Multi-companion matrices
Georgi N. Boshnakov *
Department of Mathematics, UMIST, P.O. Box 88, Manchester M60 1QD, UK
Received 30 September 1997; accepted 18 July 2001
Submitted by H.J. Werner

Abstract

In this paper, we introduce and study the class of multi-companion matrices. They generalize companion matrices in various ways and possess a number of interesting properties. We find explicit expressions for the generalized eigenvectors of multi-companion matrices such that each generalized eigenvector depends on the corresponding eigenvalue and a number of quantities which are functionally independent of the eigenvalues of the matrix and (up to a uniqueness constraint) of each other. Moreover, we obtain a parameterization of a multi-companion matrix through the eigenvalues and these additional quantities. The number of parameters in this parameterization is equal to the number of non-trivial elements of the multi-companion matrix. The results can be applied to statistical estimation, simulation and theoretical studies of periodically correlated and multivariate time series in both discrete- and continuous-time.

© 2002 Elsevier Science Inc. All rights reserved.
AMS classification: 15A; 62M; 39A
Keywords: Jordan decomposition; Continuous-time autoregression; Generalized eigenvectors

1. Introduction

In this paper, we introduce and study the class of multi-companion matrices. They generalize companion matrices in various ways. Multi-companion matrices consist of a number of rows put on top of a rectangular unit matrix. Such matrices can be
factored into products of companion matrices or can be described as accompany-
ing some matrix polynomials. The (generalized) eigenvectors of multi-companion
matrices can be fully described/parameterized by their eigenvalues and a number
of quantities which are functionally independent of the eigenvalues and, up to a
uniqueness constraint, functionally independent among them as well.

The results can be applied to estimation, simulation and probabilistic study of
periodically correlated and multivariate time series in both discrete- and continuous-
time.

Multi-companion matrices have the following pattern:

\[
F = \begin{pmatrix}
    f_{11} & f_{12} & \cdots & f_{1m-k} & f_{1m-k+1} & \cdots & f_{1m} \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\
    f_{k1} & f_{k2} & \cdots & f_{km-k} & f_{km-k+1} & \cdots & f_{km} \\
    1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
    0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
    0 & 0 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix}.
\] (1.1)

**Definition 1.1.** The \( m \times m \) matrix \( F \) is said to be multi-companion of order \( k \) (or \( k \)-companion) if
1. the first \( k \) rows of \( F \) are arbitrary;
2. the \( k \)th sub-diagonal of \( F \) consists of 1’s;
3. all other elements of \( F \) are zero;
4. \( 1 \leq k < m \).

According to this definition a 1-companion matrix is companion. Occasionally it
may be convenient to think of the identity matrix as being a 0-companion matrix.
The requirement \( k < m \) is imposed for notational convenience only. Most of the
results turn into trivialities when \( k = m \). On the other hand, factorizations consid-
ered in Section 4 make sense for \( k = m \), too, and may be useful occasionally (see
Section 6.1).

To my knowledge the results on multi-companion matrices given in this paper are
new. They seem to be new even in the particular case obtained by replacing each
element of an ordinary companion matrix by a square matrix. Such matrices appear
in control theory. Their spectral properties are studied in [7, Chapter 14] and [6], see
also the following section for further discussion.

The main result is probably that each (generalized) eigenvector of a \( k \)-companion
\( m \times m \) matrix can be represented as a simple function of the corresponding eigen-
value and \( k - 1 \) quantities which are functionally independent from all eigenvalues
of the matrix and the generalized eigenvectors corresponding to different eigenvalues
(a constraint is necessary for eigenvectors corresponding to the same eigenvalue, see
Section 5.2 for full details). Generalized eigenvectors here are the columns of a matrix which transforms the given matrix to Jordan form. Putting this the other way round, any $k$-companion matrix can be obtained by specifying: its eigenvalues, the types of its Jordan blocks, and $k - 1$ quantities for each of its $m$ generalized eigenvectors. This results in $mk$ parameters. This is exactly the number of the non-trivial (not equal to 0 or 1) elements of a $k$-companion $m \times m$ matrix.

2. Notation

In this section, we introduce the notations, which are used throughout the paper maintaining as much consistency as reasonable.

Unless otherwise stated the matrices are $m \times m$. An $m \times n$ matrix with ones on the main diagonal and zeroes otherwise is denoted by $I_{n,n}$, the zero matrix by $0_{m,n}$, omitting the second index when $m = n$. Both indices are omitted when the context allows that. Elements of matrices are denoted by attaching indices to their (lowercase) names. A range index of the form $i : j$ specifies part of a row or column. The vector formed from $i$th to $j$th element of the $l$th row of a matrix is called its $l$th $(i:j)$-row, and similarly for columns. The natural notations $a_{il}$ and $a_{sj}$ for the $i$th row and $j$th column, respectively, are handy occasionally. The columns of a matrix which transform a given matrix into Jordan form are called its generalized eigenvectors.

The first $k$ rows of a $k$-companion matrix are called its non-trivial rows. The $m \times m$ $j$-companion matrices, $j = 1, \ldots, k$, whose non-trivial rows are the last $j$ non-trivial rows of $F$ are denoted by $F_j$. With this notation $F = F_k$ and $F_1$ is the companion matrix generated by the $k$th row of $F$. When referring to the elements of $F_j$ we use either this relation or notation as in (1.1) with an additional upper index $(j)$ attached.

We use at several places the notation

\[ f_{i*}^{(j)} = (f_{im}^{(j)}, \ldots, f_{im-1}^{(j)}) \]

for the $i$th $(m-j+2:m)$-row of the matrix $F_j$, $j \geq 2$. Thus $f_{i*}^{(j)}$ contains the last $j - 1$ elements of the $i$th row of $F_j$.

Let $M$ be any $m \times m$ matrix with determinant $\det(M)$. We say that the $j \times j$ minor in its lower-right corner is its backward leading minor of order $j$, $j = 1, 2, \ldots, m$. The backward leading minors of the upper right block of a $k$-companion matrix $F$ are denoted by $\delta_j$, $j = 1, 2, \ldots, k$, and the corresponding determinants by $\Delta_j$ (or $\Delta_j(F)$ to specify the source matrix).

With these notations $F$ can be written in block form as

\[
F = \begin{pmatrix}
F_{1,k,1,m-k} & \delta_k \\
I_{m-k} & 0_{m-k,k}
\end{pmatrix}.
\] (2.1)
Consider the matrix

\[ G = \begin{pmatrix} 0_{m-k,k} & I_{m-k} \\ s_k & G_{m-k+1,m+k+1} \end{pmatrix}. \]

(2.2)

The matrices \( G, F^T, \) and \( G^T \) are equally reasonable candidates to be named multi-companion. In the case \( k = 1, \) all these matrices are called companion. For continuous-time systems the usual choice is \( G, \) and in discrete-time more natural seems \( F. \) Different authors and computer algebra systems have their own preferences as to what to call companion matrix.

The particular case of multi-companion matrices of order \( k \) whose dimension \( m \) is divisible by \( k \) is important both for matrix theory and its applications, see [7, Chapter 14] and [6]. We hope that this paper besides considering the case of general \( k \) and \( m \) gives further insight into this important particular case as well. Compare for example the study of the Jordan decomposition presented here with the theory of matrix solvents, see [7, Section 14.9]. Matrix solvents provide illuminating parallels with the theory of scalar polynomials but they do not always exist. On the other hand, we describe the spectral properties of this kind of matrices by giving explicitly their generalized eigenvectors (whose form is also interesting), a description which applies to certain matrix polynomials with degenerate leading matrix coefficient, and needs no change when \( m \) is not divisible by \( k. \) In the applications we have in mind explicit expressions for the (generalized) eigenvectors in terms of the eigenvalues and additional parameters are of some independent interest.

The abbreviation \( C[\phi_1, \ldots, \phi_m] \) is used for the companion matrix, corresponding to the bracketed variables,

\[ C[\phi_1, \ldots, \phi_m] = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{m-1} & \phi_m \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \]

3. Multiplication by multi-companion matrices

Multi-companion matrices have properties which can be viewed as a generalization of the corresponding properties of companion matrices. For example, a \( k \)-companion matrix is non-singular if and only if its upper right block \( \delta_k \) (see Eq. (2.1)) has a non-zero determinant. This is essentially, up to the sign, the determinant of \( F. \) It is convenient to have a formal statement for it. If \( k = 1, \) then \( F \) is companion, its upper-right corner \( \delta_k \) is a scalar and \( \Delta_k = f_{km}. \)

**Lemma 3.1.** The determinant of the \( k \)-companion matrix \( F \) is equal to \((-1)^{m-k}k \Delta_k, \) or \((-1)^{m+1}k \Delta_k \) or \((-1)^{m-k}(m+1) \Delta_k.\)
Note. One can arrive at each of these formulas by appropriate permutations of the rows and columns of $F$. To verify that they are equivalent add or subtract, as appropriate, each pair of exponents to obtain expressions of the type $x(x+1)$, which are always even, and hence the summands leading to them have the same parity.

Direct calculation shows that the inverse of an invertible multi-companion matrix has the pattern of the matrix $G$ in (2.2),

$$F^{-1} = \begin{pmatrix} O_{m-k,k} & I_{m-k} \\ -\delta_k^{-1} & -\delta_k^{-1} F_{1:k,1:m-k} \end{pmatrix}.$$ 

Multiplication by multi-companion matrices possesses some nice properties summarized in the following theorems. Again, setting $k = 1$ we obtain the corresponding properties of the companion matrices. Proofs are simple and therefore omitted.

