INT. J. CONTROL, 1986, VOL. 43, NO. 6, 1745–1761

ARMA models, their Kronecker indices and their McMillan degree

MICHEL R. GEVERS†

This paper highlights some difficulties with the use of ARMA models with leading unit coefficient matrix in system identification. It is shown that the McMillan degree of such models is not in any easy way related to the row degrees of the polynomial factors of the ARMA model. A rank test is given for the McMillan degree of such models and it is shown that this degree will generically be a multiple of the dimension of the observation vector.

1. Introduction

Not much mystery remains about canonical forms for state-space (SS) or matrix-fraction description (MFD) representations for linear multivariable systems of finite order. It is well-known that the McMillan degree (i.e. the order) of the system is the dimension of the state of any minimal state-space representation, and is the sum of the row (column) degrees of the denominator matrix of any row-reduced left coprime (column-reduced right coprime) MFD of its matrix transfer function. Canonical MFDs are therefore defined by their row (or column) degrees, and these will determine the number of free parameters in these canonical descriptions: see e.g. Guidorzi (1981). These row (or column) degrees are in turn determined by the left (or right) Kronecker indices of the matrix transfer function $K(z)$. It is also well established that the set $S_n(n_1, \dots, n_p)$ of all $p \times m$ matrix transfer functions $K(z)$ with, say, left Kronecker indices $(n_1, \dots, n_p)$ with $\sum_1^p n_i = n$ is an analytic manifold whose dimension d is entirely determined by these Kronecker indices: $d = d(n_1, \dots, n_p)$. In addition $d(n_1, \dots, n_p) \leq n(p+m)$, and generically $d(n_1, \dots, n_p) = n(p+m)$. In most standardly used canonical MFDs the number of free parameters is precisely $d(n_1, \dots, n_p)$. All this has been extensively described in a number of papers: see e.g. Clark (1976), Hazewinkel and Kalman (1976), Deistler (1985), Hannan and Kavalieris (1984).

So then why another paper on canonical forms? It turns out that most of the studies on canonical forms have been for SS or MFD models, because these are the most widely used in control. In econometrics, and also in system identification, it is often much more natural to use ARMA or ARMAX models. Here we shall concentrate on ARMA models:

$$A(D)y(t) = B(D)u(t) \tag{1.1 a}$$

where D is the unit delay operator $(Dy(t) = y(t-1))$, dim $y = p$, dim $u = m$ and

$$A(D) = A_0 + A_1 D + \dots + A_q D^q, \quad A_q \neq 0 \tag{1.1 b}$$

$$B(D) = B_0 + B_1 D + \dots + B_r D^r, \quad B_r \neq 0 \tag{1.1 c}$$

These models are usually referred to as ARMA(q, r) models. ARMA models are

Received 5 September 1985.

† Department of Systems Engineering, Australian National University, GPO Box 4, Canberra, ACT 2601, Australia.

Since ARMA models and MFD models are connected in such an obvious way, one would expect that the properties of canonical MFD models mentioned above should apply almost unchanged to canonical ARMA models. Perhaps surprisingly, this is not the case. In fact, the following points will be made in this paper, some of which are fairly obvious and some of which are not:

(i) if $A^{-1}(D)B(D)$ is a left coprime ARMA model for $K(z)$ with $A(D)$ row proper, this does not imply that the row degrees of $A(D)$ are connected to the left Kronecker indices of $K(z)$;

(ii) the McMillan degree of $K(z)$ cannot be obtained from such $A(D)$;

(iii) given a $K(z) \in S_n(n_1, \ldots, n_p)$, it is generally not possible to construct a monic ARMA model such that the number of free parameters is equal to $d(n_1, \ldots, n_p)$, the dimension of $S_n(n_1, \ldots, n_p)$;

(iv) monic ARMA models will generically represent systems whose McMillan degree is a multiple of p, the dimension of the observation vector.

The last problem can be relaxed by taking a non-monic ARMA model (i.e. $A_0 \neq I$). However, such models lose the major advantage of ARMA models mentioned above.

