ECONOMETRIC INSTITUTE

UPPER TRIANGULARIZATION OF MATRICES BY LOWER TRIANGULAR SIMILARITIES

H. BART AND P.S.M. KOP JANSEN

REPORT 8711/B

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM - THE NETHERLANDS

UPPER TRIANGULARIZATION OF MATRICES BY LOWER TRIANGULAR SIMILARITIES

H. Bart P.S.M. Kop Jansen

ABSTRACT

This paper is concerned with the following questions. Given a square matrix A, when does there exist an invertible lower triangular matrix L such that $L^{-1}AL$ is upper triangular ? And if so, what can be said about the order in which the eigenvalues of A may appear on the diagonal of $L^{-1}AL$? The motivation for considering these questions comes from systems theory. In fact they arise in the study of complete factorizations of rational matrix functions. There is also an intimate connection with the problem of complementary triangularization of pairs of matrices discussed in [4].

0. INTRODUCTION

Let A be an m×m (complex) matrix. As is well-known, A can be brought into upper triangular form by a similarity transformation. In other words, there exists an invertible m×m matrix S such that $S^{-1}AS$ is upper triangular. Here we are interested in the situation where S can be chosen to be lower triangular. We then say that A admits upper triangularization by a lower triangular similarity.

First let us give some motivation for considering this property. Recall from systems theory that a complete factorization of a rational n×n matrix function $W(\lambda)$ is a factorization of the form

$$(0.1)\qquad W(\lambda) = \left(I_n + \frac{1}{\lambda-\alpha_1}R_1\right) \cdots \left(I_n + \frac{1}{\lambda-\alpha_m}R_m\right),$$

where $R_1,\dots,R_m$ are n×n matrices of rank 1 and a certain minimality condition is satisfied (no "pole/zero cancellations"). Such factorizations do not always exist and if they do one cannot always choose the order of the poles $\alpha_1,\dots,\alpha_m$ of $W(\lambda)$ at will.

The questions arising naturally in this context can be most efficiently answered in terms of realizations. A realization of $W(\lambda)$ is an expression of the form

$$(0.2)\qquad W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B,$$

where A is an m×m matrix, B is an m×n matrix and C is an n×m matrix. Systems theory tells us that such a realization exists whenever $W(\lambda)$ is analytic at ∞ and $W(\infty) = I_n$. It turns out that under certain additional assumptions

$W(\lambda)$ admits a complete factorization of the form (0.1) if and only if the matrix $A-BC$ associated with the realization (0.2) admits upper triangularization by a lower triangular similarity. For details (including a brief review of the necessary background material from systems theory), see Section 4.

Next, let us describe the contents of the other three sections constituting the paper. We shall do this by stating (simplified versions of) the main results and commenting on them. All the time A will be a square complex $m \times m$ matrix.

THEOREM 0.1. *If A is diagonable, then A admits upper triangularization by a lower triangular similarity.*

This theorem is proved in Section 1. The argument is based on the observation that A admits upper triangularization by a lower triangular similarity if and only if there exists an (invertible) matrix S such that S has non-vanishing leading principal minors and $S^{-1}AS$ is upper triangular. Section 1 also contains an analysis of the 2×2 case.

THEOREM 0.2. *Let $A = Z + bc^T$, where Z is a lower triangular non-derogatory matrix, b is a cyclic vector for Z and c is a cyclic vector for the transpose Z^T of Z. Then A admits upper triangularization by a lower triangular similarity.*

This theorem is proved in Section 2. The general theme of Section 2 is the relationship between the problem studied in the present paper and the issue of *simultaneous reduction to complementary triangular forms* discussed in [4]. It turns out that the conditions of Theorem 0.2 imply that A and Z admit simultaneous reduction to complementary triangular forms. This means that there exists an invertible matrix S such that $S^{-1}AS$ is upper triangular and $S^{-1}ZS$ is lower triangular. Further analysis yields that S can be chosen to be lower triangular (see Proposition 2.1).

It is illuminating to note that the problem of simultaneous reduction to complementary triangular forms is also intimately connected to that of complete factorization of a rational matrix function. As a matter of fact, the two are practically equivalent. Indeed, if $W(\lambda)$ is given by (0.2) and m is taken as small as possible (i.e., the realization (0.2) is minimal), then $W(\lambda)$ admits a complete factorization of the form (0.1) if and only if A and A-BC admit simultaneous reduction to complementary triangular forms. For further details, see [4].

THEOREM 0.3. *If A is a first companion matrix, then A admits upper triangularization by a lower triangular similarity.*

A first companion matrix has the determining entries in the last row. Second companion matrices are the transposes of first companion matrices. So a second companion matrix has the determining entries in the last column.

THEOREM 0.4. *Let A be a second companion matrix. Then A admits upper triangularization by a lower triangular similarity if and only if A has at most one eigenvalue zero (counted according to algebraic multiplicity).*

These theorems are proved in Section 3 by explicitly defining lower triangular similarities that bring the companion matrices in upper triangular form. Two corollaries are given. They are based on the simple observation that if A admits upper triangularization by a lower triangular similarity and if Z is a lower triangular matrix, then A and Z admit simultaneous reduction to complementary triangular forms.

The theorems stated above are concerned with the existence of a lower triangular matrix L such that $L^{-1}AL$ is upper triangular. For Theorems 0.2-0.4 the information is completed by determining the order in which the eigenvalues of A may appear on the diagonal of $L^{-1}AL$. In connection with Theorem 0.1, an example is given to show that in general this order cannot be chosen at will.

