$$(\mathbf{A} - \lambda_i \mathbf{I}_r) \mathbf{L}_{M_{i-1}+k} = (1 - \delta_{k,1}) \mathbf{L}_{M_{i-1}+k-1}, \quad k = 1, \dots, m_i, \quad i = 1, \dots, s.$$

This is *the Jordan chain*. $\mathbf{L}$ is not unique and there is at present no convention or canonical choice for it. Note that $\{\mathbf{\Lambda}_i, \mathbf{L}_i, \mathbf{R}_i^*\}$ can be put in any order, for example with m_i increasing, or decreasing as done by MAPLE, or with $|\lambda_i|$ increasing. For a given ordering, $\mathbf{L}_k$, the k th column of $\mathbf{L}$, is not unique as it can be multiplied by any nonzero constant.

Section 2 gives the Jordan form of a triangular matrix when diagonal elements are either all equal, $s = 1$, or all distinct, $s = r$, and also when the number of Jordan blocks is $s = 2$ or $r - 1$. The method can be extended to general s . A corollary gives the Jordan form of a triangular banded matrix (a matrix where nonzero entries are confined to bands on one side of the diagonal) in terms of Bell polynomials.

Section 3 gives some examples of triangular banded matrices. Section 4 considers the question: given $\mathbf{V} \in \mathbb{C}^{r \times r}$, for what class of matrices does $\{\mathbf{V}^k\}$ form a basis?

We have not found any results on $\mathbf{L}$ in the literature. A result that overlaps a little with ours, but without discussing $\mathbf{L}$, is Theorem 6.25, page 424 of Horn and Johnson [4].

2 The general upper triangular matrix

Let $\mathbf{A}$ be any upper triangular matrix. That is, $A_{ij} = 0$ for $i > j$ so that $\{A_{i,i}\}$ are the eigenvalues of $\mathbf{A}$, and $\det(\mathbf{A}) = \prod_{i=1}^r A_{i,i}$. Its Jordan form depends on the number of equal eigenvalues. In this section, we give solutions for when these are all equal and for when these are all distinct, and more generally when the number of Jordan blocks, s , is 1 or 2 or $r - 1$ or r . This is enough to show how the method can be applied to the general case, as other cases can be dealt with similarly. Let $\mathbf{l}_i$ be the i th column of $\mathbf{L}$: $\mathbf{L} = (\mathbf{l}_1, \dots, \mathbf{l}_r)$.

The case $s = 1$: We begin with the case $A_{i,i} \equiv a_0$. Our first result gives necessary and sufficient conditions that $s = 1$, that is, there is only one Jordan block, $\mathbf{\Lambda} = \mathbf{J}_r(a_0)$.

Theorem 2.1. Suppose that $\mathbf{A} = a_0 \mathbf{I}_r + \Delta$, where Δ is any strictly upper triangular $r \times r$ matrix with no zero on its first superdiagonal, that is,

$$\prod_{i=1}^{r-1} A_{i,i+1}^i \neq 0.$$

Then $\mathbf{A}$ has Jordan form (1) with $s = 1$, $m_r = r$, $\lambda_1 = a_0$,

$$\mathbf{L} = (\mathbf{l}_1, \dots, \mathbf{l}_r), \quad (5)$$

where $\mathbf{l}_j = \Delta^{r-j} \mathbf{l}_r$, $j = 1, \dots, r-1$, and $\mathbf{l}_r$ is any r -vector with $L_{r,r} \neq 0$. So, $\mathbf{L}$ is given in terms of $\mathbf{l}_r$ by

$$L_{i,j} = (\Delta^{r-j} \mathbf{l}_r)_i = \sum_{k=i+r-j}^r (\Delta^{r-j})_{i,k} L_{k,r}, \quad (6)$$