**Theorem 3.1.** Let $F$ be a $k$-companion matrix, and $B$ an arbitrary matrix.

1. If $L = FB$, then for $j = 1, \ldots, m$,

$$l_{ij} = \begin{cases} f_i \cdot b_{kj}, & i = 1, \ldots, k \\ b_{i-k,j}, & i = k + 1, \ldots, m \end{cases},$$

i.e., left multiplication of any matrix $B$ by a $k$-companion matrix $F$ moves the first $m - k$ rows of $B$ $k$ rows downwards without change. The first $k$ rows of the product $L = FB$ have the usual general form.

2. If $R = BF'$, then for $i = 1, \ldots, m$,

$$r_{ij} = \begin{cases} b_{i,1} \cdot f'_{j*}, & j = 1, \ldots, k \\ b_{i,j+k}, & j = k + 1, \ldots, m \end{cases},$$

i.e., right multiplication of any matrix $B$ by the transposed of a $k$-companion matrix $F$ moves the first $m - k$ columns of $B$ $k$ columns rightward without change. The first $k$ columns of the product $R = BF'$ have the usual general form.

3. If $S = FBF'$ is the “symmetric” product of $F$ with $B$, then $s_{ij}$ is given by the following table:

<table>
<thead>
<tr>
<th>Elements $s_{ij}$ of the “symmetric” product $S = FBF'$</th>
<th>$j = 1, \ldots, k$</th>
<th>$j = k + 1, \ldots, m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i = 1, \ldots, k$</td>
<td>$l_{i \cdot f'_{j*}}$</td>
<td>$l_{i \cdot f'_{j*}}$</td>
</tr>
<tr>
<td>$i = 1, \ldots, k$</td>
<td>$l_{i \cdot f'<em>{j*}} = f_i \cdot b</em>{kj}$</td>
<td>$l_{i \cdot f'<em>{j*}} = f_i \cdot b</em>{kj}$</td>
</tr>
<tr>
<td>$i = k + 1, \ldots, m$</td>
<td>$l_{i \cdot f'<em>{j*}} = b</em>{i-k, j+k}$</td>
<td>$l_{i \cdot f'<em>{j*}} = b</em>{i-k, j+k}$</td>
</tr>
</tbody>
</table>

where $l_{ij}$ are given by Eq. (3.1) above.

Thus the symmetric product of any matrix $B$ with a companion matrix and its transpose moves each element of the upper left $(m - k) \times (m - k)$ block of $B$
k rows downwards and k columns rightwards (i.e., k positions in south-east direction). The elements in the first k rows and first k columns of the product are formed by the above formulas from which it can be seen that only the upper left \( k \times k \) block of the product is quadratic in respect to the non-trivial rows of \( F \).

The following simple corollaries of Theorem 3.1 deserve separate formulation.

**Corollary 3.1.** Let \( F = X_k Y_l \), where \( X_k \) is \( k \)-companion, \( Y_l \) is \( l \)-companion, and \( k + l < m \). Then
1. \( F \) is \( k + l \)-companion;
2. the last \( l \) non-trivial rows of \( F \) coincide with the non-trivial rows of the second factor, \( Y_l \);
3. \( A_{k+l}(F) = A_k(X_k)A_l(Y_l) \).

**Proof.** Only the last part needs further justification. By Lemma 3.1 \( \det F = (-1)^{(k+l)(m+1)}A_{k+l}(F) \). On the other hand,

\[
\det F = \det X_k \det Y_l = (-1)^{k(m+1)} A_k(X_k)(-1)^{(m+1)} A_l(Y_l).
\]

A comparison of the two equations gives the desired result. \( \square \)

In particular, the product (in any order) of a \( (k - 1) \)-companion matrix by a companion matrix is \( k \)-companion.

**Corollary 3.2.** The product \( A_k \cdots A_1 \) of companion matrices

\[
A_i = \mathbb{C}[\phi_{i1}, \ldots, \phi_{im}], \quad i = 1, \ldots, k, \ k < m,
\]

is multi-companion of order \( k \). This product is non-singular if and only if

\[
\phi_{km} \cdots \phi_{1m} \neq 0.
\]

The inverse assertion is not true even in the non-singular case—given a \( k \)-companion matrix it is not always possible to represent it as a product of companion matrices. This problem is the subject of the following section.

4. Factorizations of multi-companion matrices

4.1. Representation as a product of two matrices

4.1.1. Factoring into companion times multi-companion

We are looking for a companion matrix \( A \) for which the following equation holds:

\[
F_k = AF_{k-1}.
\]  

(4.1)

It follows from Corollary 3.1 that the non-trivial rows of \( F_{k-1} \) must coincide with the second to \( k \)th non-trivial rows of \( F_k \). Hence, the expanded form of Eq. (4.1) is
Hence, the elements of these matrices obey the equations

\[
\begin{align*}
\mathbf{f}(k)_{1j} &= \sum_{i=1}^{k-1} a_i \mathbf{f}(k)_{i+1,j} + a_{k-1+j}, \quad j = 1, \ldots, m - k + 1, \\
\mathbf{f}(k)_{1j} &= \sum_{i=1}^{k-1} a_i \mathbf{f}(k)_{i+1,j}, \quad j = m - k + 2, \ldots, m, \\
\mathbf{f}(k)_{ij} &= \mathbf{f}(k-1)_{i-1,j}, \quad i = 2, \ldots, m, \quad j = 1, \ldots, m.
\end{align*}
\]

Eqs. (4.2) are already solved explicitly for \(a_{k-1+j}, \quad j = 1, \ldots, m - k + 1\) (i.e., \(a_k, \ldots, a_m\)).

\[
a_j = \mathbf{f}(k)_{1j-1} - \sum_{i=1}^{k-1} a_i \mathbf{f}(k)_{i+1,j-1} + \mathbf{f}(k)_{1j-1}, \quad j = k, \ldots, m.
\]

The remaining equations involve operations on parts of the rows of \(\mathbf{F}_{k-1}\). Namely, let \(\mathbf{f}(k)_{i*}\) be the \(i\)th \((m - k + 2: m)\)-row of \(\mathbf{F}_{k-1}\), i.e., \(\mathbf{f}(k)_{i*} = (\mathbf{f}(k)_{Im-k+2}, \ldots, \mathbf{f}(k)_{Im})\).

Then, from (4.3) and (4.4),

\[
\begin{align*}
\mathbf{f}(k)_{11} &= \mathbf{f}(k)_{12} = \ldots = \mathbf{f}(k)_{1m-k}, \\
\mathbf{f}(k)_{1m-k+1} &= \mathbf{f}(k)_{1m-k+2} = \ldots = \mathbf{f}(k)_{1m}.
\end{align*}
\]
\[ f_i^{(k)} = \sum_{i=1}^{k-1} a_i f_{i+1}^{(k)}. \]  
\[ f_i^{(k)} = f_{i-1}^{(k-1)}, \quad i = 2, \ldots, m. \]  

(4.5)

Hence, the problem of finding \( A \) reduces to studying the solutions of the \((k-1) \times (k-1)\) system (4.3) (or (4.5)) for \( a_1, \ldots, a_{k-1} \). The dimension of this system is independent of the size, \( m \), of the matrix. The matrix of this system is the upper right \((k-1) \times (k-1)\) corner, \( \delta_k \) of \( F_k \). Hence, if \( \delta_k \) is non-singular, then the solution for the \( a \)'s exists and is unique. More careful examination of Eqs. (4.3) gives complete description of the solutions we are looking for. Namely, these equations express the sub-row vector \( f_1^{(k)} = (f_1^{(k)} - k + 2, \ldots, f_m^{(k)}) \) as a linear combination of sub-row vectors \( f_i^{(k)}, i = 2, \ldots, k \), which are located below it in \( F_k \). Hence, we have the following result.

**Theorem 4.1.** Let \( F \) be a \( k \)-companion matrix. Factorization \( F = AF_{k-1} \), with companion matrix \( A \), exists if and only if the vector \( f_1^{(k)} \) lies in the space spanned by the vectors \( f_i^{(k)}, i = 2, \ldots, k \).

In particular, \( f_1^{(k)} \) is in the desired space when \( f_i^{(k)}, i = 2, \ldots, k \), are linearly independent, or, equivalently, when the backward leading minor, \( \delta_k \), of \( F \) has a non-zero determinant.

In some cases, when factorization (4.1) does not exist for \( F \) itself, it is sufficient to have factorization for an extended form of \( F \) which is again \( k \)-companion and is obtained by appending zero columns to \( F \) and appropriate rows. The following useful corollary of Theorem 4.1 holds.

