NORTH-HOLLAND

An Algorithm for Coprime Matrix Fraction Description using Sylvester Matrices

J. C. Basilio and B. Kouvaritakis
Oxford University
Department of Engineering Science
Parks Road
Oxford OX1 3PJ, U.K.

Submitted by Richard A. Brualdi

ABSTRACT

Several algorithms for the computation of coprime matrix fraction descriptions have been proposed in the past. Here we explore some of the properties of Sylvester matrices and develop a different approach to the problem which is based on singular value decompositions and therefore avoids problems of numerical ill-conditioning. © 1997 Elsevier Science Inc.

1. INTRODUCTION

Coprime matrix fraction descriptions (MFDs) of transfer function matrices play an important role in several aspects of linear multivariable feedback system design [1–6]. Two broad approaches to the computation of such descriptions have dominated the recent literature: one using state-space system realizations and the other using Sylvester resultant matrices. The former [7, 8] starts from a controllable realization and uses an appropriate transformation that brings the state matrix into Hessenberg form. This provides the means for the derivation of a recursive algorithm that leads to an irreducible transfer function matrix fraction description. The Sylvester matrix approach, on the other hand, solves a problem which derives from the minimum design problem defined by [9] and which can be stated as follows: given a $p \times q$ transfer function matrix $H(z)$ with a left MFD $A^{-1}(z)B(z)$,

655 Avenue of the Americas, New York, NY 10010
0024-3795/97/$17.00
PII S0024-3795(96)00636-2

find a right coprime MFD $N(z)M^{-1}(z)$; it is assumed that $A(z)$, $B(z)$, $N(z)$, and $M(z)$ are $p \times p$, $p \times q$, $p \times q$, and $q \times q$ polynomial matrices, respectively. This reduces to finding the right coprime polynomial matrices $N(z)$ and $M(z)$ which satisfy the equation

$$\begin{bmatrix} B(z) & -A(z) \end{bmatrix} \begin{bmatrix} M(z) \\ N(z) \end{bmatrix} = 0. \tag{1}$$

An elegant solution to this problem is provided by the computation of a minimal polynomial basis for the right null space of $[B(z) \quad -A(z)]$ [10]: (i) form an appropriate Sylvester resultant matrix S, by equating the coefficient matrices of $B(z)M(z)$ and $A(z)N(z)$ of corresponding power; (ii) use a search algorithm to find the first q primary dependent columns of S; (iii) write these columns as linear combinations of the preceding linearly independent columns; (iv) use the coefficients of linear dependence to form $N(z)$ and $M(z)$. Step (ii) is numerically challenging, and Kung and Kailath [11] use an orthogonalization process on S to determine whether the innovation introduced by a particular column is zero or not and hence to determine whether this particular column is linearly dependent or not on the preceding columns. To improve numerical robustness, Datta and Gangopadhyay [12] deploy Householder transformations to search for the primary dependent columns of S. Further problems of ill-conditioning to do with the determination of rank and linear dependence can be ameliorated through the use of singular value decomposition, but this remedy can prove to be computationally demanding.

Here, rather than compute a polynomial basis for the kernel of S, we explore properties of S and come up with an alternative algorithm which avoids the above difficulties.

2. THE DEFINITION OF THE SYLVESTER MATRIX

Let $A(z)$, $B(z)$, $N(z)$, and $M(z)$ be polynomial matrices given as

$$A(z) = \sum_{k=0}^{a} A_k z^{a-k}, \qquad B(z) = \sum_{k=0}^{b} B_k z^{b-k},$$

$$N(z) = \sum_{k=0}^{n} N_k z^{n-k}, \qquad M(z) = \sum_{k=0}^{m} M_k z^{m-k}. \tag{2}$$

Then Equation (1) implies that the sum of the coefficients of $B(z)M(z)$ and $-A(z)N(z)$ of corresponding power is zero and thus we may write

$$\begin{bmatrix} B_0 & 0 & \cdots & 0 & -A_0 & 0 & \cdots & 0 \\ B_1 & B_0 & \cdots & 0 & -A_1 & -A_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ B_b & B_{b-1} & & B_0 & -A_a & -A_{a-1} & & -A_0 \\ 0 & B_b & \ddots & \vdots & 0 & -A_a & \ddots & \vdots \\ \vdots & \vdots & \ddots & B_{b-1} & \vdots & \vdots & \ddots & -A_{a-1} \\ 0 & 0 & \cdots & B_b & 0 & 0 & \cdots & -A_a \end{bmatrix} \begin{bmatrix} M_0 \\ M_1 \\ \vdots \\ M_m \\ N_0 \\ N_1 \\ \vdots \\ N_n \end{bmatrix} = 0, \tag{3}$$

$\leftarrow m + 1$ block cols. $\rightarrow$ $\leftarrow n + 1$ block cols. $\rightarrow$

$$S_{m,n}(B, A)X = 0.$$

Implicit in the above is the assumption that $b + m = a + n$; in the general case this condition can be brought about by introducing the prerequisite number of zero leading coefficients into $B(z)$ and $A(z)$ depending on whether $b + m < a + n$ or $b + m > a + n$. The matrix $S_{m,n}(B, A)$ above is referred to as the Sylvester resultant matrix of $A(z)$ and $B(z)$.

The problem of finding a right coprime MFD, $N(z)M^{-1}(z)$, therefore requires the computation of suitable integers m and n and the real matrix solution X of Equation (3). Clearly the columns of X must lie in the right null space, $\kappa(S)$, of $S_{m,n}(B, A)$. However, in general the dimension of $\kappa(S)$, $\nu(S) = \dim \kappa(S)$, will exceed the number of columns q of X, and thus the solution X is not unique. Matrix representations of any q-dimensional subspace of $\kappa(S)$ will satisfy Equation (3) but will not necessarily lead to a right MFD which is coprime. In this paper we explore the properties of $S_{m,n}(B, A)$ in order to determine the smallest permissible values for m and n, and propose an algorithm for the computation of matrices X which lead to coprime polynomial factorizations.

3. PROPERTIES OF NONCOPRIME POLYNOMIAL MATRICES

Most of the properties discussed in this section relate to left MFDs but have obvious counterparts for right factorizations.