Comment 1.1

The four drawbacks just mentioned arise when ARMA models are derived from canonical MFD models through the relationship (1.6). These canonical MFD models are themselves obtained by selecting a basis of the observability matrix of a state-space representation of the system (or, equivalently, of the Hankel matrix of that system), and there are several ways of selecting such a basis. It has recently been shown, however, that when the system has no poles at the origin, a canonical ARMA model can be derived directly from any state-space model by selecting a basis of the constructibility matrix (rather than the observability matrix). This requires a non-singular state-transition matrix and corresponds to constructing the present state from past observations as opposed to future observations, which is a logical thing to do for ARMA models. The column degrees of the canonical monic ARMA model then coincide with the constructibility indices: see Bokor and Keviczky. (1985).

It is apparent from our statements above that ARMA models exhibit a number of limitations and drawbacks which do not seem to be well recognized in the control literature. It was argued in Bokor and Keviczky (1982), for example, that one could always transform a canonical MFD into a canonical monic ARMA model with $n(p+m)$ parameters: we shall show that this is not correct. It was argued in Stoica (1982) that fully parametrized ARMA models could represent almost all systems and that those that could not be so represented could be considered 'pathological'. It is the purpose of this paper to correct some of these statements and to draw attention to some of the limitations of ARMA models. In addition we present a new criterion which gives the McMillan degree of a monic model as a function of the parameter matrices A_i, B_i of that model.

Statisticians have for quite some time been aware of some of the limitations mentioned above and, in particular, the fact that the McMillan degree of monic ARMA models will generically be a multiple of p: see for example Akaike (1974), Hannan (1976). The fact that they perform their search over models whose McMillan degree increases by p each time the length of the ARMA model is increased by 1 does not concern statisticians too much because their data always comes from infinite-dimensional systems anyway, and because they argue that a system whose McMillan degree is not a multiple of p can be approximated arbitrarily closely by a system of higher degree. This might also explain why statisticians have not been interested in establishing the precise connection between the parameters of an ARMA model and its McMillan degree, which we present in § 5. In engineering applications, the underlying system can often be assumed finite and it is often important to obtain a minimal model, particularly if the model is to be used to design a controller. A more important aspect, however, which applies equally well to statistical and engineering applications, is the curse of dimensionality. The number of parameters of a monic ARMA model is increased by $p(p+m)$ each time the length of the model is increased by 1; the corresponding increase with MFD models is only $p+m$ each time one of the structure indices is increased by 1. A consequence is that in practical applications multivariate ARMA models have always been limited to very low lag-lengths (typically $\max(q, r) \leq 2$).

The paper is organized as follows. In § 2 we recall some basic mathematical notions about the McMillan degree and the Kronecker indices of a rational transfer function matrix, and their connection with the row degrees of left coprime MFDs of such a transfer function matrix. In § 3 we briefly review canonical MFDs and we derive canonical ARMA models from these where A_0 is not necessarily equal to I. We also study the degree properties of these ARMA models. Monic ARMA models are studied in § 4. Finally, in § 5 we give a new formula for the calculation of the McMillan degree of a monic ARMA model with $A(D)$ and $B(D)$ left coprime, and we show that such ARMA models will generically represent systems of McMillan degree $k \times p$, where $p = \dim y$ and k is an integer.

2. Transfer functions, Kronecker indices and McMillan degree

As a starting point we consider a $p \times m$ rational matrix $K(z)$ (or equivalently $\bar{K}(D)$) that is strictly proper, i.e.

$$\lim_{z \to \infty} K(z) = \lim_{D \to 0} \bar{K}(D) = 0 \tag{2.1}$$

$K(z)$ can then be written as

$$K(z) = K_1 z^{-1} + K_2 z^{-2} + \dots \tag{2.2}$$

We shall often refer to $K(z)$ as 'the system'.

With $K(z)$ we associate a Hankel matrix $\mathscr{H}_{1,N}[K]$, defined as follows:

$$\mathscr{H}_{1,N}[K] \triangleq \begin{bmatrix} K_1 & K_2 & \dots & K_N \\ K_2 & K_3 & \dots & K_{N+1} \\ \vdots & & & \\ K_N & K_{N+1} & \dots & K_{2N-1} \end{bmatrix} \tag{2.3}$$