$\mathbf{L}$ is upper triangular,

$$\begin{aligned} L_{i,i} &= (\Delta^{r-i})_{i,r} L_{r,r} = A_{i,i+1} \cdots A_{r-1,r} L_{r,r}, \\ \det(\mathbf{L}) &= L_{r,r}^r \prod_{i=1}^{r-1} A_{i,i+1}^i \neq 0, \end{aligned}$$

and $\mathbf{A} = \mathbf{L} \mathbf{A} \mathbf{L}^{-1}$ is the Jordan form for $\mathbf{A}$. So, if $\mathbf{l}_r = \mathbf{e}_{r,r} = (0, \dots, 0, 1)'$, the r th unit vector in $\mathbb{C}^{r \times r}$, then

$$\begin{aligned} L_{i,j} &= (1 - \delta_{i,j}) (\Delta^{r-j})_{i,r}, \\ L_{i,i} &= (\Delta^{r-i})_{i,r} = A_{i,i+1} \cdots A_{r-1,r}, \\ \det(\mathbf{L}) &= \prod_{i=1}^{r-1} A_{i,i+1}^i. \end{aligned} \quad (7)$$

Proof. Let $\mathbf{0}_r$ be the r -vector of zeros. Then

$$\begin{aligned} (a_0 \mathbf{I}_r + \Delta) \mathbf{L} &= \mathbf{A} \mathbf{L} = \mathbf{L} (a_0 \mathbf{I}_r + \mathbf{U}_r), \quad \Delta \mathbf{L} = \mathbf{L} \mathbf{U}_r = (\mathbf{0}_r, \mathbf{l}_1, \dots, \mathbf{l}_{r-1}), \\ \mathbf{l}_{j-1} &= \Delta \mathbf{l}_j \quad \text{for } j \geq 1. \end{aligned}$$

So, (5) holds and the rest follows, since if Δ is strictly upper triangular, then

$$(\Delta^k)_{i,j} = 0 \text{ unless } i \leq j - k. \quad \text{Also, } (\Delta^{r-i})_{i,r} = A_{i,i+1} \cdots A_{r-1,r}. \quad (8)$$

The proof is complete. $\square$

We call $\mathbf{L}$ of (7) the canonical choice of $\mathbf{L}$. If just one $A_{i,i+1}$ is zero, then the Jordan form of $\mathbf{A}$ is given under the case $s = 2$ below. We now apply this to the triangular banded matrix. Given a sequence $\mathbf{a} = (a_1, a_2, \dots)$ of complex numbers, the ordinary partial Bell polynomials, $\widehat{B}_{k,i} = \widehat{B}_{k,i}(\mathbf{a})$, are defined by

$$\left(\sum_{k=1}^{\infty} a_k x^k \right)^i = \sum_{k=i}^{\infty} \widehat{B}_{k,i} x^k$$

for $x \in \mathbb{C}$ and $i \geq 0$. So, $\widehat{B}_{0,0} = 1$, $\widehat{B}_{k,i} = 0$ if $k < i$, or if $k = i \neq 0$. The others are tabled on page 309 of the study by Comtet [1].

Corollary 2.1. Suppose that $\mathbf{A} \in \mathbb{C}^{r \times r}$ has the form

$$\mathbf{A} = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots \\ 0 & a_0 & a_1 & \cdots \\ 0 & 0 & \cdots & \\ 0 & \cdots & 0 & a_0 \end{pmatrix} = \sum_{k=0}^{r-1} a_k \mathbf{U}_r^k. \quad (9)$$

Set $\Delta = \mathbf{A} - a_0 \mathbf{I}_r$. Then

$$\Delta^i = \sum_{k=i}^{r-1} \widehat{B}_{k,i} \mathbf{U}_r^k = \begin{pmatrix} 0 & \widehat{B}_{1,i} & \widehat{B}_{2,i} & \cdots & \widehat{B}_{r-2,i} & \widehat{B}_{r-1,i} \\ 0 & 0 & \widehat{B}_{1,i} & \cdots & \widehat{B}_{r-3,i} & \widehat{B}_{r-2,i} \\ 0 & 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \widehat{B}_{1,i} \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