**Corollary 4.1.** For some \( i \in [1, k] \) the \((m+i) \times (m+i)\) \( k \)-companion matrix \( W_i \) with non-trivial rows equal to those of \( F \) (appended with \( i \) zeroes) admits factorization of type (4.1).

\[ W_i = G_{k-1} X, \]  

(4.6)

**Proof.** First, notice that the conditions of Theorem 4.1 hold if the upper right corner of the multi-companion matrix consists of zeroes only. The matrix \( W_k \) has this pattern. Hence the corollary holds with \( i = k \). Usually, a smaller \( i \) will do the job as well since the rank of the upper right corner of \( W_i \), in general, decreases with \( i \). A simple example when the smaller possible value of \( i \) is \( k \) is given by a matrix \( F \) whose last column contains only one non-zero element in its first row. \( \square \)
where $X$ is companion and, by Corollary 3.1, its non-trivial row coincides with the $k$th row of $F_k$. We omit below the upper index of that row.

\[
\begin{pmatrix}
    f_{11}^{(k)} & f_{12}^{(k)} & \cdots & f_{1m-k}^{(k)} & f_{1m-k+1}^{(k)} & \cdots & f_{1m}^{(k)} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    f_{k1}^{(k)} & f_{k2}^{(k)} & \cdots & f_{km-k}^{(k)} & f_{km-k+1}^{(k)} & \cdots & f_{km}^{(k)} \\
    1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
    0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix}
\]

\[
\begin{pmatrix}
    g_{11}^{(k-1)} & g_{12}^{(k-1)} & \cdots & g_{1m-k}^{(k-1)} & g_{1m-k+1}^{(k-1)} & \cdots & g_{1m}^{(k-1)} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    g_{k1}^{(k-1)} & g_{k2}^{(k-1)} & \cdots & g_{km-k}^{(k-1)} & g_{km-k+1}^{(k-1)} & \cdots & g_{km}^{(k-1)} \\
    1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
    0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix}
\times
\begin{pmatrix}
    f_{k1} & f_{k2} & \cdots & f_{km-1} & f_{km} \\
    1 & 0 & \cdots & 0 & 0 \\
    0 & 1 & \cdots & 0 & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    0 & 0 & \cdots & 1 & 0
\end{pmatrix}
\]

The elements of these matrices obey the following equations:

\[
\begin{align*}
    f_{ij}^{(k)} &= g_{ij}^{(k-1)} f_{kj} + g_{ij}^{(k-1)} f_{ij+1}, & i = 1, \ldots, m-1, & j = 1, \ldots, k-1, \\
    f_{im}^{(k)} &= g_{i1}^{(k-1)} f_{km}, & i = 1, \ldots, k-1, \\
    f_{kj}^{(k)} &= f_{kj}, & j = 1, \ldots, m.
\end{align*}
\]

In (sub-) column form these equations look as

\[
\begin{align*}
    f_{1k-1,j-1}^{(k)} &= g_{1k-1,j-1}^{(k-1)} f_{kj-1} + g_{1k-1,j-1}^{(k-1)} f_{1k-1,j}, & j = 2, \ldots, m, \\
    f_{1k-1,m}^{(k)} &= g_{1k-1,m}^{(k-1)} f_{km}.
\end{align*}
\]

Now, if $f_{km} \neq 0$, the second equation gives $g_{1k-1,m}^{(k-1)} = f_{1k-1,m}^{(k)} / f_{km}$. Then the other non-trivial columns, $g_{1k-1,j}^{(k-1)}$, can be determined by the formulas

\[
\begin{align*}
    g_{1k-1,j}^{(k-1)} &= f_{1k-1,j-1}^{(k)} - g_{1k-1,j}^{(k-1)} f_{kj-1}.
\end{align*}
\]
Hence, we have the following properties.

**Lemma 4.1.** If $f_{km}^{(k)} \neq 0$, then there exists a unique factorization $F = G_{k-1}X$ with $X$ companion.

**Lemma 4.2.** If $f_{km}^{(k)} = 0$, then a factorization $F = G_{k-1}X$ exists if and only if the last column of $F$ is zero.

This lemma shows, in particular, that a non-singular $k$-companion matrix with $f_{1m}^{(k)} = 0$ cannot be represented as a product of companion matrices.

If the conditions of Lemma 4.2 are satisfied, the first column $g_{s1}^{(k-1)}$ may be chosen arbitrarily and therefore the factorization is not unique. In applications some particular choices of the factorization may be more attractive than others. One possibility is to set the first column $g_{s1}^{(k-1)}$ equal to the zero vector. It may be more natural to make zero the last column $g_{sm}^{(k-1)}$, not the first one, although this is not always possible. Suppose that in the considered case $f_{km-1} \neq 0$. Then the setting $g_{s1}^{(k-1)} = f_{1km-1}^{(k)} / f_{km-1}$ ensures that $g_{sm}^{(k-1)} = 0$. Otherwise if $f_{km-1} = 0$, then $f_{1km-1}^{(k)} = g_{sm}^{(k-1)}$.

More generally, suppose that $f_{km-\gamma} \neq 0$ and $f_{km-\gamma+1} = \cdots = f_{km} = 0$. Then

\[
\begin{align*}
    f_{1km-1}^{(k)} &= g_{s1}^{(k-1)} f_{ki} + g_{s1+1}^{(k-1)} \quad \text{for } i \leq m - \gamma, \\
    f_{1km-1}^{(k)} &= g_{s1}^{(k-1)} \quad \text{for } i = m - \gamma + 1, \ldots, m - 1, \\
    f_{1km-1}^{(k)} &= 0.
\end{align*}
\]

As above we can choose $g_{s1}^{(k-1)} = 0$ or force $g_{sm-\gamma+1}^{(k-1)} = 0$ by the choice

\[
g_{s1}^{(k-1)} = f_{1km-\gamma}^{(k)} / f_{km-\gamma}.
\]

Consider, for example, the case $k = 2$ which is of some importance by itself. In that case $F_1$ is companion and system (4.3) reduces to the single equation

\[
f_{1m}^{(2)} = a_1 f_{2m}^{(2)}.
\]

Eqs. in (4.2) are

\[
f_{1j}^{(2)} = a_1 f_{2j}^{(2)} + a_{1+j}, \quad j = 1, \ldots, m - 1.
\]

The solution is unique when $f_{2m}^{(2)} \neq 0$. There is no solution if $f_{2m}^{(2)} = 0$ and $f_{1m}^{(2)} \neq 0$ (i.e., $f_{1m}^{(2)} \neq 0$ does not belong to the space spanned on $f_{2m}^{(2)}$) and there are infinitely many solutions when both coefficients are zero, with $a_1$ arbitrary, and $a_j = f_{1j-1}^{(2)}$, for $j = 2, \ldots, m$.

See Section 6.1 for a time series example.
4.1.3. Multi-companion times multi-companion

Having a factorization (4.6) we can continue by factoring $G_{k-1}$ in the same way. However, there is an important difference between the results for left and right factorization with companion matrices, considered so far. The factorization $F_j = A_j F_{j-1}$ determines the companion matrices $A_j$ in parallel, independently for $j = k, \ldots, 1$, even when one or more $A_j$ are not unique. On the contrary, let $G_k \equiv F_k$ and consider the sequence of factorizations

$G_k = G_{k-1} X_k, G_{k-1} = G_{k-2} X_{k-1} \cdots G_2 = G_1 X_2$

$G_1 = X_1$ and suppose that $G_j$ is not unique. Then not only $X_j$ may be different for the different possible choices of $G_j$, the corresponding $G_{j-1}$ may also be different, thus affecting the subsequent factorizations $G_{j-1}$ as $G_{j-2} X_{j-1}, \ldots$. Moreover, it may happen that for some $j$ the factorization process cannot be continued even if it exists for the whole matrix $F$.

However, if we allow for $X$ to be multi-companion, this problem disappears since then at each step all the information about $F$ is used. Extending the arguments from Section 4.1.1 we obtain a necessary and sufficient condition for factorization of a multi-companion matrix into multi-companion matrices of lower order, $F_k = G_u G_v$, where $k = u + v$ and, necessarily, $G_v = F_v$. Namely, for every non-trivial row of $G_u$ we can do the same analysis as in the proof of Theorem 4.1 (where a companion matrix $A$ stands in place of $G_u$).

**Theorem 4.2.** Let $F$ be a $k$-companion matrix, $1 \leq u < k$, $v = k - u$. Then a factorization $F = G_u F_{k-u}$, where $G_u$ is $u$-companion, exists if and only if for each $l, l = 1, \ldots, u$, the $l$th $(m - v + 1 : m)$-row vector, lies in the space spanned by the $l$th $(m - v + 1 : m)$-row vectors $i = u + 1, \ldots, k$.

**4.2. Maximal factorizations of multi-companion matrices**

The representations of the multi-companion matrix $F$ as products of two matrices considered so far form the basis for its full factorization.

**Theorem 4.3.** Factorization of a $k$-companion matrix $F$ into a product $A_k \cdots A_1$ of companion matrices exists if and only if for each $j, j = k, \ldots, 2$, the vector $f_{k-j+1}^{(j)}$ lies in the space spanned by the vectors $f_{k-j+2}^{(j)}, \ldots, f_k^{(j)}$.

**Proof.** Apply Theorem 4.1 to $F_j, j = k, \ldots, 2$. □

The next result is a useful particular case of the above theorem.

**Proposition 4.1.** Let $F$ be an $m \times m$ $k$-companion matrix.

1. A necessary condition for the existence of companion matrices $A_i$, $i = 1, \ldots, k$ such that $F = A_k \cdots A_1$, is that if for some $i$, $A_i = 0$, then $A_j = 0$ for all $j > i$.
2. A sufficient condition is that all $A_i$’s are different from zero.
Proof. The necessary condition follows from the fact that if $F_i$ is singular, then the $F_j$'s with $j > i$ are also singular since they are products of $F_i$ with other matrices. The sufficient condition is obvious since system (4.3) has a unique solution when it is fulfilled. □

Note that, in general, if some of the backward leading minors of a matrix are zero, this does not imply that the determinant $\Delta_k$ is zero.

The above considerations show that each factor $A_i$ is determined independently of the others. A somewhat unexpected consequence of this fact is that if some $A_i$ is not unique, the different choices for it do not influence the other factors. This is convenient and intuitive in time series modeling where some choices may be preferable. Two typical criteria are: (i) as many of the first $a_i$'s to be zero (it possible); (ii) as many of the $a_i$'s to be zero. The two criteria may be combined if the second does not give a unique solution.

An important feature of the above theorem is that the system it involves has dimension $k$ which does not depend on the dimension of $F$. In particular the determination of the matrices $A_i$ from $F$ is simple (at least in the non-singular case).

The singular case (which allows for some of the $A_i$'s to have $\phi_{im} = 0$) is important since it corresponds to a periodic autoregression with different orders for the different seasons (see $p_i$ in (6.1)).