The rank of $\mathscr{H}_{1,\infty}[K]$ is called the order of the system or, equivalently, the McMillan degree of $K(z)$, denoted $\delta[K(z)]$. $\mathscr{H}_{1,\infty}[K]$ is made up of block rows of size p and block columns of size m. We denote by $r(i,j)$ the ith row of the jth block row of $\mathscr{H}_{1,\infty}[K]$. Similarly we denote by $c(i,j)$ the ith column of the jth block column of $\mathscr{H}_{1,\infty}[K]$. For each $i \varepsilon \{1, \dots, p\}$, let $r(i, n_i+1)$ be the first row $r(i,j)$ that is linearly dependent on all rows above it in $\mathscr{H}_{1,\infty}[K]$. For $i = 1, \dots, p$ this defines p integers $n_1, \dots, n_p$. These are called the left Kronecker indices, or the output Kronecker indices or the observability indices. This last terminology comes from the fact that the linear dependence relations on the rows of $\mathscr{H}_{1,\infty}[K]$ are identical to the linear dependence relations on the rows of the observability matrix of any minimal state space realization of $K(z)$. We shall call $\mathrm{Kr}_0 = \{n_1, \ldots, n_p\}$ the ordered set of Kronecker indices n_i for increasing i, while Kr_u will denote the unordered set. For example, if $\mathrm{Kr}_0 = \{2, 1\}$ then Kr_u is the collection of ordered sets $\{2, 1\}$ and $\{1, 2\}$. We shall denote $\rho \triangleq \max\{n_i\}$. By applying exactly the same procedure to the columns of $\mathscr{H}_{1,\infty}[K]$ one can similarly define the m right Kronecker indices $\nu_1, \ldots, \nu_m$, also called input Kronecker indices or controllability indices.

We shall now state a number of facts concerning Kronecker indices and their relationship to the McMillan degree and to coprime MFDs of $K(z)$. They will be stated for left Kronecker indices and, correspondingly, for left coprime MFDs. By duality, identical results exist for right Kronecker indices and right coprime MFDs. Most of these results are well known and will therefore be stated without proof.

Lemma 2.1 (see for example Kailath 1980)

The McMillan degree of $K(z)$ is the sum of the left (right) Kronecker indices of $K(z)$: $n = \sum_1^p n_i = \sum_1^m \nu_i$.

Lemma 2.2 (see for example Popov 1972)

A permutation of the rows of $K(z)$ (which corresponds to a relabelling of the components of the output vector $y(t)$) leaves Kr_u unchanged, but it may change Kr_0.

This result is very important for our future arguments: we illustrate it with an example.

Example 2.1

Consider the systems

$$K_1(z) = \begin{bmatrix} \dfrac{1}{z(z-0{\cdot}5)} & \dfrac{1}{z} \\ \dfrac{1}{z(z-0{\cdot}5)} & 0 \end{bmatrix} \quad \text{and} \quad K_2(z) = \begin{bmatrix} 0 & \dfrac{1}{z} \\ \dfrac{1}{z(z-0{\cdot}5)} & 0 \end{bmatrix} \tag{2.4}$$

It is easy to compute that $K_1(z)$ has $\mathrm{Kr}_0 = \{2, 1\}$ while $K_2(z)$ has $\mathrm{Kr}_0 = \{1, 2\}$. By permuting the rows we get

$$\bar{K}_1(z) = \begin{bmatrix} \dfrac{1}{z(z-0{\cdot}5)} & 0 \\ \dfrac{1}{z(z-0{\cdot}5)} & \dfrac{1}{z} \end{bmatrix} \quad \text{and} \quad \bar{K}_2(z) = \begin{bmatrix} \dfrac{1}{z(z-0{\cdot}5)} & 0 \\ 0 & \dfrac{1}{z} \end{bmatrix} \tag{2.5}$$

We now find $\mathrm{Kr}_0 = \{2, 1\}$ for $\bar{K}_1(z)$ and $\bar{K}_2(z)$. Notice that Kr_u remains unchanged by the permutation.

Lemma 2.3

Let $P^{-1}(z)Q(z) = K(z)$ be a polynomial left coprime MFD of a strictly proper $K(z)$ with $P(z)$ row reduced. Then the row degrees of $P(z)$ are the left Kronecker indices of

$K(z)$. Hence deg det $P(z) = n = \delta[K(z)]$. In addition these row degrees can be arranged in arbitrary order.