If $a_1 \neq 0$, $\mathbf{A}$ has Jordan form

$$\mathbf{A} = \mathbf{L}(a_0 \mathbf{I}_r + \mathbf{U}_r) \mathbf{L}^{-1},$$

where $\mathbf{L} = (\mathbf{l}_1, \dots, \mathbf{l}_r)$ and $\mathbf{l}_i$ is given in terms of $\mathbf{l}_r$ by

$$L_{i,j} = \sum_{k=r-j}^{r-1} \widehat{B}_{k,r-j} L_{i+k,r}.$$

So, if $\mathbf{l}_r = \mathbf{e}_{r,r} = (0, \dots, 0, 1)'$, then $L_{i,j} = \widehat{B}_{r-i,r-j}$;

$$L_{j,j} = a_1^{r-j}, \quad \det(\mathbf{L}) = \prod_{j=1}^r L_{j,j} \neq 0.$$

That is, the canonical choice of $\mathbf{L}$ is

$$\mathbf{L} = \begin{pmatrix} \widehat{B}_{r-1,r-1} & \widehat{B}_{r-1,r-2} & \cdots & \widehat{B}_{r-1,1} & \widehat{B}_{r-1,0} \\ 0 & \widehat{B}_{r-2,r-2} & \cdots & \widehat{B}_{r-2,1} & \widehat{B}_{r-2,0} \\ 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \widehat{B}_{1,1} & \widehat{B}_{1,0} \\ 0 & 0 & \cdots & 0 & \widehat{B}_{0,0} \end{pmatrix}.$$

For example, $\mathbf{l}_{r-1} = (a_{r-1}, a_{r-2}, \dots, a_1, 0)'$.

The case $s = 2$: Suppose that $A_{i,i} = a_i$, where $i = 1$ for $1 \leq i \leq m_1$ and $i = 2$ for $m_1 < i \leq r$. Then $\mathbf{A} = \text{diag}(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$, where $\mathbf{\Lambda}_i = \mathbf{J}_{m_i}(a_i)$. Write $\mathbf{A}$, $\mathbf{L}$ as 2×2 block matrices with (i, j) elements $\overline{\mathbf{A}}_{i,j}$, $\overline{\mathbf{L}}_{i,j}$, $i, j = 1, 2$. Since $\mathbf{A}$ is upper triangular, we can choose $\mathbf{L}$ to be upper triangular, so that $\overline{\mathbf{A}}_{2,1} = \overline{\mathbf{L}}_{2,1} = \mathbf{0}$. By (2), $\mathbf{A}\mathbf{L} = \mathbf{L} \text{diag}(\mathbf{\Lambda}_1, \mathbf{\Lambda}_2)$. Taking the (i, j) element for $i, j = 1, 2$ gives

$$\overline{\mathbf{A}}_{i,i} \overline{\mathbf{L}}_{i,i} = \overline{\mathbf{L}}_{i,i} \mathbf{\Lambda}_i, \quad i = 1, 2, \quad (10)$$

$$\overline{\mathbf{A}}_{1,1} \overline{\mathbf{L}}_{1,2} + \overline{\mathbf{A}}_{1,2} \overline{\mathbf{L}}_{2,2} = \overline{\mathbf{L}}_{1,2} \mathbf{\Lambda}_2 = \overline{\mathbf{L}}_{1,2} (a_2 \mathbf{I}_{m_2} + \mathbf{U}_{m_2}). \quad (11)$$

The solution of (10) for $\overline{\mathbf{L}}_{i,i}$ is just that given for $s = 1$ above with $\mathbf{A}$, r , a_0 , $\mathbf{L}$, Δ replaced by $\overline{\mathbf{A}}_{i,i}$, m_i , a_i , $\overline{\mathbf{L}}_{i,i}$, $\mathbf{\Lambda}_i$, where $\mathbf{\Lambda}_i = \overline{\mathbf{A}}_{i,i} - a_i \mathbf{I}_{m_i}$. So, it requires that the elements of the first superdiagonal of $\overline{\mathbf{A}}_{i,i}$ are nonzero, that is,

$$\prod \{ \mathbf{A}_{i,i+1} : 1 \leq i < r, \quad i \neq m_1 \} \neq \mathbf{0}. \quad (12)$$

