Upper Triangularization of Matrices by Permutations and Lower Triangular Similarity Transformations

H. Bart
Econometric Institute
Erasmus University Rotterdam
Postbus 1738
3000 DR Rotterdam, The Netherlands

Submitted by Thomas J. Laffey

ABSTRACT

Let A be an arbitrary (square) matrix. As is well known, there exists an invertible matrix S such that $S^{-1}AS$ is upper triangular. The present paper is concerned with the observation that S can always be chosen in the form $S = \Pi L$, where Π is a permutation matrix and L is lower triangular. Assuming that the eigenvalues of A are given, the matrices Π, L, and $U = L^{-1}\Pi^{-1}A\Pi L$ are constructed in an explicit way. The construction gives insight into the freedom one has in choosing the permutation matrix Π. Two cases where Π can be chosen to be the identity matrix are discussed (A diagonable, A a low order Toeplitz matrix). There is a connection with systems theory.

0. INTRODUCTION

In [6] and [7] three mutually related concepts play a role, namely

complete factorization of rational matrix functions,
complementary triangular forms of pairs of matrices,
upper triangularization of matrices by lower triangular similarities.

To put the present paper in the right perspective, we present some details.

A rational matrix function is called *elementary* if it has McMillan degree 1, i.e., it is of the form $I + (\lambda - \alpha)^{-1}R$, where R is a matrix of rank 1. A

655 Avenue of the Americas, New York, NY 10010 0024-3795/88/$3.50

complete factorization of a rational matrix function is a factorization into elementary factors which has the additional property of being *minimal*, i.e., there are no "pole/zero cancellations." Factorizations of this type are important in systems theory (see [8], [11], [4], [5], [2], [9], and [1]).

For a large class of rational matrix functions, the problem of complete factorization is equivalent to that of *simultaneous reduction to complementary triangular forms* of pairs of matrices (see [6] and the references given there): Given two matrices A and Z, when does there exist a similarity transformation S such that $S^{-1}AS$ is upper triangular and $S^{-1}ZS$ is lower triangular? In turn, the latter problem is closely related to that of upper triangularization by lower triangular similarities. This is explained in [7, Section 2].

The present paper focuses on questions concerning upper triangularization by lower triangular similarities. Let us give a formal definition. A (complex) $m\times m$ matrix M is said to admit *upper triangularization by a lower triangular similarity* if there exists an invertible lower triangular $m\times m$ matrix L such that $L^{-1}ML$ is upper triangular. We shall show that modulo a similarity transformation with a permutation matrix, each (square) matrix has this property. In fact we shall be concerned with the following more detailed result.

THEOREM. *Let A be a complex $m\times m$ matrix. Then, given an ordering $\alpha_1,\dots,\alpha_m$ of the eigenvalues of A, there exists an $m\times m$ permutation matrix Π and an invertible lower triangular $m\times m$ matrix L such that $L^{-1}\Pi^{-1}A\Pi L$ is upper triangular with diagonal $(\alpha_1,\dots,\alpha_m)$.*

The *diagonal* of a (square) matrix $K=[k_{ij}]_{i,j=1}^{m}$ is the ordered m-tuple $(k_{11},\dots,k_{mm})$.

If one is merely interested in a proof of the existence of Π and L, a simple induction argument (on the order m of A) will suffice. For details, see the technical report [3]. An even quicker existence proof can be given as follows. Choose an invertible $m\times m$ matrix S such that $S^{-1}AS$ is upper triangular with diagonal $(\alpha_1,\dots,\alpha_m)$. The matrix S can be factorized as ΠLU, where Π is a permutation matrix, L is lower triangular, and U is upper triangular [10, Section 2.10]. Clearly Π and L have the desired properties.

As indicated before, in both this proof and the induction argument given in [3], the emphasis is on the existence of L and Π. The point of the present paper, however, lies in an explicit construction of Π, L, and $U=L^{-1}\Pi^{-1}A\Pi L$. This construction (in which the eigenvalues of A are supposed to be given) is described in Section 1. As a by-product, it gives insight into the freedom that one has in choosing the permutation matrix Π. With this information we return (in Section 2) to the study of upper triangularization by lower triangular similarities (i.e., the case when Π can be taken to be the identity matrix). Attention is paid to the situation where A is diagonable (resulting in a modest refinement of [7, Theorem 0.1]) and to the case when A is a low order Toeplitz matrix.