A few remarks about notation and terminology: All matrices to be considered have complex entries. The n×n identity matrix is denoted by I_n, or simply I. Whenever this is convenient, matrices are identified with linear operators. The null space of a matrix (operator) M is denoted by Ker M. We use M^T for the transpose of M. The symbol [] stands for "end of proof" or "end of example".

1. PRELIMINARIES AND FIRST RESULTS

Let A be an m×m matrix. We say that A admits <u>upper triangularization by a lower triangular similarity</u> if there exists an invertible lower triangular m×m matrix L such that $L^{-1}AL$ is upper triangular.

Given a matrix A with this property, one can build others. Indeed, if A admits upper triangularization by a lower triangular similarity, then so does

(i) $T^{-1}AT$, where T is any invertible lower triangular m×m matrix,

(ii) p(A), where p is an arbitrary polynomial.

In (ii) one can also take an analytic function p defined on a neighbourhood of the spectrum of A. If A_1 and A_2 admit upper triangularization by a lower triangular similarity, then so does

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}.$$

Of course this result can be extended to direct sums involving more than two matrices.

Next we present a first characterization and some related material.

THEOREM 1.1. *The* $m \times m$ *matrix* A *admits upper triangularization by a lower triangular similarity if and only if there exists an invertible* $m \times m$ *matrix* S *such that* S *has non-vanishing leading principal minors and* $S^{-1}AS$ *is upper triangular*.

Proof. Suppose such an S exists. The assumption on the principal leading minors of S implies that S can be factorized as $S = LU$, where L is an invertible lower triangular matrix and U is an invertible upper triangular matrix (cf. [12, Section 2.10]). Now $L^{-1}AL = U(S^{-1}AS)U^{-1}$ is upper triangular. This proves the if part of the theorem. The only if part is a triviality: take $S = L$, where L is an invertible lower triangular matrix such that $L^{-1}AL$ is upper triangular. []

Theorem 1.1. can be used to give a quick proof of Theorem 0.1.

Proof of Theorem 0.1. Let U be an invertible $m \times m$ matrix such that $U^{-1}AU$ is a diagonal matrix. By multiplying U from the right with an appropriate permutation matrix Π, one gets a matrix $U\Pi$ with non-vanishing principal leading minors. Put $S = U\Pi$. Then $S^{-1}AS = \Pi^{-1}(U^{-1}AU)\Pi$ is again a diagonal matrix. In particular $S^{-1}AS$ is upper triangular. Now apply Theorem 1.1. []

By the diagonal of a matrix $K = [k_{ij}]_{i,j=1}^{m}$ we mean the ordered m-tuple $(k_{11},\dots,k_{mm})$. If K is a triangular matrix, then the diagonal of K contains the eigenvalues of K counted according to algebraic multiplicity. In connection with Theorem 0.1, the following question comes up. Can one choose the diagonal of $L^{-1}AL$ at will ? The following example shows that in general the answer is negative.

EXAMPLE 1.3. Let

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

Then A is diagonable and so Theorem 0.1 applies. In fact A can be brought even into diagonal form by a lower triangular similarity. For instance, if

$$L_1 = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix},$$

then L_1 is lower triangular and

$$L_1^{-1}AL_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The diagonal of $L_1^{-1}AL_1$ is (0,1). It is easy to verify that there does not exist an invertible lower triangular matrix L for which $L^{-1}AL$ is upper triangular with diagonal (1,0). []

We conclude this section by considering the 2×2 case.

THEOREM 1.4. *Let* A *be a* 2×2 *matrix, and write*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \qquad H = \begin{bmatrix} a_{11}-a_{22} & 2a_{12} \\ 2a_{21} & a_{22}-a_{11} \end{bmatrix}.$$

The following statements are equivalent:

(i) A *admits upper triangularization by a lower triangular similarity*,

(ii) A *is diagonable or* $a_{12} \neq 0$,

(iii) rank $H \neq 1$ *or* $a_{12} \neq 0$.

Proof. Put

$$\alpha = \frac{a_{11} + a_{22}}{2}.$$

Then $H = 2(A-\alpha I)$, and so A admits upper triangularization by a lower triangular similarity if and only if the same is true for H.

The discriminant of the quadratic polynomial $\det(\lambda I-A)$ is equal to $-\det H$. Thus A has two different eigenvalues if and only if rank H = 2. Clearly, A is a scalar multiple of the identity if and only if rank H = 0. So rank $H \neq 1$ implies that A is diagonable. This proves that (iii) implies (ii).

Assume (ii) is satisfied. If A is diagonable, then (i) holds by Theorem 0.1. In the case when A is non-diagonable (hence rank H = 1) and $a_{12} \neq 0$, we put

$$L = \begin{bmatrix} 2a_{12} & 0 \\ a_{22}-a_{11} & 2a_{12} \end{bmatrix}.$$

Then L is lower triangular and invertible. A straightforward computation, based on the identity det H = 0, yields

$$L^{-1}HL = \begin{bmatrix} 0 & 2a_{12} \\ 0 & 0 \end{bmatrix}.$$

It follows, that (i) is satisfied.

Finally, suppose rank H = 1 and $a_{12} = 0$. Then $a_{11} = a_{22}$, and H has the form

$$H = \begin{bmatrix} 0 & 0 \\ 2a_{21} & 0 \end{bmatrix}$$

with $a_{21} \neq 0$. But then H is non-diagonable. Since H is lower triangular too, we may conclude that H does not admit upper triangularization by a lower triangular similarity. So (i) implies (iii). []

2. THE CONNECTION WITH COMPLEMENTARY TRIANGULAR FORMS

In this section we make the connection with [4], where complementary triangular forms of pairs of matrices are investigated. Recall that two m×m matrices A and Z are said to admit <u>simultaneous reduction to complementary triangular forms</u> if there exists an invertible m×m matrix S such that $S^{-1}AS$ is upper triangular and $S^{-1}ZS$ is lower triangular.