Given $\overline{\mathbf{L}}_{2,2}$, we now solve (11) for $\overline{\mathbf{L}}_{1,2}$. Write it as

$$(a_1 \mathbf{I}_{m_1} + \mathbf{\Lambda}_{1,1}) \overline{\mathbf{L}}_{1,2} + \mathbf{B} = \overline{\mathbf{L}}_{1,2} (a_2 \mathbf{I}_{m_2} + \mathbf{U}_{m_2}),$$

where $\mathbf{B} = \overline{\mathbf{A}}_{1,2} \overline{\mathbf{L}}_{2,2}$ is $m_1 \times m_2$. That is, $\mathbf{J} \overline{\mathbf{L}}_{1,2} + \mathbf{B} = \overline{\mathbf{L}}_{1,2} \mathbf{U}_{m_2}$, where $\mathbf{J} = (a_1 - a_2) \mathbf{I}_{m_1} + \mathbf{\Lambda}_{1,1}$ and $m = m_1$. Taking the i th column gives

$$\mathbf{b}_i + \mathbf{J} \mathbf{n}_i = \mathbf{n}_{i-1}, \quad i = 1, \dots, m_2,$$

where $\mathbf{b}_i$ and $\mathbf{n}_i$ are the i th columns of $\mathbf{B}$ and $\overline{\mathbf{L}}_{1,2}$, respectively. Its solution in terms of $\mathbf{n}_{m_2}$ is

$$\mathbf{n}_{m_2-i} = \sum_{j=0}^{i-1} \mathbf{J}^j \mathbf{b}_{m_2-i+j+1} + \mathbf{J}^i \mathbf{n}_{m_2}, \quad i = 1, \dots, m - 1.$$

Also by (12), $\det(\mathbf{L}) = \prod_{i=1}^2 \det(\overline{\mathbf{L}}_{i,i}) \neq 0$ so that $\mathbf{A} = \mathbf{L} \mathbf{A} \mathbf{L}^{-1}$ is the Jordan form for $\mathbf{A}$.

Example 2.1. Suppose that $s = r = 3$. Then we can take

$$\mathbf{L} = \begin{pmatrix} 1 & \mathbf{A}_{1,2} & \mathbf{L}_{1,3} \\ 0 & \mathbf{A}_{2,2} - \mathbf{A}_{1,1} & \mathbf{L}_{2,3} \\ 0 & 0 & \mathbf{A}_{3,3} - \mathbf{A}_{1,1} \end{pmatrix},$$

where

$$\mathbf{L}_{1,3} = \mathbf{A}_{1,2}\mathbf{A}_{2,3} + \mathbf{A}_{1,3}(\mathbf{A}_{3,3} - \mathbf{A}_{2,2}), \quad \mathbf{L}_{2,3} = \mathbf{A}_{2,3}(\mathbf{A}_{3,3} - \mathbf{A}_{1,1}).$$

The case $s = r$: This holds if $\{\mathbf{A}_{j,j}\}$ are all distinct. Then $m_j \equiv 1$, $\mathbf{A}_j = \mathbf{A}_{j,j}$, $\mathbf{A} = \text{diag}(\mathbf{A}_{1,1}, \dots, \mathbf{A}_{r,r})$ and $\mathbf{L}$ satisfies $\mathbf{L}\mathbf{A} = \mathbf{A}\mathbf{L}$. So, $\mathbf{L}$ is upper triangular and the (i, j) element of $\mathbf{L}\mathbf{A} = \mathbf{A}\mathbf{L}$ gives

$$\mathbf{L}_{i,j} = (\mathbf{A}_{j,j} - \mathbf{A}_{i,i})^{-1} \sum_{k=i+1}^j \mathbf{A}_{i,k}(\mathbf{L}_{k,j} - \mathbf{L}_{k-1,j-1}), \quad j = i+1, \dots, r.$$

In particular,

$$\begin{aligned} \mathbf{L}_{j-1,j} &= (\mathbf{A}_{j,j} - \mathbf{A}_{j-1,j-1})^{-1} \mathbf{A}_{j-1,j}(\mathbf{L}_{j,j} - \mathbf{L}_{j-1,j-1}), \\ \mathbf{L}_{j-2,j} &= (\mathbf{A}_{j,j} - \mathbf{A}_{j-2,j-2})^{-1} \sum_{k=j-1}^j \mathbf{A}_{j-2,k}(\mathbf{L}_{k,j} - \mathbf{L}_{k-1,j-1}), \end{aligned}$$

and so on, giving $\mathbf{L}_j$ in terms of $\mathbf{L}_{j,j}$, which is arbitrary but nonzero.