DEFINITION 3.1. Two polynomial matrices with the same number of rows are said to be noncoprime if there exist polynomial matrices $A'(z)$ and $B'(z)$ such that $A(z) = L(z)A'(z)$ and $B(z) = L(z)B'(z)$ for some nonunimodular polynomial matrix $L(z)$, i.e. $\delta(|L|) \neq 0$, where $\delta(\cdot)$ denotes degree and $|\cdot|$ stands for determinant. The polynomial matrix $L(z)$ is referred to as a common left divisor of $A(z)$ and $B(z)$ and is said to be a greatest common left divisor (gcld) if any other common divisor $L_1(z)$ of $A(z)$ and $B(z)$ is also a left divisor of $L(z)$.

For convenience from now on we shall adopt the notation of this definition: given two noncoprime matrices $A(z)$, $B(z)$, then $L(z)$ will denote a gcld and the corresponding right factors will be denoted by $A'(z)$ and $B'(z)$. The definition above has the following two implications: (i) if $L(z)$ is a gcld, then so is $L(z)U(z)$, where $U(z)$ is unimodular; hence gcld's are not unique; (ii) the degree of $L(z)U(z)$ can be made arbitrarily large [even though $\delta(|L(z)U(z)|) = \delta(|L(z)|) =$ constant], and henceforth gcld will be used to refer to the greatest common left divisor of least degree.

DEFINITION 3.2. Two noncoprime matrices $A(z)$ and $B(z)$ are said to be *strictly not coprime* (SNCP) if $\delta(A) = \delta(L) + \delta(A')$ and $\delta(B) = \delta(L) + \delta(B')$.

The property of interest that follows from the above definition is that dividing out gcld's from a pair of SNCP matrices $A(z)$ and $B(z)$ reduces the degree of $A(z)$ and $B(z)$, but this is not the case if $A(z)$ and $B(z)$ are not SNCP.

PROPOSITION 3.1. *Let $B(z)$ and $A(z)$ ($A(z)$ square) be SNCP.* Then $|A_0| \neq 0$ implies that $|L_0| \neq 0$ and $|A'_0| \neq 0$, where A_0, A'_0, and L_0 denote the leading coefficients of $A(z)$, $A'(z)$, and $L(z)$.

Proof. By the definition of $L(z)$ and $A'(z)$ we have that $A_0 = L_0 A'_0$, so that taking determinants we get the result of the proposition. ∎

PROPOSITION 3.2. *Let $B(z)$ and $A(z)$ ($A(z)$ square) be noncoprime but not SNCP.* Then $\delta(B) < \delta(L) + \delta(B')$ implies that L_0 and B'_0 are rank deficient, and $\delta(A) < \delta(L) + \delta(A')$ implies that L_0 and A'_0 are singular.

Proof. The first inequality implies that the coefficient of $z^{\delta(L)+\delta(B')}$, $L_0 B'_0$, in $L(z)B'(z)$ must be zero, which in turn implies that both L_0 and B'_0 are rank deficient; the proof for L_0 and A'_0 is the same. ∎

REMARK 3.1. The rank deficiency of L_0 and A'_0 does not imply that the leading coefficient A_0 of $A(z)$ is itself rank deficient. Indeed, in the context of the paper, where $A(z)$ is the denominator polynomial matrix of a transfer function matrix $H(z)$, it is always possible to arrange for A_0 to be nonsingular. For example, we can choose $A(z) = d(z)I_p$, where $d(z)$ denotes the monic least common denominator of $H(z)$ and I_p is the p-dimensional identity matrix, so that $A_0 = I$. For this reason, henceforth, unless otherwise stated, A_0 will be assumed to be nonsingular. ■

PROPOSITION 3.3. *If $\delta(A') > \delta(A)$ then $L(z)$ is unimodular.*

Proof. Let $\delta(A) = a$, $\delta(L) = \lambda$, $\delta(A') = a'$, $\tau = a' - a > 0$, and define $A^{\#}(z) = A'_0 z^{\tau} + A'_1 z^{\tau-1} + \cdots + A'_{\tau}$; then the coefficients of $z^{\lambda + a' - i}$ for $i = 0, 1, \dots, \tau - 1$ in $L(z)A'(z)$ will be zero. Using this fact, it is easy to show that $L(z)A^{\#}(z) = L_{\lambda}A'_{\tau}$ and thus $\delta(|L(z)|) + \delta(|A^{\#}(z)|) = 0$, which implies that $\delta(|L(z)|) = 0$. ■

4. DETERMINATION OF DIMENSIONS FOR THE SYLVESTER MATRIX

The solution X of Equation (3) exists if $S_{m,n}(B, A)$ has a kernel which is at least q-dimensional, and this will be so for $m = a + \gamma$, $n = b + \gamma$ with γ an integer such that $\gamma \geqslant a(p - q)/q$, because then S will have at least q more columns than rows. In the interest of keeping the computational complexity to a minimum it is important to ensure that the dimensions of S are as small as possible, and thus a sensible first choice for m, n is $m = \alpha = a + \gamma_0$, $n = \beta = b + \gamma_0$, where γ_0 denotes the smallest integer such that $\gamma_0 \geqslant a(p - q)/q$. However, as will be seen below, in the general case smaller values for m and n may be possible.

LEMMA 4.1. *If $A(z)$ and $B(z)$ are noncoprime, then the number of zero singular values, $\mu_{\alpha\beta}$, of $S_{\alpha,\beta}(B, A)$ is greater than zero.*

Proof. For $m = \alpha$, $n = \beta$, S is "short and fat," and so to prove the result we need show that the $S_{\alpha,\beta}(B, A)$ has a left null space.

(1) $A(z)$, $B(z)$ *SNCP*: In this case, by Definition 3.2 we have that $a = \lambda + a'$ and $b = \lambda + b'$, which can be combined with the definitions $\alpha = a + \gamma_0$ and $\beta = b + \gamma_0$ to give that $\alpha + b' = \beta + a'$. Thus $S_{\alpha,\beta}(B, A)$

can be written as

$$S_{\alpha,\beta}(B, A) = \begin{bmatrix} L_0 & & \\ L_1 & \ddots & \\ \vdots & \ddots & L_0 \\ L_\lambda & \cdots & L_1 \\ & \ddots & \vdots \\ & & L_\lambda \end{bmatrix} \begin{bmatrix} B'_0 & & & -A'_0 & & \\ B'_1 & \ddots & & -A'_1 & \ddots & \\ \vdots & \ddots & B'_0 & \vdots & \ddots & -A'_0 \\ B'_{b'} & \cdots & B'_1 & -A'_{a'} & \cdots & -A'_1 \\ & \ddots & \vdots & & \ddots & \vdots \\ & & B_{b'} & & & -A a' \end{bmatrix}$$

$$= C_L S_{\alpha,\beta}(B', A'), \tag{4}$$

where C_L can be shown to have precisely λp more rows than columns. As a consequence C_L, and hence $S_{\alpha,\beta}(B, A)$, will have a left null space (of dimension at least λp).