A few remarks on notation and terminology: All matrices to be considered have complex entries. The $m\times m$ identity matrix is denoted by I_m, or simply I. The superscript T stands for the operation of taking the transpose of a matrix (or vector).

1. THE CONSTRUCTIVE APPROACH

In this section we shall prove the Theorem stated in the Introduction by directly constructing L, Π, and $U=L^{-1}\Pi^{-1}A\Pi L$.

Write $A_0(\lambda)=\lambda I_m - A$, and let $e_1,\dots,e_m$ be the standard basis of $\mathbb{C}^m$. Since α_1 is an eigenvalue of A, there exists a nonzero vector $x_1=(x_{11},\dots,x_{m1})^T$ such that $A_0(\alpha_1)x_1=0$. Choose $\sigma(1)$ with $x_{\sigma(1),1}\neq 0$. Now replace the $\sigma(1)$th column of $A_0(\lambda)$ by x_1 and denote the resulting matrix by $A_1(\lambda)$. A straightforward argument shows that

$$(\lambda-\alpha_1)\det A_1(\lambda)=x_{\sigma(1),1}\det A_0(\lambda);$$

hence

$$\det A_1(\lambda)=x_{\sigma(1),1}(\lambda-\alpha_2)\cdots(\lambda-\alpha_m).$$

In particular $\det A_1(\alpha_2)=0$.

But then there exists a nonzero vector $x_2=(x_{12},\dots,x_{m2})^T$ such that $A_1(\alpha_2)x_2=0$. Since $x_1\neq 0$, at least one of the components x_{s2}, $s\neq\sigma(1)$, must be nonzero. Choose $\sigma(2)$ with $\sigma(2)\neq\sigma(1)$ and $x_{\sigma(2),2}\neq 0$, and write $\tilde{x}_2$ for the vector arising from x_2 on replacing $x_{\sigma(1),2}$ by 0. Replace the $\sigma(2)$th column of $A_1(\lambda)$ by $\tilde{x}_2$, and denote the resulting matrix by $A_2(\lambda)$. Again without difficulty, one checks that

$$(\lambda-\alpha_2)\det A_2(\lambda)=x_{\sigma(2),2}\det A_1(\lambda).$$

Hence

$$\det A_2(\lambda) = x_{\sigma(1),1}x_{\sigma(2),2}(\lambda-\alpha_3)\cdots(\lambda-\alpha_m),$$

and consequently $\det A_2(\alpha_3)=0$.

We proceed by choosing a nonzero vector $x_3=(x_{13},\dots,x_{m3})^T$ such that $A_2(\alpha_3)x_3=0$. Clearly x_1 and $\tilde{x}_2$ are linearly independent, so at least one of the components x_{s3}, $s\neq\sigma(1),\sigma(2)$, is nonzero. Take $\sigma(3)$ with $\sigma(3)\neq\sigma(1),\sigma(2)$ and $x_{\sigma(3),3}\neq 0$, and write $\tilde{x}_3$ for the vector arising from x_3 on replacing both $x_{\sigma(1),3}$ and $x_{\sigma(2),3}$ by zero. Replace the $\sigma(3)$th column of $A_2(\lambda)$ by $\tilde{x}_3$, and denote the resulting matrix by $A_3(\lambda)$. Then

$$\det A_3(\lambda) = x_{\sigma(1),1}x_{\sigma(2),2}x_{\sigma(3),3}(\lambda-\alpha_4)\cdots(\lambda-\alpha_m),$$

and we can go on with the construction.