The case $s = r - 1$: Without loss of generality, we can order the blocks so that $m_i = 1$ for $1 \leq i \leq r - 2$ and $m_{r-1} = 2$. By (4), $\mathbf{L}_j = \mathbf{I}_j$, $\mathbf{A}\mathbf{L}_j = \mathbf{A}_{j,j}\mathbf{I}_j$ for $1 \leq j \leq r - 2$, and $\mathbf{L}_{r-1} = (\mathbf{I}_{r-1}, \mathbf{l}_r)$ satisfies $\mathbf{A}\mathbf{L}_{r-1} = (\mathbf{I}_{r-1}, \mathbf{l}_r)\mathbf{J}_2(\lambda_{r-1})$, where $\lambda_{r-1} = A_{r-1,r-1} = A_{r,r}$. That is,

$$\mathbf{A}\mathbf{L}_{r-1} = \lambda\mathbf{I}_{r-1}, \quad (\mathbf{A} - \lambda\mathbf{I}_r)\mathbf{l}_r = \mathbf{l}_{r-1}.$$

So, $\mathbf{l}_{r-1}$ is the eigenvector of $\mathbf{A}$ with eigenvalue λ_{r-1} , and $\mathbf{l}_{r-1} = (\mathbf{A} - A_{r,r}\mathbf{I}_r)\mathbf{l}_r$ has last element zero. So, $\mathbf{L}$ is upper triangular, and $\mathbf{l}_r$ is any vector with $L_{r,r} \neq 0$. The canonical choice $\mathbf{l}_r = \mathbf{e}_{r,r}$ implies $\mathbf{l}_{r-1} = (A_{1,r}, A_{2,r}, \dots, A_{r-1,r}, 0)'$.

3 Some triangular banded matrices

As in (1), we set $\mathbf{J}(\lambda) = \lambda\mathbf{I}_m + \mathbf{U}_m$.

Example 3.1. Suppose that $\theta \in \mathbb{C}$. By (3),

$$\mathbf{A}^\theta = \sum_{i=1}^s \mathbf{L}_i \mathbf{J}_{m_i}(\lambda_i)^\theta \mathbf{R}_i^*.$$

For $\lambda \neq 0$, $\mathbf{J}(\lambda)^\theta$ has the form (9) with

$$a_0 = \lambda^\theta, \quad a_k = a_0 \alpha_k \lambda^{-k},$$

where $\alpha_k = \binom{\theta}{k}$. So,

$$\widehat{B}_{k,i}(\mathbf{a}) = a_0^i \widehat{B}_{k,i}(\boldsymbol{\alpha}) \lambda^{-k}.$$

More generally, for $\mathbf{A}$ of Theorem 2.1, by (8),

$$\mathbf{A}^\theta = \sum_{k=0}^{r-1} \binom{r}{k} a_0^{\theta-k} \Delta^k$$

is upper triangular with (i, j) element

$$(\mathbf{A}^\theta)_{i,j} = \sum_{k=0}^{j-i} \binom{r}{k} a_0^{\theta-k} (\Delta^k)_{i,j}. \quad (13)$$

Example 3.2. By (3),

$$\exp(\mathbf{A}) = \sum_{i=1}^s \mathbf{L}_i \exp[\mathbf{J}_{m_i}(\lambda_i)] \mathbf{R}_i^*.$$