(2) *$A(z)$, $B(z)$ noncoprime but not SNCP*: Three cases need be considered here: (2i) $\alpha + b' > \beta + a'$; (2ii) $\alpha + b' = \beta + a'$; (2iii) $\alpha + b' < \beta + a'$. Case (2iii) is symmetric to (2i) and need not be considered separately. We begin with case (2i) $\alpha + b' > \beta + a'$. This condition implies that (i) $\alpha + b' - \beta - a'$ zero leading coefficients must be introduced into $A'(z)$; (ii) C_L must have less rows than before. Condition (i) is required to match the degrees of $B'(z)M(z)$ and $A'(z)N(z)$, whereas (ii) is to do with the fact that the degree of $B(z)M(z) - A(z)N(z)$ is $\alpha + b$ (or $\beta + a$), which implies that $S_{\alpha,\beta}(B, A)$ must have $(\alpha + b + 1)p$ rows. However, the C_L and hence $S_{\alpha,\beta}(B, A)$, as given above, have $(\alpha + b' + \lambda + 1)p$ rows: this implies that the first $(\lambda + b - b')p$ rows of the $C_L S_{\alpha,\beta}(B', A')$ (as given above) will be zero and must be dropped. In other words, $S_{\alpha,\beta}(B, A)$ will be given by the expression above but with C_L truncated so as to lose the first $\lambda - \tau$ blocks, where $\tau = b - b'$. As a consequence C_L now will have τp more rows than columns. We distinguish three possibilities.

(a) $\tau < 0$: This implies $b' > b$, which in turn, by Proposition 3.3, implies that $L(z)$ is unimodular; this however cannot be the case, since $A(z)$ and $B(z)$ are assumed to be noncoprime.
(b) $\tau > 0$: In this case C_L and hence $S_{\alpha,\beta}(B, A)$ (as required by the lemma) will possess a left null space.
(c) $\tau = 0$: Here C_L will be square and upper triangular, with L_λ appearing on its diagonal block positions. Thus if L_λ is singular, then C_L will be rank

deficient, and hence $S_{\alpha,\beta}(B, A)$ will have a left null space. If on the other hand L_λ is full rank, C_L will be invertible, and hence $S_{\alpha,\beta}(B, A)$ will have a left null space if and only if $S_{\alpha,\beta}(B', A')$ does. However, since $b < b' + \lambda$ (by the proof of Proposition 3.2), we have that $L_0 B'_0 = 0$; $L_0 A'_0$ is also zero on account of the introduction of zero leading coefficients in $A'(z)$. Hence premultiplying the first row block of $S_{\alpha,\beta}(B', A')$ by L_0 would result in a zero row block, and this establishes the existence of a left null space.

To complete the proof we finally need to consider case (2ii), for which $\alpha + b' = \beta + a'$. This case is in essence the same as above, except that now it is no longer necessary to introduce zero leading coefficients into $A'(z)$, but instead the fact that $L_0 A'_0 = 0$ is implied by the inequality $a < \lambda + a'$ and the proof of Proposition 3.2. It is noted that a cannot be equal to $\lambda + a'$, because this together with $\alpha + b' = \beta + a'$ would contradict the assumption $A(z)$, $B(z)$ not SNCP made earlier. ■

The implication of Lemma 4.1 is that for noncoprime $A(z)$ and $B(z)$ the dimension of the kernel of $S_{\alpha,\beta}(B, A)$ will be greater than q, and this raises the question whether m and n can be chosen to be smaller than α and β without affecting the existence of a solution X to Equation (3). This issue is explored in the following theorem.

THEOREM 4.1. *Let $\rho_{m,n}$, $\nu_{m,n}$, and $\mu_{m,n}$ denote respectively the rank, nullity, and number of zero singular values of $S_{m,n}(B, A)$. Then:*

(i) $\rho_{\alpha-k,\beta-k} \leqslant \rho_{\alpha,\beta} - kp$;

(ii) $\nu_{\alpha-k,\beta-k} \geqslant \mu_{\alpha,\beta} - (k-1)q$;

(iii) *If there exist integers χ, $\eta > 0$, $\eta < q$, such that $\nu_{\alpha,\beta} = (\chi + 1)q + \eta$, then*

$$\nu_{\alpha-\chi,\beta-\chi} = q + \eta; \tag{5}$$

(iv) $\nu_{\alpha-k-1,\beta-k-1} \geqslant \nu_{\alpha-k,\beta-k-1} - \nu_{B_0}$;

(v) $\nu_{\alpha-k,\beta-k-1} \geqslant \nu_{\alpha-k,\beta-k} - \rho_{B_0}$.

Proof. (i): Rearranging the columns of $S_{\alpha,\beta}(B, A)$, we may write the equivalence relationship

$$S_{\alpha,\beta}(B, A) \sim \begin{bmatrix} B_{11} & A_{11} & 0 \\ B_{21} & A_{21} & S_{\alpha-k,\beta-k} \end{bmatrix}, \tag{6}$$

where $B_{11} \in \mathbb{R}^{kp \times kq}$, $B_{21} \in \mathbb{R}^{(\alpha+b-k+1) \times kq}$, $A_{11} \in \mathbb{R}^{kp \times kp}$, $A_{21} \in \mathbb{R}^{(\alpha+b-k+1) \times kp}$, and where A_{11} is a block lower triangular matrix whose diagonal blocks are all A_0. Hence A_{11} is full rank, and this in conjunction with the equivalence relationship above implies (i).