We shall need (part of) the following proposition. An m×m matrix Z is called <u>non-derogatory</u> if all its eigenvalues have geometric multiplicity 1. In other words, if $\mathrm{rank}(Z-\zeta I_m) \geq m-1$ for all complex ζ.

PROPOSITION 2.1. <u>Let</u> Z <u>be a lower triangular</u> m×m <u>matrix. The following two statements are equivalent</u>:

(i) Z <u>is non-derogatory</u>,

(ii) <u>Every invertible</u> m×m <u>matrix</u> S <u>with the property that</u> $S^{-1}ZS$ <u>is lower triangular with the same diagonal as</u> Z, <u>is lower triangular</u>.

<u>Proof</u>. First we shall prove that (i) implies (ii). In fact we shall establish a slightly more general result: Let Z, Y and S be m×m matrices, and assume that Z is non-derogatory, Y and Z are lower triangular with the same diagonal, and ZS = SY. Then S is lower triangular. The argument is as follows.

Let $e \in \mathbb{C}^m$ have all coordinates zero except for the last one, which is equal to 1. Then $Z(Se) = S(Ye) = \mu Se$, where μ is the last element on the diagonal of Y. Our assumptions imply that μ is the last element on the diagonal of Z too. Clearly e is an eigenvector of Z corresponding to the eigenvalue μ. Since Z is non-derogatory, the eigenspace of Z corresponding to the eigenvalue μ has dimension 1. So Se is a scalar multiple of e. This means that all entries in the last column of S vanish, except perhaps for the last one.

Let Z_o be the (m-1)×(m-1) matrix obtained from Z by striking out the last column and the last row of Z. Define Y_o and S_o in the same way. Then $S_oY_o = Z_oS_o$. Since Z is lower triangular, we have that

$$\operatorname{rank}(Z-\zeta I_m) \leqq 1 + \operatorname{rank}(Z_o-\zeta I_{m-1}), \qquad \zeta \in \mathbb{C}.$$

Hence Z_o is again non-derogatory. This part of the proof can now be finished by an induction argument.

Next we show that (ii) implies (i). Suppose (i) is not satisfied. So the (lower triangular) matrix Z has an eigenvalue ζ such that $\dim \operatorname{Ker}(Z-\zeta I_m) \geqq 2$. Without loss of generality, we may assume that $\zeta = 0$. So $\dim \operatorname{Ker} Z \geqq 2$. Let Z_1, respectively Z_m, be the matrix obtained from Z by striking out the first row and the first column, respectively the last row and the last column, of Z.

We may also assume that Ker Z_1 and Ker Z_m are 1-dimensional. Indeed, one can reach this situation by removing appropriate columns and rows from Z.

Under these circumstances the first column of Z depends linearly on the other columns. Likewise, the last row of Z depends linearly on the other rows. In particular the first and last element in the diagonal of Z are equal to 0. It is now easy to find an invertible lower triangular m×m matrix L such that $L^{-1}ZL$ is lower triangular with the same diagonal as Z, while the entries in the first row, the last row, the first column and the last column of $L^{-1}ZL$ are all zero. Define the m×m matrix R by

$$R = \begin{bmatrix} 0 & 0 & & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & 1 & 0 & & & & & 0 \\ \cdot & & \cdot & & & & & \cdot \\ \cdot & & & & \cdot & & & \cdot \\ \cdot & & & & & \cdot & & \cdot \\ 0 & & & & & & 1 & 0 \\ 1 & 0 & \cdot & & \cdot & \cdot & 0 & 0 \end{bmatrix},$$

and put S = LR. Then $S^{-1}ZS$ is lower triangular with the same diagonal as Z, but S is not lower triangular. []

We are now ready to give a second characterization of the property studied in this paper. It involves an auxiliary lower triangular non-derogatory matrix Z.

COROLLARY 2.2. *Let A and Z be m×m matrices, and assume that Z is lower triangular and non-derogatory. Then A admits upper triangularization by a lower triangular similarity if and only if there exists an invertible m×m matrix S such that $S^{-1}AS$ is upper triangular, $S^{-1}ZS$ is lower triangular and $S^{-1}ZS$ has the same diagonal as Z.*

In the particular case when Z has only one eigenvalue of algebraic multiplicity m (for instance, Z is a single lower triangular m×m Jordan block), one can drop the requirement that $S^{-1}ZS$ and Z have the same diagonal.

The if part of Corollary 2.2 is an immediate consequence of Proposition 2.1. We shall say something about the only if part at the end of the section.

THEOREM 2.3. *Suppose* A *can be written as* A = Z + R, *where* A *is a lower triangular* m×m *matrix*, rank R = 1, *and* A *and* Z *have no common eigenvalues*. *Then, given an ordering* $\alpha_1,\ldots,\alpha_m$ *of the eigenvalues of* A, *there exists an invertible lower triangular* m×m *matrix* L *such that* $L^{-1}AL$ *is upper triangular with diagonal* $(\alpha_1,\ldots,\alpha_m)$.

Proof. Since rank R = 1, we have

$$\operatorname{rank}(Z-\zeta I_m) \geqq -1 + \operatorname{rank}(A-\zeta I_m), \qquad \zeta \in \mathbb{C}.$$

By assumption A and Z have no common eigenvalues. So Z is non-derogatory.