In m steps, we obtain vectors $x_1=\tilde{x}_1, x_2, \tilde{x}_2,\dots,x_m,\tilde{x}_m$, a permutation $\sigma=(\sigma(1),\dots,\sigma(m))$ of the numbers $1,\dots,m$, and λ-matrices $A_0(\lambda),\dots,A_m(\lambda)$ such that

(i) $A_0(\lambda)=\lambda I_m - A$,

(ii) $A_{k-1}(\alpha_k)x_k=0$, $x_{\sigma(k),k}\neq 0$,

(iii) $\tilde{x}_k$ is obtained from $x_k=(x_{1k},\dots,x_{mk})^T$ by replacing the components $x_{\sigma(1),k},\dots,x_{\sigma(k-1),k}$ by zero,

(iv) $A_k(\lambda)$ is obtained from $A_{k-1}(\lambda)$ by replacing the $\sigma(k)$th column of $A_{k-1}(\lambda)$ by the vector $\tilde{x}_k$.

Note that $A_m(\lambda)$ does not depend on λ any more. Write $S=A_m(\lambda)$. Then S has $\tilde{x}_k$ as its $\sigma(k)$th column. Also, let V be the $m\times m$ matrix having $y_k=\alpha_k e_{\sigma(k)}+x_k-\tilde{x}_k$ as its $\sigma(k)$th column. Observe that the sth component y_{sk} of y_k is equal to α_k when $s=\sigma(k)$, to zero when $s\neq\sigma(1),\dots,\sigma(k)$, and to x_{sk} for s otherwise.

Next we establish the identity

$$S^{-1}AS=V.$$

Since $\tilde{x}_1,\dots,\tilde{x}_m$ are linearly independent, the matrix S is invertible. By (ii), we have $A_{k-1}(\alpha_k)\tilde{x}_k=A_{k-1}(\alpha_k)(\tilde{x}_k-x_k)$. In the expression $A_{k-1}(\alpha_k)\tilde{x}_k$ the only columns of $A_{k-1}(\alpha_k)$ that matter are those for which the corresponding components of $\tilde{x}_k$ do not vanish. So $A_{k-1}(\alpha_k)\tilde{x}_k=(\alpha_k I_m - A)\tilde{x}_k$. Analogously $A_{k-1}(\alpha_k)(\tilde{x}_k-x_k)=S(\tilde{x}_k-x_k)$. It follows that $A\tilde{x}_k=\alpha_k\tilde{x}_k+S(x_k-\tilde{x}_k)$. Now $\tilde{x}_k=Se_{\sigma(k)}$ and $y_k=Ve_{\sigma(k)}$. Hence

$$ASe_{\sigma(k)}=S\big(\alpha_k e_{\sigma(k)}+x_k-\tilde{x}_k\big)=Sy_k=Sve_{\alpha(k)},$$

and we may conclude that $AS = SV$. Since S is invertible, this intertwining relation can be rewritten as $S^{-1}AS = V$.

Finally, introduce the permutation matrix Π by stipulating that Π has $e_{\sigma(k)}$ as its kth column. Put $L = \Pi^{-1}S\Pi$ and $U = \Pi^{-1}V\Pi$. Then L is an invertible lower triangular matrix, U is an upper triangular matrix with diagonal $(\alpha_1, \dots, \alpha_m)$, and $L^{-1}\Pi^{-1}A\Pi L = U$. Thus L and Π have the desired properties.

2. UPPER TRIANGULARIZATION BY LOWER TRIANGULAR SIMILARITIES

There is quite a bit of freedom in the construction described in Section 1. Indeed, there is the freedom in the ordering of the eigenvalues of A, there is the choice of the vectors x_k, etc. Thus, the construction gives rise to a collection $\mathscr{S}$ of "admissible" permutations: for each permutation $\sigma = (\sigma(1), \dots, \sigma(m)) \in \mathscr{S}$, there exists a permutation matrix Π_σ such that $\Pi_\sigma^{-1}A\Pi_\sigma$ admits upper triangularization by a lower triangular similarity. Here $\Pi_\sigma(e_k) = e_{\sigma(k)}$, $k = 1, \dots, m$. Conversely, if σ is a permutation of the numbers $1, \dots, m$ and $\Pi_\sigma^{-1}A\Pi_\sigma$ admits upper triangularization by a lower triangular similarity, then $\sigma \in \mathscr{S}$. In fact, the construction described in Section 1 covers all possible ways of upper triangularizing A by means of a permutation and a lower triangular similarity.