(ii): Given that the number of columns of $S_{\alpha-k,\,\beta-k}(B, A)$ is $(\alpha - k + 1)q + (\beta - k + 1)p$, we have

$$\begin{aligned}
\nu_{\alpha-k,\beta-k} &= (\alpha - k + 1)q + (\beta - k + 1)p - \rho_{\alpha-k,\beta-k} \\
&\geqslant (\alpha - k + 1)q + (\beta - k + 1)p - \rho_{\alpha,\beta} + kp \\
&= \left[(\alpha + 1)q + (\beta + 1)p - \rho_{\alpha,\beta}\right] - kq = \nu_{\alpha,\beta} - kq \\
&\geqslant [\mu_{\alpha,\beta} + q] - kq = \mu_{\alpha,\beta} - (k - 1)q,
\end{aligned} \tag{7}$$

where use has been made of the fact that the nullity of $S_{\alpha,\,\beta}(B, A)$ is equal to the excess of columns over rows plus the number of zero singular values.

(iii): Substituting $\nu_{\alpha\beta} = (\chi + 1)q + \eta$ into (ii), we obtain $\nu_{\alpha-\chi,\beta-\chi} \geqslant q + \eta$, which implies that MFD's of degree greater than or equal to $\alpha - \chi$ exist. Then by Proposition 3.2 of Anderson and Jury [13]

$$\rho_{\alpha-\chi,\beta-\chi} = (\alpha - \chi - 1)p + \delta_M(H(z)), \tag{8}$$

where $\delta_M(H(z))$ denotes the Smith-McMillan degree of $H(z)$. Making recursive use of the above, we deduce

$$\rho_{\alpha,\beta} = \rho_{\alpha-\chi,\beta-\chi} + \chi p. \tag{9}$$

Then using the fact that the nullity is given by the number of columns minus the rank, we can write

$$\begin{aligned}
\nu_{\alpha-\chi,\beta-\chi} &= (\alpha - \chi + 1)q + (\beta - \chi + 1)p - \rho_{\alpha-\chi,\beta-\chi} \\
&= (\alpha + 1)q + (\beta + 1)q - \chi q = \nu_{\alpha,\beta} - \chi q = q + \eta,
\end{aligned} \tag{10}$$

where use has been made of the assumption of the theorem that $\nu_{\alpha\beta} = (\chi + 1)q + \eta$.

(iv): The relationship between $S_{\alpha-k,\beta-k-1}(B, A)$ and $S_{\alpha-k-1,\beta-k-1}(B, A)$ is

$$S_{\alpha-k,\beta-k-1}(B, A) = \begin{bmatrix} B_0 & 0 \quad \cdots \quad 0 \\ B_1 & \\ \vdots & \\ B_b & S_{\alpha-k-1,\beta-k-1}(B, A) \\ 0 & \\ \vdots & \\ 0 & \end{bmatrix}, \tag{11}$$

from which it is apparent that

$$\nu_{\alpha-k,\beta-k-1} \leqslant \nu_{\alpha-k-1,\beta-k-1} + \nu_{B_0}. \tag{12}$$

(v): After some column block rearrangement it is possible to write the following equivalence relationship:

$$S_{\alpha-k,\beta-k}(B, A) \sim \begin{bmatrix} & A_0 \\ & \vdots \\ S_{\alpha-k,\beta-k-1}(B, A) & A_a \\ & 0 \\ & \vdots \\ & 0 \end{bmatrix}, \tag{13}$$

and since A_0 is assumed to be full rank and the first row block of $S_{\alpha-k,\beta-k-1}(B, A)$ is $[B_0, 0, \dots, 0]$, it is easy to establish that condition (v) of the theorem holds true. ■

We can now use the results of Theorem 4.1 to develop an algorithm for the computation of allowable values of m and n. The idea is to start with a $S_{m,n}(B, A)$ for which m and n are such that the dimension of the kernel of $S_{m,n}(B, A)$ is at least q. Now the difficulty with such a choice of m and n is that the kernel of $S_{m,n}(B, A)$ will not necessarily lead to a right coprime MFD of $H(z)$. Indeed, if $A(z)$ and $B(z)$ are noncoprime, then by Lemma 4.1 we have that the number of zero singular values of $S_{m,n}(B, A)$ is going to be greater than zero, and this suggests that it may be possible to reduce the dimensions of $S_{m,n}(B, A)$. The answer to when this is possible and how it can be done in a systematic way is provided by Theorem 4.1 and is described in the algorithm below. The advantage of the algorithm is that it always leads to a matrix $S_{m,n}(B, A)$ whose kernel has dimension $\nu_{m,n}$ such that $q \leqslant \nu_{m,n} < 2q$. Furthermore, this is achieved through at most three singular value decompositions: (i) one to identify the number of zero singular values in the initial choice of $S_{m,n}(B, A)$ (step 2); (ii) another to compute the rank/nullity of B_0, the leading coefficient of $B(z)$, with the view to reducing the degree of $N(z)$ (step 4); and (iii) a last singular decomposition (performed when certain conditions are satisfied) which allows a further reduction in the degree of $M(z)$ (step 5).

ALGORITHM 4.1 (Systematic reduction of m and n).

Step 1. Compute the smallest integer $\gamma = \gamma_0$ such that $\gamma \geqslant a(p-q)/q$, and set $\alpha = a + \gamma_0$, $\beta = b + \gamma_0$.

Step 2. Compute the number $\mu_{\alpha\beta}$ of the zero singular values of $S_{\alpha,\beta}(B, A)$, and hence compute $\nu_{\alpha\beta} = \tau + \mu_{\alpha\beta}$, where τ is the excess of columns over rows in $S_{\alpha,\beta}(B, A)$.

Step 3. Compute the nonnegative integers χ, η with $\eta < q$, such that $\nu_{\alpha\beta} = (\chi + 1)q + \eta$; in addition compute ρ_{B_0}, $\nu_{B_0} = q - \rho_{B_0}$.

Step 4. If $\rho_{B_0} > \eta$, then set $n = \beta - \chi$, $m = \alpha - \chi$, and stop. Otherwise set $n = \beta - \chi - 1$, $m = \alpha - \chi$.

Step 5. Compute $\nu_{\alpha-\chi,\beta-\chi-1}$: if $\nu_{\alpha-\chi,\beta-\chi-1} - \nu_{B0} \geqslant q$, then set $m = \alpha - \chi - 1$; otherwise set $m = \alpha - \chi$. Stop.