Using Theorem 4.2 we obtain a canonical (maximal) factorization for $F$.

Theorem 4.4. There exist integer numbers $r$, $k_i$, $i = 1, \ldots, r$, with $\sum_i k_i = k$, and multi-companion matrices $G_{k_i}$, $i = 1, \ldots, r$ such that

$$F_k = G_{k_1} G_{k_2} \cdots G_{k_r}$$

and in any other factorization $H_1 \cdots H_s$, $H_i$ are products of $G_{k_i}$.

Proof. The proof is constructive. In the trivial case when there is no $v$ such that $F = G_u F_v$, put $r = 1$, $k_1 = k$, $G_{k_1} = F_k$. Otherwise, take the maximal $v < k$ for which $F = G_u F_v$. Then $G_u$ cannot be factored further (otherwise we would have a contradiction with the choice of $v$). Do the same with $F_v$, etc. □

Although not all multi-companion matrices can be factored fully into companion matrices there always exists a permutation of the first $k$ rows which makes this possible. To this end, consider the last column of $F$.

1. If the last column is zero, then there exists a factorization $F = G_{k-1} A_1$.
2. If $f_{km} \neq 0$, then there exists factorization $F = G_{k-1} A_1$.
3. If $f_{km} = 0$ and the last column is not entirely zero, then a permutation $P_{k-1}$ of the first $k$ rows of $F$ will move a non-zero element to position $(km)$. From condition 2 follows that there exists a factorization of $P_{k-1} F$ as $G_{k-1} A_1$, i.e., $F = P_{k-1}' G_{k-1} A_1$, since permutation matrices are orthogonal.
Applying this procedure recursively to $G_{k-1}, \ldots, G_2$ we end up with the representation $F = P'_{k-1} \cdots P'_1 A_k \cdots A_1$. Products of permutation matrices give permutation matrices. Hence, we have the following result.

**Theorem 4.5.** Any $k$-companion matrix $F$ can be factored as

$$F = PA_k \cdots A_1,$$

where $P$ is a (row) permutation matrix and $A_i, i = 1, \ldots, k$, are companion matrices.

5. Eigenvalues and eigenvectors of multi-companion matrices

5.1. Characteristic polynomial of $F$

To describe the eigen-structure of $F$ consider its characteristic matrix

$$
\lambda I - F = \begin{pmatrix}
\lambda - f_{11} & -f_{12} & \cdots & -f_{1m-k} & -f_{1m-k+1} & \cdots & -f_{1m} \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
-f_{k1} & -f_{k2} & \cdots & -f_{km-k} & -f_{km-k+1} & \cdots & -f_{km} \\
-1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & -1 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & -1 & 0 & \cdots & \lambda 
\end{pmatrix}.
$$

(5.1)

To write down $F$ in block form for convenient manipulation, let us denote by $s$ the smallest integer greater than or equal to $m/k$, i.e., $s = \lceil m/k \rceil$, and by $t$ the remainder $(m \mod k)$. Then $t = m - (s-1)k$ and (note the minus signs)

$$F = \begin{pmatrix}
-A_1 & -A_2 & \cdots & -A_{s-1} & -A_s \\
I_{t,t} & 0_{k,k} & \cdots & 0_{k,k} & 0_{t,k} \\
0_{k,t} & I_k & \cdots & 0_{k,k} & 0_{t,k} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0_{k,t} & 0_{k,k} & \cdots & I_k & 0_{k,k} 
\end{pmatrix},$$

(5.2)

where $A_1$ is $k \times t$, $A_i$ are $k \times k$ for $i = 2, \ldots, s$. So, there are $s$ block columns/rows in $F$. Denoting by $H_i$ a $k \times k$ matrix with ones on the $i$th sub-diagonal and zeroes elsewhere and, similarly by $H_{t,k}$ a $t \times k$ matrix with ones on the $(k-t)$th super-diagonal and zeroes elsewhere, i.e.,

$$H_i = \begin{pmatrix}
0_{i,k} \\
I_{k-i,k}
\end{pmatrix}, \quad H_{t,k} = \begin{pmatrix}
0_{t,k-t} & I_{t,t}
\end{pmatrix}. \quad (5.2)$$
we can write
\[
\lambda I - F = \begin{pmatrix}
I_{k,t} \lambda + A_1 & \lambda H_t + A_2 & \cdots & A_{s-1} & A_s \\
-I_t & H_{t,k} \lambda & \cdots & 0_{t,k} & 0_{t,k} \\
0_{k,t} & -I_k & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0_{k,t} & 0 & \cdots & -I_k & \lambda I_k
\end{pmatrix}.
\]

Note that for \( t = k \) we have \( H_t = 0_k \) and \( H_{t,k} = I_{t,t} = I_t = I_k \).

It is helpful to look more closely at the upper left corner of this matrix, the example is for \( t = 3, k = 5 \); \( A_1 \) and \( A_2 \) are omitted for readability,
\[
\begin{pmatrix}
I_{k,t} \lambda & \lambda H_t \\
-I_t & H_{t,k} \lambda
\end{pmatrix} =
\begin{pmatrix}
\lambda & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \lambda & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \lambda & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \lambda & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda & 0 & 0 \\
-1 & 0 & 0 & 0 & 0 & \lambda & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & \lambda \\
0 & 0 & -1 & 0 & 0 & 0 & \lambda
\end{pmatrix}.
\]

Notice that the first block column of \( \lambda I - F \) has \( t \) columns.

Standard manipulations on the block columns of \( \lambda I - F \) allow us to transform it to a simpler form. To turn into zeroes all the blocks of the last block column except the first one, we multiply the \((s - i)\)th block column by \( \lambda^i \) and add it to the \( s \)th block column for \( i = 1, \ldots, s - 2 \). Each of these operations makes zero the corresponding \((s - i + 1)\)st block element and introduces an element \( \lambda^{i+1} I_k \) in the \((s - i)\)th block row. Finally, we right multiply the first block column by \( H_{t,k} \lambda^{s-1} \) and add it to the last column. The multiplication by \( H_{t,k} \) is needed to adjust the dimensions of the corresponding matrices, it prepends \( k - t \) zero columns to \( A_1 \) and \( I_{k,t} \). After these manipulations the block in the upper right corner of \( \lambda I - F \) looks as
\[
A_s + A_{s-1} \lambda + \cdots + A_3 \lambda^{s-3} + (\lambda H_t + A_2) \lambda^{s-2} + (I_{k,t} \lambda + A_1) H_{t,k} \lambda^{s-1},
\]
which can be rearranged as
\[
T(\lambda) = A_s + A_{s-1} \lambda + \cdots + A_3 \lambda^{s-3} + A_2 \lambda^{s-2} + (H_t + A_1 H_{t,k}) \lambda^{s-1} + I_{k,t} H_{t,k} \lambda^2.
\] (5.3)

It is easy to show now that the nontrivial invariant polynomials of \( \lambda I - F \) and \( T(\lambda) \) coincide. First, up to now we have used only elementary operations in the transformation of \( \lambda I - F \). Second, we can make the first block row of the above matrix zero by obvious elementary row operations. Then a permutation matrix will move \( T(\lambda) \) in the lower right corner, rising all other rows upwards. The final transformations to obtain the desired canonical form are given by pre-multiplying and post-multiplying by
block matrices, having in their lower right corners the corresponding transformation matrices for $T(\lambda)$.

If $k$ is a divisor of $m$, then $I_{k,t} = I_k$, $H_k = 0_k$, $H_{t,k} = I_k$ and $T(\lambda)$ simplifies to

$$A_k + A_{k-1}\lambda + \cdots + A_3\lambda^{s-3} + A_2\lambda^{s-2} + A_1\lambda^{s-1} + \lambda^s.$$  

The above manipulations can be written in matrix form, details are given in Appendix A.

An alternative way to simplify the matrix $I\lambda - F$ is to multiply each block column by $\lambda$ and add the result to the block column to its right, starting this time from the first block column. Again some care must be taken with the non-square blocks which must be pre-multiplied by appropriate matrices. After such manipulations the matrix $\lambda I - F$ is transformed to

$$M(\lambda) = \begin{pmatrix} \alpha & T(\lambda) \\ -I_{m-k} & 0 \end{pmatrix},$$

where

$$\alpha = (I_{k,t}\lambda + A_1, (I_{k,t}\lambda + A_1)H_{t,k}\lambda + (\lambda H_t + A_2), (I_{k,t}\lambda + A_1)H_{t,k}\lambda + (\lambda H_t + A_2)\lambda + A_3, \ldots).$$

This approach eliminates directly all non-trivial elements from block rows 2 to $s$. Matrix formulation and details can be found in Appendix B.

5.2. Eigenvectors and Jordan chains

Consider the equation

$$Fx = \lambda x,$$  \hspace{1cm} (5.4)

which relates $F$ to an eigenvalue $\lambda$ and a corresponding eigenvector $x$. From Theorem 3.1 we know that $F$ shifts the elements of $x$ downwards by $k$ rows. Hence, using notation $x_j$ for the $j$th element of $x$, we have

$$x_{j-k} = \lambda x_j, \quad j > k.$$  

This equation shows that the first (or any consecutive) $k$ elements of $x$ determine uniquely the entire eigenvector $x$. An immediate consequence is the following proposition.

**Proposition 5.1.** There can be at most $k$ linearly independent eigenvectors of a $k$-companion matrix corresponding to a given eigenvalue.