Let $\zeta_1,\ldots,\zeta_m$ be the eigenvalues of Z in the order in which they appear on the diagonal of Z. According to [4, Theorem 7.2], there exists an invertible m×m matrix L such that

$$L^{-1}AL = \begin{bmatrix} \alpha_1 & \alpha_1-\zeta_1 & \alpha_1-\zeta_1 & \cdot & \cdot & \cdot & \alpha_1-\zeta_1 \\ 0 & \alpha_2 & \alpha_2-\zeta_2 & \cdot & \cdot & \cdot & \alpha_2-\zeta_2 \\ \cdot & \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & \alpha_m \end{bmatrix},$$

and $L^{-1}ZL$ has diagonal $(\zeta_1,\ldots,\zeta_m)$. In particular, $L^{-1}AL$ is upper triangular with diagonal $(\alpha_1,\ldots,\alpha_m)$ and $L^{-1}ZL$ has the same diagonal as Z. Now apply Proposition 2.1. []

Note that this proof contains additional information about the form that $L^{-1}AL$ can take.

Next we shall see that the hypotheses of Theorem 2.3 amount to the same as those of Theorem 0.2. As a consequence, *the conclusion of Theorem* 2.3 *also holds under the hypotheses of Theorem* 0.2. In this sense Theorem 2.3 is just a refined version of Theorem 0.2. Recall that b is a *cyclic vector* of the $m\times m$ matrix Z if $b, Zb, \ldots, Z^{m-1}b$ are linearly independent. The matrix Z has cyclic vectors if and only if it is non-derogatory, and in that case the set of cyclic vectors is open and dense in $\mathbb{C}^m$.

PROPOSITION 2.4. *Let* A *be an* $m\times m$ *matrix. The following statements are equivalent*:

(i) A *can be written as* $A = Z + R$, *where* A *and* Z *have no common eigenvalues and* rank $R = 1$,

(ii) A *can be written as* $A = Z + bc^T$, *where* b *is a cyclic vector for* Z *and* c *is a cyclic vector for* Z^T.

Proof. Assume A can be written as in (i). Choose vectors b and c in $\mathbb{C}^m$ such that $R = bc^T$, and put

$$M = \mathrm{Ker}\begin{bmatrix} c^T \\ c^T Z \\ \vdots \\ c^T Z^{m-1} \end{bmatrix}.$$

Then M is an invariant subspace for Z (use the Cayley-Hamilton theorem). Since $A = Z + bc^T$, we have that M is an invariant subspace for A too. Clearly A and Z coincide on M. So if M is non-trivial, A and Z have a common eigenvalue. By assumption this is not the case. Hence M = (0). But this means that c is a cyclic vector for Z^T. Applying the same reasoning to A^T, Z^T and R^T, one sees that b is a cyclic vector for Z. This proves that (i) implies (ii).

Suppose that A can be written as in (ii). We need to show that Z and $A = Z + bc^T$ have no common eigenvalues. One way to do this, is to assume (without loss of generality) that Z has Jordan canonical form and to establish that $\det(\mu I_m - A) \neq 0$ whenever μ is an eigenvalue of Z. The details are quite tedious, and will be omitted.

Another approach uses (some of) the elements from systems theory reviewed in Section 4 below. Put

$$w(\lambda) = \frac{\det(\lambda I_m - A)}{\det(\lambda I_m - Z)}.$$

Then $w(\lambda)$ is a scalar rational function having the value 1 at ∞. Observe that

$$\begin{aligned} w(\lambda) &= \det(\lambda I_m - A).\det(\lambda I_m - Z) \\ &= \det(\lambda I_m - A)(\lambda I_m - Z)^{-1} \\ &= \det\left(I_m - bc^T(\lambda I_m - Z)^{-1}\right) \\ &= \det\left(1 - c^T(\lambda I_m - Z)^{-1} b\right). \end{aligned}$$

So $w(\lambda) = 1 - c^T(\lambda I_m - Z)^{-1}b$, and this realization of $w(\lambda)$ is minimal because b is a cyclic vector for Z and c is a cyclic vector for Z^T. On the other hand, the McMillan degree of $w(\lambda)$ is equal to m-k, where k is the number of common zeros of the monic scalar polynomials $\det(\lambda I_m - A)$ and $\det(\lambda I_m - Z)$. Hence k = 0, i.e., the matrices A and Z have no common eigenvalues. []

It is illuminating to consider the following special case. Let Z be a diagonal m×m matrix with m different complex numbers on the diagonal. Then b and c are cyclic vectors of $Z = Z^T$ if and only if all elements in b and c are non-zero. This means that $R = bc^T$ is a rank one matrix having no zero entries.

As announced earlier, we conclude this section with some remarks about the only if part of Corollary 2.2. The first is that this part of the corollary is trivial and holds even without the requirement that Z is non-derogatory. So we have the following result on simultaneous reduction to complementary triangular forms. _Let_ A _and_ Z _be_ m×m _matrices, and assume that_ A _admits upper triangularization by a lower triangular similarity and that_ Z _is lower triangular. Then there exists an invertible (lower triangular)_ m×m _matrix_ L _such that_ $L^{-1}AL$ _is upper triangular and_ $L^{-1}ZL$ _is lower triangular (with the same diagonal as_ Z). We shall use this simple observation in Section 3.

The next remark is that the condition of lower triangularity imposed on Z is not too restrictive. Indeed, if A and Z are arbitrary m×m matrices and T is any invertible m×m matrix, then A and Z admit simultaneous reduction to complementary triangular forms if and only if the same is true for $T^{-1}AT$ and $T^{-1}ZT$. By choosing T appropriately, one can always see to it that $T^{-1}ZT$ becomes lower triangular.

3. COMPANION MATRICES

In this section we study upper triangularization by lower triangular similarities of companion matrices. Recall that companion matrices are the building blocks of what is often referred to as the first natural normal form (cf. [6, Section VI.6] and [12, Section 7.6]).