REMARK 4.1. The above algorithm is based on conditions which are sufficient only, and has been devised to keep the number of singular value decompositions (of large matrices) to a minimum. In particular, it is possible that the n of the first part of step 4 could be reduced further, but this can only be confirmed after a singular value decomposition has been performed on $S_{\alpha-\chi,\beta-\chi-1}(B, A)$; to avoid the extra computational load, this reduction is omitted.

5. COMPUTATION OF COPRIME FACTORS

5.1. *Necessary Conditions for Right Coprime MFDs*

The purpose of Algorithm 4.1 is to reduce, at a small computational cost, the dimensions of the problem defined in Equation (3). This is done by determining suitably small values for n, m; for these values the kernel of $S_{m,n}(B, A)$ will be of dimension $q + \eta$ and can be used for the computation of solutions $M(z)$, $N(z)$ to Equation (1) as follows.

LEMMA 5.1. *Let X_M be the matrix defined by the first $(m+1)q$ rows of X, and let X_N be the matrix defined by the last $(n+1)p$ rows of X. Furthermore let $D_0, D_1, \dots, D_m$ be the matrices defined by the first q rows of X_M, the next q rows, ..., the last q rows of X_M, respectively, and similarly using the rows of X_N, taken p at a time, define matrices $C_0, C_1, \dots, C_n$. Then the polynomial matrices*

$$N(z) = C(z)R(z), \quad M(z) + D(z)R(z),$$

$$\text{with} \quad C(z) = \sum_{i=0}^{n} C_i z^{n-i}, \; D(z) = \sum_{i=0}^{m} D_i z^{m-i}, \tag{14}$$

where $R(z)$ is a $(q+\eta)\times q$ polynomial matrix, define an MFD for $H(z)$.

Proof. By definition X satisfies Equation (3), and hence $D(z)$, $C(z)$, and therefore $D(z)R(z)$, $C(z)R(z)$ satisfy Equation (1), which implies that the polynomial matrices $N(z)$, $M(z)$ of Equation (14) define an MFD for $H(z)$. ■

The polynomial matrices $N(z)$ and $M(z)$ of the lemma above satisfy Equation (1), but do not give a factorization of $H(z)$ which is right coprime for all $R(z)$. The results below provide the basis of a procedure for the choice of $R(z)$ which result in irreducible MFDs.

LEMMA 5.2. *The denominator matrices of irreducible matrix fraction descriptions of a transfer function $H(z)$ all have the same nonunit invariant factors and therefore the same determinant.*

Proof. See [3, p. 446]. ■

THEOREM 5.1. *The pole polynomial of $H(z)$ is the product of the invariant factors of $D(z)$. Furthermore, a necessary condition for the $N(z)$ and $M(z)$ of Equation* (14) *to be right coprime is that $R(z)$ is irreducible, namely that all its invariant factors are trivial* (*i.e.*, *$R(z)$ does not have zeros*).

Proof. Let $Z(z)$ be the zero polynomial of $D(z)$ [i.e., let $Z(z)$ be the product of the invariant factors of $D(z)$]; then $\det M(z) = Z(z)f(z)$, where $f(z)$ is some polynomial for all $R(z)$. By considering the Smith form of $D(z)$ it is easy to show that there exists an $R(z)$ such that $f(z)$ is a constant, and thus $Z(z)$ is the greatest common divisor of $M(z)$ for all $R(z)$. The implication of this is that the roots of $Z(z)$ are the poles of $H(z)$. Now the zeros of a reducible $R(z)$ will be roots of $f(z)$, which will appear as roots of $\det M(z)$ in addition to the roots of $Z(z)$, and therefore by Lemma 5.2 such an $R(z)$ cannot give rise to a coprime MFD. ■

COROLLARY 5.1. *All $R(z)$ which lead to a right coprime factorization can be written as*

$$R(z) = U(z)E \tag{15}$$

where $U(z)$ is square and unimodular, whereas E is the matrix comprising the first q columns of the identity matrix of dimension $q + \eta$.

Proof. By Theorem 5.1 we have that the invariant factors of the $R(z)$ considered by the corollary are trivial. Thus from the Smith decomposition of $R(z)$ we have

$$R(z) = L_R(z)\begin{bmatrix} I_q \\ 0 \end{bmatrix} V_R(z) = U(z)E,$$

$$\text{where} \quad U(z) = L_R(z)\begin{bmatrix} V_R(z) & 0 \\ 0 & I_\eta \end{bmatrix} \text{ and } E = \begin{bmatrix} I_q \\ 0 \end{bmatrix}, \tag{16}$$

where $L(z)$, $V(z)$, and hence $U(z)$ are all unimodular. ■

COROLLARY 5.2. *There exists a unimodular $U(z)$ such that the solutions $N(z)$, $M(z)$ of Lemma 4.1 for the $R(z)$ of Corollary 4.1 are right coprime.*

Proof. Let the Smith decomposition of $D(z)$ be given as

$$D(z) = L_D(z)\begin{bmatrix} S_D(z) & 0 \end{bmatrix} V_D(z). \tag{17}$$

Then choosing the matrix $U(z)$ of Corollary 5.1 to be the inverse of $V_D(z)$ results in an $R(z)$ for which

$$M(z) = L_D(z)S_D(z). \tag{18}$$

The determinant of this $M(z)$ is $Z(z)$, and hence by Lemma 5.1 and Theorem 5.1 the pair $M(z)$, $N(z)$ for this particular choice of $R(z)$ define a right coprime MFD for $H(z)$. ∎

The corollary above indicates a procedure for the computation of $N(z)$ and $M(z)$, but it is based on the computation of the Smith canonical form of $D(z)$. This could be prone to problems of ill-conditioning, and for this reason we next consider an alternative procedure which is based on singular value decompositions and therefore is robust.

5.2. *Robust Algorithm for the Computation of Right Coprime $N(z)$, $M(z)$*

The overall strategy behind the procedure to be discussed in this section is simple, but the presentation of detailed steps is somewhat involved; the reader may prefer to get insight into the proposed algorithm by consulting the numerical example of Section 6, Example 6.2, first.