The Jordan chain for the eigenvalue $\lambda$ and the eigenvector $x$ can be written explicitly for the matrix $F$. 
Suppose that the geometric multiplicity of the eigenvalue $\lambda$ is less than its algebraic multiplicity, $x$ is one of the eigenvectors corresponding to it, with chain $x^{(1)} \equiv x, x^{(2)}, \ldots, x^{(s)}$ of length $s$. Then

$$Fx^{(1)} = \lambda x^{(1)},$$
$$Fx^{(i)} = \lambda x^{(i)} + x^{(i-1)} \quad \text{for } i = 2, \ldots, s,$$

and the two equations can be squeezed into one if we adopt the convention $x^{(0)} \equiv 0$.

As in the case of eigenvectors, the generalized eigenvectors of $F$ have only $k$ free components. Indeed, from (5.5) we have

$$(Fx^{(i)})_j = \lambda x^{(i)}_j + x^{(i-1)}_j, \quad j = 1, \ldots, k,$$

where, by Theorem 3.1, $(Fx^{(i)})_j = x^{(i)}_{j-k}$ for $j > k$. Hence,

$$(x^{(i)})_{j-k} = \lambda x^{(i)}_j + x^{(i-1)}_j, \quad j > k.$$  \hfill (5.6)

The discussion above shows that the elements of any eigenvector $x$ of a $k$-companion matrix can be written as

$$x_j = c_j \lambda^{l_j},$$  \hfill (5.7)

where $c_{j-k} = c_j$ and $l_{j-k} - l_j = 1$, $j = 1, \ldots, m$, and there are only $k$ different $c_j$'s. The powers $l_j$ can be normalized in different ways, e.g. so that either $l_1 = \cdots = l_k = 0$, or $l_{m-k+1} = \cdots = l_m = 0$. We will use the latter convention which simply states that the last $k$ elements of the eigenvector are taken as independent parameters while the remaining are computed from them. We will do the same with the generalized eigenvectors from a Jordan chain.

Namely, let $r_j$, $j = 1, \ldots, k$, be such that $l_j = j + r_j k \leq m$, but $l_j + k > m$. Then we denote

$$c^{(i)}_j = x^{(i)}_j.$$  \hfill \(5.8\)

Obviously, $c^{(1)}_j = c_j$ under the adopted convention for $l$.

5.2.1. Generalized eigenvectors of $F$—first approach

Suppose that $\lambda \neq 0$. Then we can write the $j$th element of $x^{(i)}$ in the form $d^{(i)}_j \times \lambda^{l_j-i+1}$, where $l_j = l_{j-k} - 1$, and $d^{(i)}_j$ may depend on $\lambda$. Then

$$d^{(i)}_{j-k} \lambda^{l_j-i+1} = \lambda d^{(i)}_{j} \lambda^{l_j-i+1} + d^{(i-1)}_{j} \lambda^{l_j-i+2}.$$  

Hence,

$$d^{(i)}_{j-k} \lambda^{l_j-i+2} = d^{(i)}_{j} \lambda^{l_j-i+2} + d^{(i-1)}_{j} \lambda^{l_j-i+2}.$$  

Now division by $\lambda^{l_j-i+2}$ gives

$$d^{(i)}_{j-k} = d^{(i)}_{j} + d^{(i-1)}_{j},$$  \hfill (5.8)
i.e.,
\[ d_j^{(i)} = d_{j-k}^{(i)} - d_j^{(i-1)}. \]

Writing down the last equation for a range of values of the index we have
\[
\begin{align*}
  d_j^{(i)} &= d_j^{(i-1)} - k - d_{j-1}^{(i-1)} \\
  d_j^{(i+1)} &= d_{j+k}^{(i-1)} - d_j^{(i-1)} \\
  d_j^{(i+2)} &= d_{j+k(2)}^{(i-1)} - d_j^{(i-1)} \\
  \vdots \\
  d_j^{(i+k)} &= d_j^{(i)} - d_{j+k}. \\
\end{align*}
\]

Summing them up we obtain
\[
\begin{align*}
  d_j^{(i)} &= d_j^{(i)} - \sum_{m=1}^{l} d_j^{(i-1)} \\
  \text{(5.9)}
\end{align*}
\]

In particular,
\[
\begin{align*}
  d_j^{(1)} &= d_j^{(1)} \\
  d_j^{(2)} &= d_j^{(2)} - \sum_{m=1}^{l} d_j^{(1)} \\
  &= d_j^{(2)} - ld_j^{(1)} \\
  d_j^{(3)} &= d_j^{(3)} - \sum_{m=1}^{l} d_j^{(2)} \\
  &= d_j^{(3)} - \sum_{m=1}^{l} (d_j^{(2)} - md_j^{(1)}) \\
  &= d_j^{(3)} - ld_j^{(2)} + \sum_{m=1}^{l} md_j^{(1)} \\
  &= d_j^{(3)} - ld_j^{(2)} + \frac{(1 + l)l}{2} d_j^{(1)}. \\
\end{align*}
\]

Let
\[
\begin{align*}
  g_0(l) &= 1, \quad g_1(l) = \sum_{i=1}^{l} g_0(i) = l, \quad g_2(l) = \sum_{i=1}^{l} g_1(i) = l(l + 1)/2, \\
  \text{and} \quad g_k(l) &= \sum_{i=1}^{l} g_{k-1}(i) \quad \text{for } k > 1.
\end{align*}
\]
We will prove by induction that
\[
d_{j+lk}^{(i)} = \sum_{m=0}^{i-1} (-1)^m g_m(l)d_j^{(i-m)}.
\] (5.10)

Indeed, for the first several elements this is true,
\[
d_{j+lk}^{(1)} = (-1)^0 g_0(l)d_j^{(1-0)} = d_j^{(1)},
\]
\[
d_{j+lk}^{(2)} = d_j^{(2)} + (-1)^1 g_1(l)d_j^{(2-1)}
\]
\[= d_j^{(2)} - ld_j^{(1)},
\]
\[
d_{j+lk}^{(3)} = d_j^{(3)} - g_1(l)d_j^{(2)} + g_2(l)d_j^{(1)}
\]
\[= d_j^{(3)} - ld_j^{(2)} + (l(l+1)/2)d_j^{(1)}.
\]

Suppose that Eq. (5.10) is true for \(i\). We will show that then it is true also for \(i+1\), i.e., that
\[
d_{j+lk}^{(i+1)} = \sum_{m=0}^{i} (-1)^m g_m(l)d_j^{(i+1-m)}.
\]

Using Eq. (5.9) and the inductive assumption we have
\[
d_{j+lk}^{(i+1)} = d_j^{(i+1)} - \sum_{m=1}^{i} d_{j+mk}^{(i)}
\]
\[= d_j^{(i+1)} - \sum_{m=1}^{i} \sum_{x=0}^{i-1} (-1)^x g_x(m)d_j^{(i-x)}
\]
\[= d_j^{(i+1)} - \sum_{x=0}^{i-1} (-1)^x d_j^{(i-x)} \sum_{m=1}^{i} g_x(m)
\]
\[= d_j^{(i+1)} - \sum_{x=0}^{i-1} (-1)^x d_j^{(i-x)} g_{x+1}(l)
\]
\[= d_j^{(i+1)} + \sum_{x=0}^{i-1} (-1)^x d_j^{(i-x)} g_{x+1}(l)
\]
\[= d_j^{(i+1)} + \sum_{x=1}^{i} (-1)^x d_j^{(i+1-x)} g_x(l)
\]
\[= \sum_{x=0}^{i} (-1)^x g_x(l)d_j^{(i+1-x)},
\]
as desired. By induction the formula is true.
From Eq. (5.10) we can see also that if $d_{j}^{(1)}$ do not depend on $\lambda$, then the same is true for $d_{i}^{(1)}$ for $i > 1$ as well. We did not assume such an independence from $\lambda$ in advance.

5.2.2. Generalized eigenvectors of $F$—second approach

A more illuminating way to generalized eigenvectors of a $k$-companion matrix is to utilize the fact that they obey a difference equation with respect to the upper index. Denote $Z_{j,r} = (x_{j+rk}, \ldots, x_{j+(r+1)k})'$, where $j = 1, \ldots, k$. Then

$$Z_{j,r-1} = \lambda Z_{j,r} + \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix} Z_{j,r}.$$ 

Let

$$J_{0} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$ 

Then

$$Z_{j,l+1-r} = (\lambda I + J_{0})Z_{j,l+1-r+1},$$

where $l$ can be arbitrary but we assume for concreteness (as in the previous subsection) that it is such that $j + (l + 1)k \leq m$ while $j + (l + 2)k > m$, i.e.,

$$Z_{j,l+1} = (c_{j}^{(s)}, \ldots, c_{j}^{(1)})'.$$  

(5.11)

Hence

$$Z_{j,l+1-r} = (\lambda I + J_{0})^{r} Z_{j,l+1}.$$

But $(\lambda I + J_{0})$ is a Jordan block. Suppose for the moment that $\lambda \neq 0$. Then

$$ (\lambda I + J_{0})^{r} = \begin{pmatrix} \lambda^{r} & (r_{1}) \lambda^{r-1} & \cdots & (r_{s-1}) \lambda^{r-s+1} \\ 0 & \lambda^{r} & \cdots & (r_{s-2}) \lambda^{r-s+2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda^{r} \end{pmatrix}.$$  

(5.12)

(cf. [7, p. 311]). Hence
Writing down the rows of the last equation and using Eq. (5.11) we obtain

\[
\begin{align*}
Z_{j,l+1-r} &= 
\begin{pmatrix}
\lambda^r & (x_j^{(s-r+1)})^{(s-r+1)} & \cdots & (x_j^{(s-1)})^{(s-1)} & (x_j^{(s)})^{(s)} \\
0 & \lambda^r & \cdots & (x_j^{(s-2)})^{(s-2)} & (x_j^{(s-1)})^{(s-1)} \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & \lambda^r & 0 \\
\end{pmatrix}
Z_{j,l+1}.
\end{align*}
\]

(5.13)

This derivation holds also for \(\lambda = 0\). In that case \((\lambda I + J_0)^{r}\) is the zero matrix for \(r \geq s\). For \(r < s\) it is a matrix with ones on the \(r\)th super-diagonal and zeroes elsewhere. The modification of the equations for \(x^{(i)}\) are then obvious. Indeed, in that case \((\lambda I + J_0)^{r}\) is a matrix with 1s on the \(r\)th super-diagonal. Then

\[
Z_{j,l+1-r} =
\begin{cases}
0, & r \geq s, \\
N_r Z_{j,l+1}, & 1 \leq r \leq s - 1.
\end{cases}
\]

Hence, \(Z_{j,l+1-r} = (c_j^{(s-r)} , \ldots , c_j^{(1)} , 0 , \ldots , 0)^t\) for \(r = 1 , \ldots , s - 1\) (there are \(r\) zeroes).