First we consider a second companion m×m matrix A, i.e., A has the form

$$(3.1) \qquad A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{bmatrix},$$

where $a_0,\ldots,a_{m-1}$ are complex numbers. The next two results are inspired by the material contained in [4, Section 3].

LEMMA 3.1. _Let A be a second companion_ m×m _matrix as in_ (3.1), _and let_ $\alpha_1,\ldots,\alpha_m$ _be the eigenvalues of_ A. _Define_

$$(3.2) \qquad \tilde{A} = \begin{bmatrix} \alpha_1 & \alpha_1 & \alpha_1 & \cdots & \alpha_1 \\ 0 & \alpha_2 & \alpha_2 & \cdots & \alpha_2 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \alpha_m \end{bmatrix},$$

and introduce $S = [s_{kj}]_{k,j=1}^{m}$ _by stipulating that_

$$(3.3) \qquad \sum_{k=1}^{m} s_{kj}\lambda^{k-1} = \lambda^{j-1}(\lambda-\alpha_{j+1})\cdots(\lambda-\alpha_m).$$

Then

$$(3.4) \qquad \det S = (-1)^{\frac{1}{2}m(m-1)}\ \alpha_2\alpha_3^2\cdots\alpha_m^{m-1},$$

and

$$(3.5) \qquad AS = S\tilde{A},$$

_i.e.,_ S _intertwines_ A _and_ $\tilde{A}$.

The matrix $\tilde{A}$ is upper triangular with diagonal $(\alpha_1,\ldots,\alpha_m)$ and S is lower triangular. The inverse $S^{-1} = [s'_{kj}]_{k,j=1}^m$ is given by

$$s'_{kj} = (-1)^{m-k} \sum_{\substack{p_k+\ldots+p_m = k-j-1 \\ p_k,\ldots,p_m = 0,\ldots,k-j-1}} \alpha_k^{-1-p_k}\cdots\alpha_m^{-1-p_m}$$

$$+ (-1)^{m-k} \sum_{\substack{q_{k+1}+\ldots+q_m = k-j \\ q_{k+1},\ldots,q_m = 0,\ldots,k-j}} \alpha_{k+1}^{-1-q_{k+1}}\cdots\alpha_m^{-1-q_m}.$$

Of course S^{-1} is lower triangular too.

Proof. Express $a_0,\ldots,a_{m-1}$ and the coefficients s_{kj} of the polynomials (3.3) in $\alpha_1,\ldots,\alpha_m$ and compute. (Or apply [4, Lemma 3.1] with the matrix Z appearing there equal to the lower triangular nilpotent m×m Jordan block). []

THEOREM 3.2. _Let_ A _be a second companion matrix as in_ (3.1). _Assume there exists an invertible lower triangular_ m×m _matrix_ S _such that_ $S^{-1}AS$ _is upper triangular with diagonal_ $(\alpha_1,\ldots,\alpha_m)$. _Then_

(3.6) $\alpha_k \neq 0, \quad k = 2,\ldots,m.$

Conversely, suppose $\alpha_1,\ldots,\alpha_m$ _are the eigenvalues_ of A _and_ (3.6) _is satisfied. Define the matrices_ $\tilde{A}$ _and_ $S = [s_{kj}]_{k,j=1}^m$ _by_ (3.2) _and_ (3.3), _respectively. Then_ S _is invertible_, S _is lower triangular and_

(3.7) $S^{-1}AS = \tilde{A}.$

In particular $S^{-1}AS$ *is upper triangular with diagonal* $(\alpha_1,\ldots,\alpha_m)$.

The eigenvalues of A can be ordered in such a way that (3.6) is satisfied if and only if A has at most one eigenvalue zero (counted according to algebraic multiplicity). In terms of the elements of A, this means that either $a_0 \neq 0$ or $a_1 \neq 0$. Condition (3.6) can be rephrased as follows: At most one eigenvalue of A is zero, and a possible zero comes first. Theorem 3.2 is a refined version of Theorem 0.4 from the Introduction.

Proof. Condition (3.6) is equivalent to requiring that the matrix $S = [s_{kj}]_{k,j=1}^{m}$ given by (3.3) is invertible. This follows from (3.4). If S is invertible, (3.5) and (3.7) amount to the same. Therefore the second part of the theorem is an immediate consequence of Lemma 3.1.

Assume S has the properties mentioned in the first part of the theorem. If $\alpha_k = 0$ for some $k \geq 2$, there exists a non-zero row vector $x = [x_1 \ \ldots \ x_m]$ satisfying $xS^{-1}AS = 0$ and $x_j = 0$, $j = 1,\ldots,k-1$. In particular $x_1 = 0$. Clearly xS^{-1} is a left eigenvector of A corresponding to the eigenvalue zero of A. Since A is of second companion type, it follows that xS^{-1} is a scalar multiple of $[1 \ 0 \ \ldots \ 0]$. Hence x is a scalar multiple of the first row of the lower triangular matrix S. Combining this with $x_1 = 0$, one gets $x = 0$ contradicting the fact that x is a non-zero vector. Hence (3.6) is satisfied. []

First companion matrices are the transposes of second companion matrices. So in the next theorem $Z = A^T$, where A is as in (3.1).

THEOREM 3.3. *Let* Z *be a first companion* m×m *matrix. Given an ordering* $\zeta_1,\ldots,\zeta_m$ *of the eigenvalues of* Z, *introduce the matrix* $T = [t_{kj}]_{k,j=2}^{m}$

by stipulating that

$$(3.8)\qquad \sum_{j=1}^{m} t_{kj}\lambda^{j-1} = (\lambda-\zeta_1)\ldots(\lambda-\zeta_{k-1}).$$

Then T is invertible, T is lower triangular and

$$(3.9)\qquad TZT^{-1} = \begin{bmatrix} \zeta_1 & 1 & 0 & \ldots & 0 \\ 0 & \zeta_2 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & 1 \\ 0 & 0 & 0 & \ldots & \zeta_m \end{bmatrix}.$$

In particular TZT^{-1} is upper triangular with diagonal $(\zeta_1,\ldots,\zeta_m)$.