We begin by defining $D^{(1)}(z) = D(z)$ and denoting the degree of the ith column of $D^{(1)}(z)$ by $m_i^{(i)}$. Furthermore we assume that $m = m_1^{(1)} \geqslant m_2^{(1)} \geqslant \cdots \geqslant m_{q+\nu}^{(1)}$; this condition can always be achieved if $D^{(1)}(z)$ is postmultiplied by a square permutation matrix $E^{(0)}$ with the view to reordering appropriately the columns of $D^{(1)}(z)$. Then $D^{(1)}(z)$ can be written as

$$D^{(1)}(z) = D_{hc}^{(1)} \operatorname{diag}\{z^{m_1^{(1)}}, z^{m_2^{(1)}}, \dots, z^{m_{q+\eta}^{(1)}}\} + D_{lc}^{(1)}(z), \tag{19}$$

where $D_{hc}^{(1)}$ is the matrix of the same dimensions as $D^{(1)}(z)$ and whose ith column comprises the coefficients of $z^{m_i^{(1)}}$ in the ith column of $D^{(1)}(z)$. Note that $D_{lc}^{(1)}$ is the difference between $D^{(1)}(z)$ and $D_{hc}^{(1)} \operatorname{diag}\{z^{m_1^{(1)}}, z^{m_2^{(1)}}, \dots, z^{m_{q+n}^{(1)}}\}$ and therefore is a polynomial matrix with column degrees strictly less than the corresponding column degrees of $D^{(1)}(z)$ (see also [3, p. 384]). The excess of columns over rows in $D_{hc}^{(1)}$ is η, and hence this matrix possesses a κ-dimensional kernel, where $\kappa \geqslant \eta$. An orthogonal matrix representation, $Y^{(1)}$, of the kernel can be computed by: (i) performing the singular value decomposition of $D^{(1)}(z)$; (ii) collecting the singular vector $y_i^{(1)}$, $i = 1, \dots, \kappa$, which satisfy the condition $D_{hc}^{(1)} y_i^{(1)} = 0$; and (iii) writing

$$Y^{(1)} = \left[y_1^{(1)}, y_2^{(1)}, \dots, y_k^{(1)} \right]. \tag{20}$$

Assume without loss of generality, that the order of the columns of $Y^{(1)}$ is such that the first nonzero element of $y_i^{(1)}$ is the p_ith (where $0 < p_i \leqslant q + \eta$) and is preceded by κ_i zeros, where $\kappa_{i+1} \geqslant \kappa_i$ for all $i = 1, \dots, \kappa$, with $\kappa_1 \geqslant 0$. Next discard from $Y^{(1)}$ all columns $y_j^{(1)}$ for which $j > i$ and $p_j = p_i$ to obtain a subset $\{\bar{y}_1^{(1)}, \bar{y}_2^{(1)}, \dots, \bar{y}_{\bar{\kappa}}^{(1)}\}$ with $\bar{\kappa} \leqslant \kappa$, and let $\bar{p}_i$ denote the position of the first nonzero element of $\bar{y}_i^{(1)}$. Using the notation $\bar{y}_{ji}^{(1)}$ for the jth element of $\bar{y}_i^{(1)}$ define the polynomial vector $u_{\bar{p}_i}^{(1)}(z)$ whose jth component $u_{j,\bar{p}_i}^{(1)}(z)$ is as follows:

$$u_{j,\bar{p}_i}^{(1)} = \begin{cases} 0, & j \leqslant \bar{p}_i - 1, \\ \bar{y}_{j,\bar{p}_i}^{(1)} z^{m_{\bar{p}_i} - m_j}, & \bar{p}_i \leqslant j \leqslant q + \eta, \end{cases} \tag{21}$$

and form the matrix $U^{(1)}(z)$ by replacing the $\bar{p}_i$th column of $I_{q+\eta}$ by $u_{\bar{p}_i}^{(1)}(z)$ for $i = 1, \dots, \bar{\kappa}$. Note therefore that by construction $U^{(1)}(z)$ is lower triangular with real constant diagonal elements, and hence is unimodular.

Because of the orthogonality of the $\bar{y}_i^{(1)}$ to the vectors defined by the rows of $D_{hc}^{(1)}$ and on account of the definition of the powers of z of the elements of $U^{(1)}(z)$, it is easy to show that the $\bar{p}_i$th column of the product $[D^{(1)}(z) - D_{lc}^{(1)}(z)]U^{(1)}(z)$ will be zero. The consequence of this is that the degree of the $\bar{p}_i$th column of the product $D^{(1)}(z)U^{(1)}(z)$ is less than or equal to $m_{\bar{p}_i} - 1$, $i = 1, \dots, \bar{\kappa}$. On account of this, the degrees of the columns in this product will not necessarily appear in descending order any longer, but this can be arranged for by a postmultiplication by an appropriate permutation matrix $E^{(1)}$ which will give

$$D^{(2)}(z) = D^{(1)}(z)U^{(1)}(z)E^{(1)} = D_{hc}^{(2)} \operatorname{diag}\{z^{m_1^{(2)}}, z^{m_2^{(2)}}, \dots, z^{m_{q+\eta}^{(2)}}\} + D_{lc}^{(2)}(z). \tag{22}$$

The matrix above is of exactly the same form as the matrix $D^{(1)}(z)$ of Equation (19), and therefore, from this point on, the whole procedure described above can be reapplied to $D^{(2)}(z)$, in order to reduce further the degree of some of its columns.

We can carry on until no further column degree reduction is possible. This will happen at the rth iteration, for which $D^{(r)}(z)$ has precisely η columns which are identically zero, and hence only the last η rows of $Y^{(r)}$ will be nonzero. Then defining

$$U(z) = E^{(0)} \prod_{i=1}^{r} U^{(i)}(z) E^{(i)} \tag{23}$$

and using the matrix E of Equation (16), we obtain an $R(z)$ for which $M(z) = D(z)R(z)$ comprises the first q nonzero columns of $D^{(r)}(z)$, and hence the invariant factors of $M(z)$ will be exactly those of $D(z)$. Hence $\det M(z) = Z(z)$, and so the pair $N(z) = C(z)R(z)$, $M(z) = D(z)R(z)$ define a right coprime MFD for $H(z)$.