Note that $\exp[\mathbf{J}(\lambda)]$ has the form (9) with

$$a_0 = \exp(\lambda), \quad a_k = \frac{a_0}{k!}.$$

So,

$$\widehat{B}_{k,i}(\mathbf{a}) = i! a_0^i \frac{B_{k,i}(\mathbf{1})}{k!},$$

where $B_{k,i}(\mathbf{a})$ is the exponential partial Bell polynomial, tabled on pages 307–308 of [1]. But by page 135 in the study by Comtet [1], $B_{k,i}(\mathbf{1}) = S(i, k)$, the Stirling number of the second kind, tabled on pages 310–311 in the study by Comtet [1]. More generally, for $\mathbf{A}$ of Theorem 2.1, by (8),

$$\exp(\mathbf{A}) = \exp(a_0) \sum_{k=0}^{r-1} \frac{\Delta^k}{k!} \quad (14)$$

is upper triangular with (i, j) element

$$(\exp(\mathbf{A}))_{i,j} = \sum_{k=0}^{j-i} \frac{(\Delta^k)_{i,j}}{k!}.$$

Example 3.3. By (3),

$$\log \mathbf{A} = \sum_{i=1}^s \mathbf{L}_i (\log \mathbf{J}_{m_i}(\lambda_i)) \mathbf{R}_i^*.$$

For $\lambda \neq 0$, $\log[\mathbf{J}(\lambda)]$ has the form (9) with

$$a_0 = \log(\lambda), \quad a_k = (-1)^{k-1} \frac{\lambda^{-k}}{k}.$$

So,

$$\widehat{B}_{k,i}(\mathbf{a}) = (-1)^{i-k} i! a_0^i B_{k,i}(\mathbf{y}) \frac{\lambda^{-k}}{k!},$$

where $y_k = (k-1)!$. By page 135 in the study by Comtet [1], $B_{k,i}(\mathbf{y}) = |s(i, k)|$, where $s(n, k)$ is the Stirling number of the first kind, tabled on pages 310–311 in the study by Comtet [1]. More generally, for $\mathbf{A}$ of Theorem 2.1, by (8),

$$\log \mathbf{A} = \mathbf{I}_r \log a_0 + \sum_{k=1}^{r-1} (-1)^{k-1} \frac{\Delta^k}{k} \quad (15)$$

is upper triangular with (i, j) element

$$(\log \mathbf{A})_{i,j} = \log a_0 + \sum_{k=1}^{j-i} \frac{(-1)^{k-1}}{k} (\Delta^k)_{i,j}.$$

4 When is $\{\mathbf{V}^k\}$ a basis in $\mathbb{C}^{r \times r}$?

In this section, we address the following question: given matrices $\mathbf{A}, \mathbf{V} \in \mathbb{C}^{r \times r}$, when can we write

$$\mathbf{A} = \sum_{k=0}^{\infty} a_k \mathbf{V}^k, \quad (16)$$

where a_k are scalar?

Clearly, if $\mathbf{V}$ is diagonal, then so must $\mathbf{A}$ be. If $\mathbf{V}$ is upper triangular, then so must $\mathbf{A}$ be. Experimenting with various $\mathbf{V}$, such as $\mathbf{V} = a\mathbf{U} + b\mathbf{U}'$, gives some interesting shapes for $\{\mathbf{V}^k\}$. We can write $\mathbf{A}$ in terms of the Jordan form of $\mathbf{V}$. For example, if $\mathbf{V}$ has only one Jordan block, say $\mathbf{V} = \mathbf{LJ}_r(\lambda)\mathbf{L}^{-1}$, then

$$\mathbf{A} = \mathbf{L}\mathbf{B}\mathbf{L}^{-1},$$

where

$$\mathbf{B} = \sum_{k=0}^{r-1} b_k \mathbf{U}_r^k, \quad b_k = \sum_{i=k}^{\infty} \binom{i}{k} a_i \lambda^{i-k}.$$

How many linearly independent $\{\mathbf{V}^k, \quad k \geq 0\}$ are there? If $\mathbf{V}$ has only one Jordan block, that is, if $s = 1$, then the answer is $r + 1$, as we now show.