The inverse $T^{-1} = [t'_{kj}]_{k,j=1}^{m}$ of T is given by

$$t'_{kj} = \sum_{\substack{p_1+\ldots+p_j = k-j \\ p_1,\ldots,p_j = 0,\ldots,k-j}} \zeta_1^{p_1}\ldots\zeta_j^{p_j}.$$

Of course T^{-1} is lower triangular with diagonal (1,...,1) too. Theorem 3.3 is a refined version of Theorem 0.3 from the Introduction.

Proof. Express the entries in the last row of Z and the coefficients t_{kj} of the polynomials (3.8) in $\zeta_1,\ldots,\zeta_m$ and compute. []

Combining Theorems 3.2 and 3.3 with the observation (concerning the first part of Corollary 2.2) formulated in the last paragraph of Section 2, one immediately obtains the following results.

COROLLARY 3.4. *Let* A *be a second companion* m×m *matrix, and let* Z *be a lower triangular* m×m *matrix. Suppose* A *has at most one eigenvalue zero (counted according to algebraic multiplicity). Then, given an ordering* $\alpha_1,\ldots,\alpha_m$ *of the eigenvalues of* A *satisfying* (3.6), *there exists an invertible lower triangular* m×m *matrix* S *such that* $S^{-1}AS$ *is upper triangular with diagonal* $(\alpha_1,\ldots,\alpha_m)$ *and* $S^{-1}ZS$ *is lower triangular (with the same diagonal as* Z).

For S one can take the matrix defined by (3.3) in Lemma 3.1. In that case $S^{-1}AS = \tilde{A}$, where $\tilde{A}$ is given by (3.2).

COROLLARY 3.5. *Let* A *be an upper triangular* m×m *matrix, and let* Z *be a second companion* m×m *matrix. Then, given an ordering* $\zeta_1,\ldots,\zeta_m$ *of the eigenvalues of* Z, *there exists an upper triangular* m×m *matrix* S *such that* $S^{-1}AS$ *is upper triangular (with the same diagonal as* A) *and* $S^{-1}ZS$ *is lower triangular with diagonal* $(\zeta_1,\ldots,\zeta_m)$.

For S one can take the transpose of the matrix T introduced in Theorem 3.3 with α_j replaced by ζ_j. In that case $S^{-1}ZS$ is the transpose of the matrix appearing in the right hand side of (3.9).

We conclude this section with a remark about Corollary 3.4. The particular case when $Z = Z_o$, where Z_o is the lower triangular nilpotent m×m Jordan block was discussed already in [4, Section 3, Case 2]. It is interesting to see what happens when, more generally, one takes for Z a lower triangular m×m Toeplitz matrix,

$$(3.10)\qquad Z = \begin{bmatrix} z_0 & 0 & 0 & \cdots & 0 \\ z_1 & z_0 & 0 & \cdots & 0 \\ z_2 & z_1 & z_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ z_{m-1} & z_{m-2} & z_{m-3} & \cdots & z_0 \end{bmatrix},$$

i.e., Z is a polynomial in Z_o. Clearly, if Z is of the form (3.10), Corollary 3.4 applies. The following converse result holds true. Let A be a second companion m×m matrix and let Z be the lower triangular Toeplitz matrix given by (3.10). Assume there exists an invertible m×m matrix S such that $S^{-1}AS$ is upper triangular with diagonal $(\alpha_1,\ldots,\alpha_m)$ and $S^{-1}ZS$ is lower triangular. If $z_1 \neq 0$, then (3.6) is satisfied. Indeed, Z_o is of the form p(Z) for some polynomial p. So our assumptions imply that A and Z_o admit simultaneous reduction to complementary triangular forms. Apply now the results of [4, Section 3].

4. THE CONNECTION WITH COMPLETE FACTORIZATION

We begin bij reviewing some material from systems theory. The material is concerned with rational n×n matrix functions. We shall always assume that these functions are proper (i.e. analytic at ∞). The relevant references are [10], [11], [9], [5], [7], [13], [2] and [3].

Let $W(\lambda)$ be a rational n×n matrix function. Then $W(\lambda)$ can be written in the form

$$(4.1)\qquad W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B,$$

where A is an m×m matrix, B is an m×n matrix and C is an n×m matrix. The identity (4.1) implies that

$$(4.2) \qquad W(\lambda)^{-1} = I_n - C(\lambda I_m - Z)^{-1}B,$$

where $Z = A-BC$ (= $A^{\times}$ in the notation of [2], [3], [8] and [1]). Expressions of the type (4.1) and (4.2) are called realizations.

The smallest possible (non-negative) integer m for which a given rational n×n matrix function $W(\lambda)$ admits a realization (4.1) is called the McMillan degree of $W(\lambda)$ and is denoted by $\delta(W)$. It is equal to the total number of poles of $W(\lambda)$ counted according to pole multiplicity. For a discussion of this notion, see [2]. Note that $\delta(W) = 0$ if and only if $W(\lambda)$ is identically equal to I_n.