REMARK 5.1. During each cycle of the procedure above, a certain number of singular vectors are discarded in forming $U^{(i)}(z)$, and as a result we end up reducing the degree of $\bar{\kappa}$ rather than κ columns, where $\bar{\kappa} \leqslant \kappa$. There exist circumstances where discarding is not necessary and the maximum possible number, κ, of degree reductions can be implemented. Thus assume that: (i) there exists an integer t, $1 \leqslant t \leqslant q + \eta$, such that $m_j^{(i)} = m_t^{(i)}$ for all $j \geqslant t$, (ii) $m_j^{(i)} > m_t^{(i)}$ for all $j < t$, and (iii) the first $t - 1$ rows of $Y^{(i)}$ are zero; and let the nonzero block of the resulting matrix be V. Then the matrix

$$U^{(i)} = \begin{bmatrix} e_1 & \cdots & e_{t-1} & \begin{matrix} 0_{t-1,\kappa} \\ V \end{matrix} & \begin{matrix} 0_{t-1,q+\eta-t-\kappa+1} \\ V^c \end{matrix} \end{bmatrix}, \tag{24}$$

where V^c is a matrix such that the matrix $[V, V^c]$ is square and full rank, can be used in place of $U^{(i)}(z)$, and it is easy to show that it will have the effect of reducing the degree of κ columns.

REMARK 5.2. Implicit in the development so far is the assumption that Algorithm 4.1 has been invoked in order to reduce n, m so that the rank defect $\nu_{m,n}$ is $q + \eta$ and therefore is never greater than $2q$. It is pointed out that this was done for computational convenience and in particular in order to reduce the dimension of $S_{m,n}(B, A)$. The procedure described in this section does not make use of the fact that $\eta < q$ and can be applied whether this is true or not.

6. NUMERICAL EXAMPLES

Two numerical examples to illustrate the results of Sections 3 and 4, respectively, are now presented. For simplicity the first of these is chosen such that $\eta = 0$, and hence does not require the application of the procedure of Section 5.2.

EXAMPLE 6.1. Consider an $H(z)$ for which

$$B(z) = \begin{bmatrix} z^4 - z^3 - 3z^2 + 2z & -2z^3 + z^2 + 5z - 3 \\ -z^2 - z + 1 & z^5 - 2z^3 - z^2 - 2z + 2 \end{bmatrix},$$

$$A(z) = d(z) I_2, \quad \text{where} \quad d(z) = z^5 - z^4 - 2z^3 - 4z + 3. \tag{25}$$

Clearly the factorization above is both a left and a right MFD for $H(z)$, but the key point is that it is not coprime. For this example $\gamma_0 = 0$ and thus $\alpha = \beta = 5$, and the resulting Sylvester matrix $S_{5,5}(B, A)$ is a 22×24 matrix for which $\mu_{5,5} = 7$, and so $\nu_{5,5} = 9$. Thus $\chi = 3$, $\eta = 1$, and by Theorem 4.1(iii) we have that $\nu_{2,2} = 3$, and hence it is clear that n, m can be chosen to be $5 - \chi = 2$ or less. However, since $\rho_{B_0} = 1$, Theorem 4.1(v) implies that $\nu_{2,1} \geqslant 2$, and so n can be chosen to be 1. Application of Algorithm 4.1 confirms that $m = 2$ and $n = 1$ represent the smallest allowable values for m, n. The kernel of $S_{2,1}(B, A)$ is two dimensional, and therefore X is a 10×2 matrix comprising five 2×2 blocks: D_0, D_1, D_2, C_0, C_1. Given these dimensions, it is clear that the matrix $R(z)$ of Equation (16) can be chosen to be a constant, and here for convenience we select $R(z) = C_1^{-1}$ and get as our right coprime factorization

$$M(z) = D_0 C_1^{-1} z^2 + D_1 C_1^{-1} z + D_2 C_1^{-1} = \begin{bmatrix} z+2 & z^2+3 \\ z-1 & z \end{bmatrix},$$

$$N(z) = C_0 C_1^{-1} z + I_2 = \begin{bmatrix} 1 & z \\ z & z+1 \end{bmatrix}, \tag{26}$$

EXAMPLE 6.2. Consider a left MFD with denominator and numerator matrices

$$A(z) = \begin{bmatrix} 14z^3 + 5z^2 + 36z - 21 & -31z^3 - 34z^2 - 39z + 3 & 41z^3 + 35z^2 + 70z - 13 \\ -19z^3 + 34z^2 - z - 4 & 33z^3 + z^2 + 4z + 1 & -52z^3 + 22z^2 - 3 \\ -2z^3 + 34z^2 - 7z - 7 & 2z^3 - 148z^2 - 35z + 3 & -10z^3 + 163z^2 + 29z - 6 \end{bmatrix},$$

$$B(z) = \begin{bmatrix} -18z^3 - 21z^2 + 76z - 27 & -5z^3 - 3z^2 + 58z + 1 & -z^3 + 11z^2 - 91z + 14 \\ 7z^3 + 86z^2 + 5z - 8 & -24z^3 + 42z^2 + 22z - 1 & 26z^3 + 69z^2 - 15z + 5 \\ 2z^3 - 230z^2 + 55z - 13 & -28z^3 - 128z^2 + 1 & 18z^3 + 133z^2 - 41z + 7 \end{bmatrix}. \tag{27}$$

Following the steps of Algorithm 5.1, it is found that $n = m = 1$ represents the smallest choice for m, n and that $\nu_{1,1} = 5$ and $\eta = 2$, which is nonzero: hence the need for the procedure of Section 5.2 for the computation of $R(z)$.