Now let us return to the problem studied in [7] which motivated the present paper. The remarks made in the preceding paragraph show that A (itself) admits upper triangularization by a lower triangular similarity if and only if the identical permutation belongs to the set $\mathscr{S}$. For this to be the case, one must be able to carry out the above construction in such a way that $\sigma(j) = j$, $j = 1, \dots, m$. This includes an appropriate choice for the ordering of the eigenvalues of A. We give two examples.

EXAMPLE 1. Suppose A is diagonable. Then we know from [7, Theorem 0.1] that A admits upper triangularization by a lower triangular similarity. This can also be seen along the lines suggested above. In fact the analysis leads to the following modest refinement of Theorem 0.1 in [7]:

Let A be a diagonable matrix, and let $\alpha_1, \dots, \alpha_m$ be an ordering of the eigenvalues of A. Then the following two statements are equivalent:

(i) *There exists an invertible lower triangular matrix L such that $L^{-1}AL$ is upper triangular with diagonal $(\alpha_1, \dots, \alpha_m)$;*

(ii) *There exist eigenvectors $w_1, \ldots, w_m$ of A corresponding to the eigenvalues $\alpha_1, \ldots, \alpha_m$, respectively, such that the matrix with w_j as its jth column has nonvanishing leading principal minors.*

For a detailed proof, see [3]. Here we only give an outline of the argument.

To prove that (ii) implies (i), write

$$x_k = w_k + \sum_{j=1}^{k-1} \beta_{kj} w_j + \sum_{j=1}^{k-1} \gamma_{kj} e_j,$$

$$\tilde{x}_k = w_k + \sum_{j=1}^{k-1} \beta_{kj} w_j,$$

and choose the scalars β_{kj} and γ_{kj} such that the vectors $x_1, \tilde{x}_1, \ldots, x_m, \tilde{x}_m$ have the desired properties. This is possible because of the condition on the eigenvectors $w_1, \ldots, w_m$ imposed in (ii). For the implication in the other direction, note that one may assume that A itself is upper triangular. But then one can apply the following observation: if D is an upper tringular diagonable matrix, then there exists an invertible upper triangular matrix U such that $U^{-1}DU$ is a diagonal matrix. The proof can be given by induction (on the order of the matrix D). What makes the induction step possible is that D has an eigenvector x, corresponding to the last eigenvalue in the diagonal of D, such that the last component of x is nonzero. This fact can be proved by induction too.

EXAMPLE 2. Using the technique developed in Section 1, the following result on low order Toeplitz matrices can be proved:

Let A be a Toeplitz matrix of order $m \leqslant 3$. Then A admits upper triangularization by a lower triangular similarity if and only if A is a scalar multiple of the identity or A is not lower triangular.

For details, see [3].

The result in Example 2 does not extend to the case $m \geqslant 4$. For a counterexample, take

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

Then A is nilpotent, and the nullspace of A is spanned by the single vector $(0,1,0,0)^T$. Therefore each permutation $\sigma=(\sigma(1),\sigma(2),\sigma(3),\sigma(4))$ belonging to $\mathscr{S}$ has $\sigma(1)=2$, so $\mathscr{S}$ does not contain the identical permutation.

REFERENCES

1 D. Alpay and I. Gohberg, Unitary rational matrix functions and orthogonal matrix polynomials, to appear.

2 H. Bart, Transfer functions and operator theory, *Linear Algebra Appl.* 84:33–61 (1986).

3 H. Bart, Upper Triangularization of Matrices by Permutations and Lower Triangular Similarities, Report 8741B, Econometric Inst., Erasmus Univ., Rotterdam, 1987.

4 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.

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

6 H. Bart and H. Hoogland, Complementary triangular forms of pairs of matrices, realizations with prescribed main matrices, and complete factorization of rational matrix functions, *Linear Algebra Appl.*, 103:193–228 (1988).

7 H. Bart and P. S. M. Kop Jansen, Upper triangularization by lower triangular similarities, *Linear Algebra Appl.*, 103:229–248 (1988).

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

9 I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, Canadian Math. Soc. Series of Monographs and Advanced Texts, Wiley, New York, 1986.

10 P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd ed., Academic, Orlando, Fla., 1985.

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

Received 4 November 1987; final manuscript received 25 March 1988