The realization (4.1) is called minimal if $m = \delta(W)$. An equivalent requirement is that

$$(4.3) \qquad \operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{m-1} \end{bmatrix} = m$$

and

$$(4.4) \qquad \operatorname{rank} [B \; AB \; \ldots \; A^{m-1}B] = m.$$

The matrices appearing in (4.3) and (4.4) have sizes mn×m and m×mn, respectively. The minimality of (4.1) implies that of (4.2). In particular, the McMillan degrees of $W(\lambda)$ and $W(\lambda)^{-1}$ are the same.

Minimal realizations are essentially unique: if (4.1) is a minimal realization of $W(\lambda)$, then all possible minimal realizations of $W(\lambda)$ can be obtained by replacing A, B and C by (respectively) $S^{-1}AS$, $S^{-1}B$ and CS, where S is any invertible m×m matrix. This result is known as the state space isomorphism theorem.

Suppose (4.1) is a minimal realization of $W(\lambda)$. Then the poles of $W(\lambda)$ coincide with the eigenvalues of A. More precisely, the following results hold true:

(i) λ_o is a pole of $W(\lambda)$ of pole multiplicity k if and only if λ_o is an eigenvalue of A of algebraic multiplicity k,

(ii) λ_o is a pole of $W(\lambda)$ of order p if and only if λ_o is a pole of $(\lambda I_m - A)^{-1}$ of order p.

A pole λ_o of $W(\lambda)$ is called geometrically simple if its pole multiplicity happens to be equal to its pole order. Note that λ_o is a geometrically simple pole if and only if λ_o is an eigenvalue of A of geometric multiplicity 1. So all poles of $W(\lambda)$ are geometrically simple poles if and only if A is non-derogatory. Scalar rational functions have geometrically simple poles only.

The McMillan degree is sublogarithmic in the following sense. If $W(\lambda) = W_1(\lambda)\ldots W_k(\lambda)$ is a factorization of $W(\lambda)$, then

$$\delta(W) \leq \sum_{j=1}^{k} \delta(W_j).$$

Of special interest are factorizations for which equality holds (no "pole-zero cancellations"). These are called minimal factorizations. There are rational matrix functions that do not allow for any non-trivial minimal factorization (cf. [5] or [1, Subsection 7.1]).

A rational nxn matrix function is called elementary if it has McMillan degree 1. A complete factorization is a minimal factorization involving elementary factors only. Thus a complete factorization is a factorization of the form

$$(4.5) \qquad W(\lambda) = \left(I_n + \frac{1}{\lambda-\alpha_1}R_1\right) \cdots \left(I_n + \frac{1}{\lambda-\alpha_m}R_m\right),$$

where m is the McMillan degree of $W(\lambda)$, $\alpha_1,\ldots,\alpha_m$ are the poles of $W(\lambda)$ counted according to pole multiplicity and $R_1,\ldots,R_m$ are $n\times n$ matrices of rank 1.

Two questions arise in this context. When does a given rational $n\times n$ matrix function $W(\lambda)$ admit a complete factorization (4.5) ? And if it does, what can be said about the order in which the poles $\alpha_1,\ldots,\alpha_m$ may appear in (4.5) ?

For the general case, an answer can be given in terms of realizations and simultaneous reduction to complementary triangular forms: Let (4.1) be a minimal realization of $W(\lambda)$. Then, given an ordering $\alpha_1,\ldots,\alpha_m$ of the poles of $W(\lambda)$, there exists a complete factorization of $W(\lambda)$ of the form (4.5) if and only if there exists an invertible $m\times m$ matrix S such that $S^{-1}AS$ is upper triangular with diagonal $(\alpha_1,\ldots,\alpha_m)$ and $S^{-1}(A-BC)S$ is lower triangular. For details, see [4, section 6].

Here we shall discuss the special situation (including the scalar case) where $W(\lambda)$ has geometrically simple poles only. In terms of a minimal realization (4.1) of $W(\lambda)$ this means that A is non-derogatory. This condition is certainly satisfied when A has no multiple eigenvalues, i.e., each pole of $W(\lambda)$ has pole multiplicity 1.

THEOREM 4.1. Let $W(\lambda)$ be a rational $n\times n$ matrix function with $W(\infty) = I_n$, and assume each pole of $W(\lambda)$ is geometrically simple. Let $\alpha_1,\ldots,\alpha_m$ be the poles of $W(\lambda)$ counted according to pole multiplicity (so $m = \delta(W)$), and let

$$W(\lambda) = I_n + C(\lambda I_m - A)^{-1}B$$

be a minimal realization of $W(\lambda)$ *such that* A *is lower triangular with diagonal* $(\alpha_m,\ldots,\alpha_1)$. *Then* $W(\lambda)$ *admits a complete factorization of the form*

$$W(\lambda) = \left(I_n + \frac{1}{\lambda-\alpha_1}R_1\right) \cdots \left(I_n + \frac{1}{\lambda-\alpha_m}R_m\right)$$

(with the poles in the given order) if and only if A-BC *admits upper triangularization by a lower triangular similarity*.

A realization of the type hypothesized in the theorem always exists.

Proof. The conclusion of the theorem holds if and only if there exists an invertible m×m matrix S such that $S^{-1}(A-BC)S$ is lower triangular and $S^{-1}AS$ is upper triangular with diagonal $(\alpha_1,\ldots,\alpha_m)$. An equivalent requirement is that there exists an invertible m×m matrix L such that $L^{-1}(A-BC)L$ is upper triangular and $L^{-1}AL$ is lower triangular with diagonal $(\alpha_m,\ldots,\alpha_1)$. By hypothesis $W(\lambda)$ has geometrically simple poles only, i.e., A is non-derogatory. Also A is assumed to be lower triangular with diagonal $(\alpha_m,\ldots,\alpha_1)$. Apply now Corollary 2.2 (with A replaced by A-BC and Z replaced by A). []

As was already observed in the last but one paragraph of Section 2, the only if part of Corollary 2.2 holds even without the condition that the matrix Z appearing in the corollary is non-derogatory. Correspondingly, *the if part of Theorem* 4.1 *is true without the assumption that* $W(\lambda)$ *has geometrically simple poles only*.