Performing the singular value decomposition of $S_{1,1}(B, A)$, we get

$$D(z) = 0.1\begin{bmatrix} -3.847 & 0.332 & -0.216 & 1.086 & -0.034 \\ -4.085 & -0.598 & 0.153 & 0.656 & 2.456 \\ -7.384 & 0.117 & 0.605 & -0.539 & 2.632 \end{bmatrix} z$$

$$+ 0.1\begin{bmatrix} -0.299 & -2.312 & 2.552 & 0.932 & -1.719 \\ -0.063 & -1.328 & 4.606 & 0.394 & -0.613 \\ 0.450 & -3.985 & 5.976 & 1.051 & -2.466 \end{bmatrix},$$

$$C(z) = 0.1\begin{bmatrix} -1.798 & 1.936 & -2.113 & 5.20 & -6.822 \\ 0.002 & -2.881 & -1.760 & 6.956 & 2.751 \\ 2.983 & -2.916 & 0.548 & 2.940 & 4.061 \end{bmatrix} z$$

$$+ 0.1\begin{bmatrix} -0.501 & -3.054 & -1.305 & -0.978 & -1.283 \\ 0.763 & 4.090 & 4.852 & 1.801 & 1.685 \\ 0.853 & 5.470 & 1.739 & 2.769 & 1.594 \end{bmatrix}, \quad (28)$$

and so clearly $D(z)$ in the form of $D^{(1)}(z)$. Hence, performing the singular value decomposition of the matrix coefficient of z in $D(z)$, we derive

$$Y^{(1)} = \begin{bmatrix} -0.1362 & 0 \\ -0.8727 & 0.1671 \\ -0.1526 & -0.9441 \\ -0.2574 & -0.2336 \\ -0.3608 & 0.1617 \end{bmatrix},$$

$$U^{(1)}(z)E^{(1)} = \left[y_1^{(1)}, y_2^{(2)}, e_3, e_4, e_5 \right]\left[e_3, e_4, e_5, e_1, e_2 \right], \quad (29)$$

where e_i, $i = 1, \ldots, 5$, denotes the ith column of I_5. The column degrees of the resulting $D^{(2)}(z)$ are $\{1, 1, 1, 0, 0\}$, and the corresponding matrix of column coefficients and a matrix representation of its kernel are

$$D_{hc}^{(2)} = 0.1\begin{bmatrix} -0.216 & 1.086 & -0.034 & 2.049 & -3.292 \\ 0.153 & 0.0656 & 2.456 & 0.585 & -4.761 \\ 0.605 & -0.539 & 2.632 & 3.246 & -6.952 \end{bmatrix},$$

$$Y^{(2)} = 0.1\begin{bmatrix} -9.577 & 0 \\ -2.548 & -4.047 \\ 1.263 & -6.250 \\ 0.439 & -5.022 \\ 0.048 & -4.398 \end{bmatrix}. \tag{30}$$

Then using Equation (21) we obtain

$$U^{(2)}(z)E^{(2)} = \begin{bmatrix} -0.9577 & 0 & & & \\ -0.2548 & -0.4047 & & & \\ 0.1263 & -0.6250 & e_3 & e_4 & e_5 \\ 0.0439z & -0.5022z & & & \\ 0.0048z & -0.4398z & & & \end{bmatrix}, \tag{31}$$

and for this the column degrees of $D^{(3)}(z)$ are $\{1, 0, 0, 0, 0\}$; the matrix representation of the kernel of $D_{hc}^{(3)}$ and the corresponding $U^{(3)}(z)E^{(3)}$ are given as

$$Y^{(3)} = \begin{bmatrix} 0 & 0 \\ -0.6061 & -0.3663 \\ 0.5176 & -0.7968 \\ -0.0699 & 0.3236 \\ 0.5999 & 0.3552 \end{bmatrix},$$

$$U^{(3)}(z)E^{(3)} = \begin{bmatrix} e_1 & Y^{(3)} & e_4 & e_5 \end{bmatrix}\begin{bmatrix} e_1 & e_4 & e_5 & e_2 & e_3 \end{bmatrix}. \tag{32}$$

This terminates the procedure of Section 5.2, because the last two columns of the product $D^{(3)}(z)U^{(3)}(z)E^{(3)}$ are zero. Then forming $R(z)$ as per Equations (15) and (23), we end up with the right coprime MFD:

$$M(z) = \begin{bmatrix} -0.034z - 1.719 & 2.049 & -3.292 \\ 2.456z - 0.613 & 0.585 & -4.761 \\ 2.632z - 2.466 & 3.246 & -6.952 \end{bmatrix},$$

$$N(z) = \begin{bmatrix} -6.822z - 1.283 & 3.648 & 0.743 \\ 2.751z + 1.685 & -5.486 & -4.045 \\ 4.061z + 1.594 & -6.444 & 1.117 \end{bmatrix}. \tag{33}$$

6. CONCLUSIONS

In this paper we have reexamined the problem of computing right coprime MFDs and have proposed a simple and robust algorithm. In concluding we point out that the algorithm proposed in Section 5.2 can also be used to compute column reduced equivalents of polynomial matrices.

The authors wish to thank the Brazilian Ministry of Education (CAPES) for financial support.

REFERENCES

1 H. H. Rosenbrock, *State Space and Multivariable Theory*, Nelson, London, 1970.
2 W. A. Wolovich, *Linear Multivariable Systems*, Springer-Verlag, New York, 1974.
3 T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1980.
4 D. C. Youla, H. A. Jabr, and J. J. Bongiorno, Modern Wiener-Hopf design of optimal controllers—part II: The multivariable case, *IEEE Trans. Automat. Control* 21:319–338 (1976).
5 V. Kucera, *Discrete Linear Control: The Polynomial Equation Approach*, Wiley, Chichester, U.K., 1979.
6 B. Kouvaritakis and J. A. Rossiter, Multivariable stable generalized predictive control, *Proc. IEEE Pt. D* 140:364–372.
7 R. V. Patel, Computation of matrix fraction description of linear time-invariant systems, *IEEE Trans. Automat. Control* 26:148–161 (1981).
8 K. B. Datta and S. Gangopadhyay, Coprime matrix fraction description via orthogonal structure theorem, *IEEE Trans. Automat. Control* 37:1517–1520 (1992).
9 S. Wang and E. J. Davison, A minimization algorithm for the design of linear multivariable systems, *IEEE Trans. Automat. Control* 18:220–225 (1973).
10 G. D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, *SIAM J. Control* 13:493–520 (1975).
11 S. Kung and T. Kailath, Fast projection methods for minimal design problems in linear system theory, *Automatica* 16:399–403 (1980).
12 K. B. Datta and S. Gangopadhyay, Coprime matrix fraction description via Sylvester resultant matrix, *Control Theory Adv. Tech.* 4:971–981 (1993).
13 B. D. O. Anderson and E. I. Jury, Generalized Bezoutian Sylvester matrices in multivariable linear control, *IEEE Trans. Automat. Control* 212:551–556 (1976).

Received 21 August 1995; final manuscript accepted 28 October 1996