Theorem 4.1. Suppose that $\lambda \neq 0$ and $\mathbf{V} = \mathbf{LJ}_r(\lambda)\mathbf{L}^{-1}$. Then,

$$\sum_{k=0}^r a_k \mathbf{V}^k = \mathbf{0} \in \mathbb{C}^{r \times r} \quad (17)$$

if $a_k = \binom{r}{k} (-\lambda)^k$.

Proof. Set $\mathbf{T}$ equal to the left hand side of (17). Then $\mathbf{T} = \mathbf{L}\mathbf{C}\mathbf{L}^{-1}$ is an upper triangular matrix with $T_{i,j} = C_{j-i}$, where

$$\mathbf{C} = \sum_{k=0}^r \mathbf{J}_r(\lambda)^k a_k.$$

So, $\mathbf{T} = \mathbf{0} \in \mathbb{C}^{r \times r}$ if and only if $C_a = 0$ for $a = 0, 1, \dots, r-1$. Putting $a = r-1$, then $r-2$ confirms that $C_a = 0$. Now use induction. $\square$

Our final result gives expressions for $\mathbf{A}^\theta$, $\exp(\mathbf{A})$ and $\log \mathbf{A}$.

Theorem 4.2. For $\mathbf{A}$ given by (16),

$$\mathbf{A}^\theta = \sum_{i=0}^{\infty} a_{\theta,i} \mathbf{V}^i, \quad \exp(\mathbf{A}) = \sum_{i=0}^{\infty} b_i \mathbf{V}^i, \quad \log \mathbf{A} = \sum_{i=0}^{\infty} c_i \mathbf{V}^i,$$

where

$$\begin{aligned} a_{\theta,i} &= \sum_{k=0}^{\min(i,r-1)} \binom{r}{k} a_0^{\theta-k} \widehat{B}_{k,i}(\mathbf{a}), \\ b_i &= \exp(a_0) \sum_{k=0}^{\min(i,r-1)} \frac{\widehat{B}_{k,i}(\mathbf{a})}{k!}, \\ c_0 &= \log a_0, \quad c_i = \sum_{k=0}^{\min(i,r-1)} \frac{(-1)^{k-1}}{k} \widehat{B}_{k,i}(\mathbf{a}) \end{aligned}$$

for $i \geq 1$.

Proof. Set $\Delta = \mathbf{A} - a_0 \mathbf{I}_r$. By (16),

$$\Delta^k = \sum_{i=k}^{\infty} \widehat{B}_{i,k}(\mathbf{a}) \mathbf{V}^i.$$

Applying this to Examples 3.1–3.3, (13)–(15) give the results. $\square$

Funding information: Authors state no funding involved.

Author contributions: Both authors contributed equally to the manuscript.

Ethical approval: Not applicable.

Conflict of interest: Authors state no conflict of interest.

Data availability statement: Not applicable.

References

- [1] L. Comtet, *Advanced Combinatorics*, Reidel, Dordrecht, 1974.
- [2] P. Dellaportas and M. Pourahmadi, *Cholesky-GARCH models with applications to finance*, Stat. Comput. **22** (2012), 849–855, DOI: <https://doi.org/10.1007/s11222-011-9251-2>.
- [3] M. Fujii, *Pre-whitening QR-decomposition maximum likelihood detection for co-channel interference rejection in MIMO systems*, IEICE Trans. Commun. **E92B** (2009), 2529–2532, DOI: <https://doi.org/10.1587/transcom.E92.B.2529>.
- [4] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [5] Q. Wang, X. Wang, and X. Pan, *Adaptive sonar beamformer based on inverse QR decomposition and recursive least squares filter for underwater target detection*, Int. J. Remote Sens. **33** (2012), 3987–3998, DOI: <https://doi.org/10.1080/01431161.2010.523026>.
- [6] C. S. Withers and S. Nadarajah, *The n th power of a matrix and approximations for large n* , N Z J Math. **38** (2008), 171–178.
- [7] C. S. Withers and S. Nadarajah, *Fredholm equations for non-symmetric kernels with applications to iterated integral operators*, Appl. Math. Comput. **204** (2008), 499–507, DOI: <https://doi.org/10.1016/j.amc.2008.06.022>.