Eq. (5.13) provides another formula for the quantities \(d^{(i)}_{j}\) introduced in the previous subsection. Indeed, in that subsection we expressed \(x^{(i)}_{j}\) in the form \(d^{(i)}_{j} \lambda^{s-i+1}\).

Hence, putting

\[
c_j^{(r)} = x^{(r)}_{j+(l+1)k} = d^{(r)}_{j+(l+1)k} \lambda^{s-i+1}\]

and

\[
x^{(r)}_{j+(l+1-r)k} = d^{(r)}_{j+(l+1-r)k} \lambda^{s-i+1}\]

in (5.13), its left-hand side becomes

\[
\begin{pmatrix}
\lambda^r & (x_j^{(s-r+1)})^{(s-r+1)} & \cdots & (x_j^{(s-1)})^{(s-1)} & (x_j^{(s)})^{(s)} \\
0 & \lambda^r & \cdots & (x_j^{(s-2)})^{(s-2)} & (x_j^{(s-1)})^{(s-1)} \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & \lambda^r & 0 \\
\end{pmatrix}
\]
while its right-hand side becomes
\[
\begin{pmatrix}
\lambda^r & \binom{r}{1} \lambda^{r-1} & \ldots & \binom{r}{s-1} \lambda^{r-s+1} \\
0 & \lambda^r & \ldots & \binom{r}{s-2} \lambda^{r-s+2} \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & \lambda^r
\end{pmatrix}
\begin{pmatrix}
d(s)_{j} + (l+1)k \lambda^l j_{l+1}k^{s+1} \\
d(s-1)_{j} + (l+1)k \lambda^l j_{l+1}k^{s-1+1} \\
\vdots \\
d(1)_{j} + (l+1)k \lambda^l j_{l+1}k
\end{pmatrix}
\]

\[
= \lambda^l j_{l+1}k^{r-s+1} + r
\begin{pmatrix}
1 & 0 & \ldots & 0 \\
0 & \lambda & \ldots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & \lambda^{r-1}
\end{pmatrix}
\begin{pmatrix}
1 & \binom{r}{1} & \ldots & \binom{r}{s-1} \\
0 & 1 & \ldots & \binom{r}{s-2} \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & 1
\end{pmatrix}
\times
\begin{pmatrix}
d(l)_{j+1}k \\
d(l-1)_{j+1}k \\
\vdots \\
d(1)_{j+1}k
\end{pmatrix}
\]

Now using the fact that \( l_{j-rk} - l_{j} = r \), we get
\[
\begin{pmatrix}
d(l)_{j+1-rk} \\
d(l-1)_{j+1-rk} \\
\vdots \\
d(1)_{j+1-rk}
\end{pmatrix}
= \begin{pmatrix}
1 & \binom{r}{1} & \ldots & \binom{r}{s-1} \\
0 & 1 & \ldots & \binom{r}{s-2} \\
\vdots & \ddots & \ddots & \vdots \\
0 & \ldots & 0 & 1
\end{pmatrix}
\begin{pmatrix}
d(l)_{j+1}k \\
d(l-1)_{j+1}k \\
\vdots \\
d(1)_{j+1}k
\end{pmatrix}
\]

In particular, when \( r = 1 \) the binomial coefficients are zero except for \( \binom{r}{1} = 1 \), which is consistent with Eqs. (5.8).

### 5.2.3. Generalized eigenvectors of \( F \)—parameterization

Of course, the Jordan chains for a matrix are not unique. For the purpose of parameterization it is useful to know when two sets of Jordan chains correspond to the same multi-companion matrix. Consider a Jordan chain (5.5). Straightforward calculations show that it can be replaced by, e.g., \( x(1), x(2) + \alpha x(1), \ldots, x(s) + \alpha x(s-1) \), or
\[
x(1), x(2), x(3) + \alpha x(1), \ldots, x(s) + \alpha x(s-2), \ldots, x(s) + \alpha x(s-2).
\]

In fact, these variations are particular cases of a general result. Denote by \( X \) a Jordan chain of length \( s \). Then
\[
FX = XJ,
\]
where \( X \) is \( m \times s \), \( J \) is \( s \times s \). Let \( K \) be a non-singular upper-triangular \( s \times s \) Toeplitz matrix. Then

\[
FXK = XJK = XKJ,
\]
since \( KJ = JK \). Hence \( XK \) is a Jordan chain too. In fact, every matrix which commutes with \( J \) is upper-triangular Toeplitz, see [7, 12.4, pp. 416–418]. Moreover, \( f(J) \) itself is upper-triangular Toeplitz for each \( (s - 1) \) times differentiable function \( f \), whose entry in the \( j \)th column of the first row is \( f(j^{(-1)}(\lambda))/(j - 1)! \) (see [7, p. 311]).

This last formula can be used to express each upper-triangular Toeplitz matrix commuting with \( J \) as a function of \( J \). Hence, the following auxiliary result holds.

**Lemma 5.1.** An \( s \times s \) matrix \( K \) commutes with a Jordan block \( J \) if and only if \( K = f(J) \) for some \( s - 1 \) differentiable function \( f \).

A canonical chain can be obtained by imposing some additional condition such as that some row of \( X \) with non-zero first element be \( (1, 0, \ldots, 0) \). When there are more than one eigenvectors corresponding to \( \lambda \), it is necessary to adopt some canonical choice for them too.

Hence, the generalized eigenvectors of a \( k \)-companion \( m \times m \) matrix \( F \) contain \((k - 1)m\) free parameters. If we add \( m \) for the number of eigenvalues, we obtain a parameterization with \( km \) parameters, a number equal to that of the non-trivial elements of such matrices.

Another parameterization of a \( k \)-companion matrix can be obtained from the results of Section 3 as a product of companion matrices. Although such a factorization does not always exist it may happen that the problem under study itself restricts the study to factorizable matrices only (see Section 6.1). In that case one may be interested in a number of related matrices. More specifically, let \( F_k = A_k \cdots A_1 \), with companion matrices \( A_i \). Consider also the rotated products \( F_1 = A_1A_k \cdots A_2, F_2 = A_2A_1A_k \cdots A_3, \) etc. If \( F_kX = XJ \), then \( F_1(A_1X) = A_1A_k \cdots A_1X = A_1F_kX = (A_1X)J \). Hence, the columns of \( A_1X \) are the generalized eigenvectors of \( F_1 \), provided that \( A_1 \) is non-singular. If \( A_1 \) is singular, then \( A_1X \) is also singular and cannot be a Jordan basis for \( F_1 \). However this does not complicate the relationship between the Jordan bases of \( F_k \) and \( F_1 \). Indeed, in that case the rank of \( A_1X \) is exactly \( m - 1 \). Furthermore, from Theorem 4.2 it follows that if \( A_1 \) is singular then the entire last column of \( F_k \) is zero, otherwise its decomposition into companion matrices would be impossible. Hence, the vector \((0, \ldots, 0, 1)'\), being an eigenvector of \( F_k \), can be put as a column of \( X \). Clearly, \( A_1(0, \ldots, 0, 1)' = 0 \), which shows that the other columns of \( A_1X \) are linearly independent. Combining everything said in this paragraph we see that only one column of \( A_1X \) should be changed to give a Jordan basis for \( F_1 \).

Note that in the case of companion matrices \( (k = 1) \) the \( c^{(1)}_1 \) are virtually unique since, as shown above, one can always choose \( c^{(1)}_1 = 1 \) and \( c^{(1)}_i = 0 \) for \( i > 1 \). The
generalized eigenvectors of companion matrices can be written in the form (see [7, p. 70]),

$$x_{ij}^{(i+1)} = \frac{1}{i!} \frac{d^i x_j}{d \lambda^i} = c_j \frac{l_j^i}{i!(l_j - i)!} \lambda^{l_j - i} = c_j \frac{l_j - i + 1}{i!} x_j^{(i)}$$

for $j = 1, \ldots, k$, with the remaining elements defined to obey (5.6). This is a particular case of (5.13). In the case of multi-companion matrices, i.e., $k > 1$, such a choice can be made for some $j$, but in general not for all $j \in [1,k]$ at the same time.

6. Applications

Here we outline two models where multi-companion matrices appear and discuss how the results about such matrices may be useful. Exposition of specific results requires a lot of background information from stochastic processes and time series analysis and will be published elsewhere.