A sufficient condition for $W(\lambda)$ to have geometrically simple poles only is that each pole of $W(\lambda)$ has pole multiplicity 1. In terms of a minimal realization (4.1) of $W(\lambda)$ this means that A has no multiple eigenvalues. In particular A is diagonable, which implies that $W(\lambda)$ does admit complete factorization (see [2, Theorems 1.6 and 3.4]; cf. also [5], [13] and [1]). The following example shows that in such a factorization the order of the poles cannot always be chosen at will.

EXAMPLE 4.2. Let

$$W(\lambda) = \begin{bmatrix} 1 & 0 \\ \frac{1}{\lambda} & \frac{\lambda}{\lambda-1} \end{bmatrix}.$$

Then $W(\lambda)$ has 0 and 1 as poles. Both poles have pole multiplicity 1. Clearly

$$W(\lambda) = \begin{bmatrix} 1 & 0 \\ \frac{1}{\lambda} & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & \frac{\lambda}{\lambda-1} \end{bmatrix}$$

is a complete factorization of $W(\lambda)$. There does not exist a complete factorization of $W(\lambda)$ of the form

$$W(\lambda) = \left(I + \frac{1}{\lambda-1}R_1\right)\left(I + \frac{1}{\lambda}R_2\right).$$

To see this, write $W(\lambda) = I + C(\lambda I-A)^{-1}B$ with

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix},$$

and note that A-BC (being lower triangular and non-diagonable) does not admit upper triangularization by a lower triangular similarity. []

REFERENCES

[1] H. Bart, Transfer functions and operator theory, Linear Algebra Appl. 84: 33-61 (1986).

[2] H. Bart, I. Gohberg, and M.A. Kaashoek, Minimal Factorization of Matrix and Operator Functions, Operator Theory: Advances and Applications, Vol. 1, Birkhäuser, Basel, 1979.

[3] H. Bart, I. Gohberg, M.A. Kaashoek, and P. Van Dooren, Factorizations of transfer functions, SIAM J. Control Optim. 18: 675-696 (1980).

[4] H. Bart and H. Hoogland, Complementary triangular forms of pairs of matrices, realizations with prescribed main matrices, and complete factorization of rational matrix functions, to appear.

[5] P. DeWilde and J. Vandewalle, On the factorization of a nonsingular rational matrix, IEEE Trans. Circuits and Systems, Vol. CAS-22(8): 387-401 (1975).

[6] F.R. Gantmacher, The Theory of Matrices, Vol. 1, Chelsea, New York, 1959.

[7] M.L.J. Hautus, Controllability and observability conditions of linear autonomous systems, Nederl. Akad. Wetensch., Proc., Ser. A 72: 443-448 (1969).

[8] M.A. Kaashoek, Minimal factorization, linear systems and integral operators, in Operators and Function Theory, NATO ASI Series, C 153 (S.C. Power, Ed.), Reidel, Dordrecht, 1985, pp. 41-86.

[9] T. Kailath, Linear Systems, Prentice-Hall, Englewood Cliffs, N.J., 1980.

[10] R.E. Kalman, Mathematical description of linear dynamical systems, SIAM J. Control 1(2): 152-192 (1963).

[11] R.E. Kalman, P.L. Falb, and M.A. Arbib, Topics in Mathematical System Theory, McGraw-Hill, New York, 1969.

[12] P. Lancaster and M. Tismenetsky, The Theory of Matrices, Second Edition, Academic Press, Orlando, Fl., 1985.

[13] L.A. Sahnovic, On the factorization of an operator-valued transfer function, Soviet Math. Dokl. 17: 203-207 (1976).

Address:

Econometric Institute

Erasmus University Rotterdam

Postbus 1738

3000 DR Rotterdam

LIST OF REPORTS 1987

8700 Publications of the Econometric Institute Second Half 1986: List of Reprints 458-489.

8701/A **J. van Daal and D.A. Walker**, "The problem of aggregation in Walrasian general equilibrium theory", 23 pages

8702/A **L. de Haan and I. Weissman**, "The index of the outstanding observation among n independent ones", 15 pages.

8703/B **H. Bart and H. Hoogland**, "Complementary triangular forms of pairs of matrices, realizations with prescribed main matrices, and complete factorization of rational matrix functions", 48 pages.

8704/A **H.K. van Dijk**, "Some advances in Bayesian estimation methods using Monte Carlo integration", 48 pages.

8705/A **B.M.S. van Praag and J.P. Hop**, "Estimation of continuous models on the basis of set-valued observations", 34 pages.

8706/A **J. Csirik, J.B.G. Frenk, G. Galambos and A.H.G. Rinnooy Kan**, "Probabilistic analysis of algorithms for dual bin packing problems", 16 pages.

8707/C **R. de Zwart en E. de Vries**, "Van kinderbijslag naar basisbeurs; effecten van een nieuwe wet", 49 pagina's.

8708/B **R.J. Stroeker and N. Tzanakis**, "On the application of Skolem's p-adic method to the solution of Thue equations", 30 pages.

8709/A **M. Meanti, A.H.G. Rinnooy Kan, L. Stougie and C. Vercellis**, "A probabilistic analysis of the multiknapsack value function", 10 pages.

8710/A **L. de Haan and S. Resnick**, "On regular variation of probability densities", 15 pages.

8711/B **H. Bart and P.S.M. Kop Jansen**, "Upper triangularization of matrices by lower triangular similarities", 28 pages.