6.1. Periodic autoregression

A process $\{X_t\}$ is said to be periodically correlated with period $d$ (or $d$-periodically correlated) if its mean $\mu_t = EX_t$ and autocovariance function $K(t,k) = E(X_t - \mu_t)(X_{t-k} - \mu_{t-k})$ exist and are $d$-periodic in $t$, see, e.g., [3]. We suppose below for simplicity that $\mu_t = 0$. A periodic autoregressive process is a periodically correlated process which satisfies a stochastic difference equation of the form

$$X_t = \sum_{i=1}^{p_t} \phi_{ti} X_{t-i} + \epsilon_t, \quad (6.1)$$

where $p_t$ are orders of the model, $\phi_{ti}, i = 1, \ldots, p_t$, coefficients, and $\epsilon_t$ is a periodic white noise with variance $\sigma^2_t$. All parameters $p_t, \phi_{ti}, i = 1, \ldots, p_t, \sigma^2_t$, are $d$-periodic with respect to the time index $t$.

The usual stationary autoregression model can be obtained from (6.1) by putting $d = 1$. In that case the parameters of the model do not depend on $t$ and the polynomial $1 - \sum \phi_t z^t$ or its companion matrix can be used to study the process, e.g. its spectrum. In the general case, $d > 1$, the polynomials $\phi_t(z) = 1 - \sum \phi_{ti} z^t$ cannot be used with the same success but a natural generalization exists (see [2]). Let $m = \max(d, p_1, \ldots, p_d)$, $Z_t = (X_t, X_{t-1}, \ldots, X_{t-m+1})^T$. Define the companion matrices $A_t = C[\phi_{t1}, \phi_{t2}, \ldots, \phi_{tm}], t = 1, \ldots, d$, and the matrices $M_t = A_t \cdots A_{t-d+1}$. Then $Z_t = A_t Z_{t-1} + E_t, Z_t = M_t Z_{t-d} + \mu_t$, where $E_t$ and $\mu_t$ are uncorrelated with $Z_{t-1}$ and $Z_{t-d}$, respectively. Without loss of information we can take every $d$th element of the sequence $Z_t$, e.g. $Y_t = Z_{td}, t = \ldots, -1, 0, 1, \ldots$. The process $Y_t$ is multivariate stationary AR(1).
\[ Y_t = F Y_{t-1} + u_t, \quad (6.2) \]

where \( F = M_d \).

Eq. (6.2) can be used to give full description of the properties of the periodic autoregressive process \( \{X_t\} \), see [2].

The matrix \( F \) is multi-companion of order \( d \) and its decomposition into a product of companion matrices is \( A_d \times \cdots \times A_1 \). Some interesting properties of the periodic autoregression model can be derived using the results from Section 3.

Also, for given periodic white noise sequence \( \varepsilon_t \) and coefficients \( p_t, \phi_{t,i}, i = 1, \ldots, p_t \), Eq. (6.1) is a stochastic difference equation, for which the existence, uniqueness of periodically correlated solutions and their construction can be studied most successfully by the Jordan decomposition of the multi-companion matrix \( F \).

This decomposition has a direct interpretation in that the eigenvalues of \( F \) are poles of the spectra and cross-spectra of the series \( \{X_{j+d}\} \) of the individual seasons for \( j = 1, \ldots, d \). A necessary and sufficient condition for existence (and uniqueness) of a periodically correlated solution to Eq. (6.1) is that all eigenvalues of \( F \) have modulus different from 1 [2]. The solution is causal (expressed in terms of \( \varepsilon_{t-k}, k \geq 0 \), only) if all the moduli are less than 1. A more general class of non-stationarity can be obtained by allowing for roots with modulus 1. This leads to periodically integrated processes (see [3]) which generalize the familiar Box–Jenkins integrated models.

Knowledge of the Jordan form of multi-companion matrices provides a way to generate periodic models with specified properties by constructing the matrix \( F \) and then deriving the parameters of the model by factorizing \( F \) into a product of companion matrices. This can be useful for selection of appropriate models in simulation studies and in estimation of restricted models for the purpose of hypothesis testing (for example, to test periodic integration it is necessary to estimate a restricted model with the zero-hypothesis roots on the unit circle). Conditions of this kind are difficult to handle when working with the parameters \( \phi_{t,i} \) themselves.

### 6.2. Multivariate autoregression in continuous time

The continuous-time autoregression model of order \( p \) (CAR(\( p \))) for a vector process, \( x_t \), can be defined as follows. Let \( X_t = (x_{t-(p-1)}, \ldots, x_{t(0)})' \), where the upper index \( i \) denotes the \( i \)th derivative of \( x_t \) and \( x_{t(0)} = x_t \). Let also \( A_i, i = 1, \ldots, p, \) be square, \( d \times d \), matrices,

\[
A = \begin{pmatrix}
A_1 & \cdots & A_{p-1} & A_p \\
I_d & \cdots & I_d & 0 \\
& \ddots & & \\
& & I_d & \\
& & & 0
\end{pmatrix},
\quad \sigma = \begin{pmatrix}
\Sigma & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix},
\]

where \( \Sigma \) is some \( d \times d \) matrix.

The process \( \{x(t)\} \) is said to be CAR(\( p \)) if it obeys the following stochastic differential equation:
\[ dX_t = AX_t \, dt + \sigma dW_t, \quad (6.3) \]

where \( W_t \) is a standard multivariate Wiener process.

Eq. (6.3) with a general matrix \( A \) has been studied extensively, see e.g. [1]. The role of the spectral decomposition of a general matrix \( A \) (in the case \( \sigma \equiv I \)) for statistical inference is exemplified in [5, pp. 487–490].

A CAR\((p)\) model (6.3) is defined once the matrices \( A_i, i = 1, \ldots, p \) and \( \Sigma \) (or, equivalently, \( A \) and \( \sigma \)) are specified. These are the parameters of the CAR\((p)\) model. The properties of the CAR\((p)\) process \( X \) are determined by the eigenstructure of the \( d \)-companion matrix \( A \) in a straightforward way. We propose to parameterize the CAR\((p)\) model through the eigensystem of the associated multi-companion matrix, \( A \). This parameterization can be useful for estimation, simulation and statistical inference for the CAR\((p)\) model.

Such a parameterization would have been difficult and hardly possible to do without explicit expressions for eigenvalues and eigenvectors since \( A \) is a \( pd \times pd \) matrix with only \( pd \times p \) non-trivial elements. If general eigenvectors and eigenvalues are used as parameters, then not only the number of parameters would increase to \( pd \times pd \) but they would be related by complicated functional relationships. The former problem is serious enough while the latter poses significant estimation and identification problems.

Since \( A \) is \( d \)-companion our description of the eigen-system of multi-companion matrices makes this task straightforward. The results on Jordan decomposition of multi-companion matrices give the possibility to parameterize the model with free, functionally independent, parameters. The functional independence of the proposed parameters paired with their intimate relation with the model properties is especially valuable.

For example, in the scalar case, \( d = 1 \), \( A \) is companion matrix and the matrix exponent \( e^{At} \) can be written down explicitly. This fact has been exploited by Jones [4] to develop an algorithm for maximum likelihood estimation of the parameters of the univariate model \((d = 1)\). The results on multi-companion matrices allow for that procedure to be generalized to the multivariate case \((d > 1)\). Note that the case \( d > 1 \) involves additional issues not present in the scalar case.

Acknowledgements

The comments of the referees were very helpful in improving the paper.

Appendix A. Canonical form of \( \lambda I - F \)

Here are the details of the transformations discussed in Section 5.2.1. Recall that \( \lambda I - F \) is given by (5.1). The transformations outlined in the paragraphs before Eq. (5.3) can be carried out with the matrices
\[ K_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ I & & & \\ \vdots & & & \\ I & A^{(1)} & & I \end{pmatrix}, \quad K_i = \begin{pmatrix} I & 0 & \cdots & 0 \\ & I & \vdots & \\ & & & A^{(i)} \\ & & I & \vdots \\ & & & I \end{pmatrix}, \quad (A.1) \]

where \( A^{(i)} = \lambda^i I, i = 1, \ldots, s - 3, A^{(s-2)} = \lambda^{s-2} I_{k,t}, A^{(s-1)} = \lambda^{s-1} H_{t,k}, \) i.e., \( A^{(i)} \) equals \( \lambda^i I \), except for the two top block rows where the dimensions should be adjusted. Also, \( K_i \) (as a right factor) multiplies the \((s-i)\)th block column by \( \lambda^i \) of \( \lambda I - F \) and adds it to the last block column. Let \( K = K_1 K_2 \cdots K_{s-1} \). Then

\[
M = (\lambda I - F)K_1 K_2 \cdots K_{s-1}
= (\lambda I - F)K
= \begin{pmatrix} I_{k,t} + A_1 & \lambda I + A_2 & A_3 & \cdots & A_{s-1} & T(\lambda) \\ -I & \lambda I & 0 & \cdots & 0 & 0 \\ H_{t,k} \lambda & 0 & & \vdots & & \\ -I & \lambda I & 0 & \vdots & & \\ -I & \lambda I & 0 & \vdots & & \\ -I & 0 & \vdots & & & \end{pmatrix}
\]

where \( T(\lambda) \) is given by (5.3). The matrix \( K \) is unimodular and it is easy to check that

\[
K = \begin{pmatrix} I & 0 & \cdots & 0 & A^{(s-1)} \\ & I & \cdots & \vdots & \vdots \\ & & & \ddots & \vdots \\ & & & & I \\ & & & & A^{(1)} \end{pmatrix}, \quad K^{-1} = \begin{pmatrix} I & 0 & \cdots & 0 & -A^{(s-1)} \\ & I & \cdots & \vdots & \vdots \\ & & & \ddots & \vdots \\ & & & & I \\ & & & & -A^{(1)} \end{pmatrix}
\]

We now left-multiply \( M \) by \( K_1 \) to remove \( \lambda \) from the next to last block row of \( M \) (\( K_1 \) multiplies the last block row by \( \lambda \) and adds it to the block row above it). Proceeding similarly we can eliminate \( \lambda \) from the diagonal of \( M \) by

\[
L = \begin{pmatrix} I & I_{n^n} & H_{t,k} \lambda & \cdots & \cdots & \cdots & I \\ I & \ddots & \cdots & \cdots & \cdots & \cdots & I \\ \vdots & \ddots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \ddots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \ddots & \ddots & \vdots \\ I & \cdots & \cdots & \cdots & \cdots & \cdots & I \\ I & \cdots & \cdots & \cdots & \cdots & \cdots & I \\ I & \cdots & \cdots & \cdots & \cdots & \cdots & I \\ I & \cdots & \cdots & \cdots & \cdots & \cdots & I \\ I & \cdots & \cdots & \cdots & \cdots & \cdots & I \end{pmatrix}
\]
to get

\[
LM = \begin{pmatrix}
I_{k,t} + A_1 & \lambda H_t + A_2 & A_3 & \cdots & A_{s-1} & T(\lambda)
\end{pmatrix}
\begin{pmatrix}
-I_t & 0_{t,k} & 0 & \cdots & 0 & 0 \\
-I & 0 & \ddots & \vdots & \vdots & \\
-I & \ddots & \ddots & \vdots & \vdots & \\
\ddots & \ddots & \ddots & 0 & 0 & \\
-I & 0 & & & & \\
-I & 0 & & & & \\
\end{pmatrix}.
\]

(A.2)

The determinant of \(T(\lambda)\) gives the characteristic polynomial of \(F\). We now left-multiply \(LM\) by

\[
L_1 = \begin{pmatrix}
I & I_{k,t} + A_1 & \lambda H_t + A_2 & A_3 & \cdots & A_{s-1} \\
-I_t & 0 & 0 & \cdots & 0 \\
-I & 0 & \ddots & \vdots & \vdots & \\
-I & \ddots & \ddots & \vdots & \vdots & \\
\ddots & \ddots & \ddots & 0 & 0 & \\
-I & 0 & & & & \\
\end{pmatrix}
\]

and obtain

\[
\begin{pmatrix}
I & I_{k,t} + A_1 & \lambda H_t + A_2 & A_3 & \cdots & A_{s-1} \\
-I_t & 0 & 0 & \cdots & 0 \\
-I & 0 & \ddots & \vdots & \vdots & \\
-I & \ddots & \ddots & \vdots & \vdots & \\
\ddots & \ddots & \ddots & 0 & 0 & \\
-I & 0 & & & & \\
\end{pmatrix}
\times
\begin{pmatrix}
I_{k,t} + A_1 & \lambda H_t + A_2 & A_3 & \cdots & A_{s-1} & T(\lambda) \\
-I_t & 0_{t,k} & 0 & \cdots & 0 & 0 \\
-I & 0 & \ddots & \vdots & \vdots & \\
-I & \ddots & \ddots & \vdots & \vdots & \\
\ddots & \ddots & \ddots & 0 & 0 & \\
-I & 0 & & & & \\
\end{pmatrix}
= \begin{pmatrix}
0_k & T(\lambda) \\
I_{m-k} & 0_{m-k,k} \\
\end{pmatrix}.
\]
If we left-multiply the last matrix by 
\[ L_2 = \begin{pmatrix} 0_{m-k,k} & I_{m-k} \\ I_k & 0_{k,m-k} \end{pmatrix}, \]
we get 
\[ \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & T(\lambda) \end{pmatrix} = L_2 L_1 L(\lambda I - F) K. \]

Let \( T(\lambda) = P_T(\lambda) D_T(\lambda) Q_T(\lambda) \) be the canonical decomposition of \( T(\lambda) \) with uni-

\text{modal matrices} \( P_T(\lambda) \) and \( Q_T(\lambda) \). Let

\[ P(\lambda) = \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & P_T(\lambda) \end{pmatrix}, \quad D(\lambda) = \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & D_T(\lambda) \end{pmatrix}, \]

\[ Q(\lambda) = \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & Q_T(\lambda) \end{pmatrix}. \]

Then 
\[ \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & T(\lambda) \end{pmatrix} = \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & P_T(\lambda) \end{pmatrix} \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & D_T(\lambda) \end{pmatrix} \]

\[ \times \begin{pmatrix} I_{m-k} & 0_{m-k,k} \\ 0_{k,m-k} & Q_T(\lambda) \end{pmatrix} = P(\lambda) D(\lambda) Q(\lambda). \]

Hence,
\[ P(\lambda) D(\lambda) Q(\lambda) = L_2 L_1 L(\lambda I - F) K, \]
and therefore,
\[ \lambda I - F = L_1^{-1} L_2^{-1} P(\lambda) D(\lambda) Q(\lambda) K^{-1}, \]
where \( D(\lambda) \) has the required form for a canonical form and the remaining matrices are unimodal matrix polynomials (some of order zero). Therefore \( D(\lambda) \) is indeed the canonical form of \( \lambda I - F \).

**Appendix B. Canonical form of $\lambda I - F$ (second manipulation)**

The element \( \lambda I \) in block column \( i \) of the matrix \( \lambda I - F \) can be made zero by replacing block column \( i \) by its sum with block column \( i - 1 \) multiplied by \( \lambda \). This operation can be repeated for each \( i \in [2, s] \). To avoid introducing new non-zero elements it is important to perform these transformations starting with \( i = 2 \) and increasing it by one each time. In matrix form the transformations can be written as follows. Let \( a_1 = I_{k,i} + A_1, \ a_2 = (I_{k,i} + A_1) H_{s,k} + (\lambda H_s + A_2). \) Then
\[
\begin{pmatrix}
\alpha_1 & \alpha_2 & A_3 & \ldots & A_{s-1} & A_s \\
-I & 0 & 0 & \ldots & 0 & 0 \\
-I & \lambda I & \ddots & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
-I & \lambda I & 0 & \ldots & -I & \lambda I \\
\end{pmatrix}
\begin{pmatrix}
I_{k,t}\lambda + A_1 & \lambda H_t + A_2 & A_3 & \ldots & A_{s-1} & A_s \\
-I & H_{k,k}\lambda & 0 & \ldots & 0 & 0 \\
-I & \lambda I & \ddots & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
-I & \lambda I & 0 & \ldots & -I & \lambda I \\
\end{pmatrix}
\begin{pmatrix}
I_{k,k}\lambda & 0 & 0 & \ldots & 0 \\
I & 0 & 0 & \ldots & 0 \\
I & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
I & \ddots & \ddots & \ddots & 0 \\
\end{pmatrix},
\]

i.e., multiplication of \( \lambda I - F \) by the second matrix on the right-hand side above performs the required operation and transforms the block entry at position (2, 2) to a zero matrix. Now multiplication of the second block column by \( \lambda \) and addition to the third block column nullifies the element \( \lambda I \) in the latter. No new non-zero entries turn up,

\[
\begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \ldots & A_{s-1} & A_s \\
-I & 0 & 0 & \ldots & 0 & 0 \\
-I & 0 & \ddots & \vdots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
-I & \lambda I & 0 & \ldots & -I & \lambda I \\
\end{pmatrix}
\begin{pmatrix}
I & 0 & 0 & \ldots & 0 \\
I & \lambda I & 0 & \ldots & 0 \\
I & 0 & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
I & \ddots & \ddots & \ddots & 0 \\
\end{pmatrix},
\]

where \( \alpha_3 = ((I_{k,t}\lambda + A_1)H_{k,k}\lambda + (\lambda H_t + A_2))\lambda + A_3. \)
Proceeding in the same way at step \( i - 1 \) we nullify \( \lambda I \) in block column \( i \). The element in the first block row is also changed from \( A_i \) to \( \alpha_i \) but the remaining elements in block column \( i \) remain zero.

The entire transformation therefore can be written as

\[
R_s = \begin{pmatrix}
I & H_{i,k}\lambda & 0 & \cdots & 0 \\
I & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \vdots \\
I & \cdots & \cdots & \cdots & I
\end{pmatrix}
\begin{pmatrix}
I & 0 & 0 & \cdots & 0 \\
I & \lambda I & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \vdots \\
I & \cdots & \cdots & \cdots & I
\end{pmatrix}
\begin{pmatrix}
I & 0 & \cdots & 0 \\
I & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & I \\
I & \cdots & \cdots & I
\end{pmatrix}
\times \cdots \times
\begin{pmatrix}
I & 0 & \cdots & 0 \\
I & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & I \\
I & \cdots & \cdots & I
\end{pmatrix}
\begin{pmatrix}
I & 0 & 0 & \cdots & 0 \\
I & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & 0 & 0 \\
\vdots & \ddots & \ddots & I & \lambda I \\
I & \cdots & \cdots & I & \lambda I
\end{pmatrix}
\begin{pmatrix}
I & 0 & \cdots & 0 \\
I & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & I \\
I & \cdots & \cdots & I
\end{pmatrix}
\end{pmatrix}
\]

Hence we have

\[
(\lambda I - F)R_s = \begin{pmatrix}
\alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_{s-1} & \alpha_s \\
-I & 0 & 0 & \cdots & 0 & 0 \\
-I & 0 & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\
-I & 0 & 0 & \cdots & 0 & 0 \\
-I & 0 & 0 & \cdots & 0 & 0
\end{pmatrix},
\]

where \( \alpha_i = \alpha_i(\lambda) \) is a polynomial in \( \lambda \) of degree \( i \) for \( i = 1, \ldots, s \). Also, up to a multiplication by \(-1\), the matrix on the right-hand side is a matrix polynomial which itself is \( k \)-companion.

Transformations analogous to that on \( LM \) (see Eq. (A.2)) can be used to complete the derivation of the canonical form of \( \lambda I - F \).
References