Proof

The first part of the proof can be found in Wolowich (1974). That the row degrees of $P(z)$ can be arranged in arbitrary order follows from the fact that, if a particular (say canonical) MFD has $P(z)$ with row degrees $n_1, \ldots, n_p$ in that order, then other row-reduced MFDs of $K(z)$ can be obtained by permuting the rows of $P(z)$ and the corresponding rows of $Q(z)$ in the same way.

Hence there is no connection between the ordered left Kronecker indices of $K(z)$ and the ordered row degrees of a row-reduced left coprime MFD of $K(z)$, except that the unordered sets are identical. Notice also that permuting the rows of $P(z)$ and, correspondingly, of $Q(z)$ changes the ordering of the I/O equations, but not the ordering of the input or output components. Lemma 2.3 has led a number of authors to construct canonical left coprime MFDs or to derive properties of left coprime MFDs under the simplifying assumption that the row degrees could be ordered in a specified way, e.g. $n_1 \leqq \ldots \leqq n_p$ or $n_1 \geqq \ldots \geqq n_p$. There is nothing wrong with that, but it has in turn led some people to believe that the Kronecker indices of $K(z)$ could be so arranged by a permutation of its rows. This is not correct. The next lemma examines what can be achieved by permuting the output components; this corresponds to permuting the rows of $K(z)$ and the columns of $P(z)$ in $K(z) = P^{-1}(z)Q(z)$, while leaving $Q(z)$ unchanged. This lemma is probably well known to researchers in the field, but we have not been able to find a statement or a proof of this result. We therefore give a complete proof.

Lemma 2.4

By permuting the rows of $K(z)$ one can always arrange the left Kronecker indices in decreasing order (i.e. $n_1 \geqq n_2 \geqq \ldots \geqq n_p$), but not always in increasing order.

Proof

The second part is proved by Example 2.1. We now prove the first part. Let $n_1, \ldots, n_p$ be the ordered Kronecker indices of $K(z)$. We denote by $r(i, k)$ the ith row of the kth block of $\mathscr{H}_{1,\infty}[K]$ and by ant $r(i, k)$ the antecedents of $r(i, k)$, i.e. the set of rows above $r(i, k)$ in the Hankel matrix. The proof is best illustrated by the crate diagrams of Figs. 1*a* and 1*b*. In these diagrams the element in position (i, k) represents $r(i, k)$. Each column of the crate therefore represents a block of rows of $\mathscr{H}_{1,\infty}[K]$, with the leftmost column representing the first block. The antecedents of $r(i, k)$ correspond to all the elements above and to the left of (i, k) in the crate diagram. The crosses in Fig. 1 indicate the linearly independent rows obtained by searching from top to bottom in the Hankel matrix, or from top left to bottom right in the crate, going down column by column. The circles indicate the first linearly dependent rows of $\mathscr{H}_{1,\infty}[K]$. The figures are drawn for a system with five outputs.

By definition of the n_i we have

$$r(i, n_i + 1) = \sum_{k=1}^{p} \sum_{l=1}^{n_{ik}} \alpha_{ikl} r(k, l) \quad i = 1, \ldots, p \tag{2.6}$$

×	×	×	×	×	○	
×	×	×	×	○		
×	×	○				
×	×	×	○			
×	×	×	○			

Figure 1*a*. ($\bar{K}(z)$).

×	×	×	×	×	○	
×	×	×	×	○		
×	×	×	○			
×	×	○				
×	×	×	○			

Figure 1*b*. ($\tilde{K}(z)$).

where

$$n_{ik} \triangleq \min(n_i, n_k) \quad \text{if } i \leq k$$
$$\triangleq \min(n_i + 1, n_k) \quad \text{if } i > k \tag{2.7}$$

This expresses that $r(i, n_i + 1)$ is a linear combination of its linearly independent antecedents. Our proof uses an initialization step and an induction step.

Initialization step

Let $n_j \triangleq \max_{i=1,\ldots,p} \{n_i\}$. If $j \neq 1$, permute the first and jth rows of $K(z)$. Denote the new matrix $\bar{K}(z)$, its rows $\bar{r}(k, l)$ and its associated Kronecker indices $\mu_1, \ldots, \mu_p$. Since ant $\bar{r}(1, n_j + 1) \subset$ ant $r(j, n_j + 1)$, it follows that $\mu_1 \geq n_j$, and since $n_j = \max\{n_i\}$, it follows by Lemma 2.2 that $\mu_1 = n_j$.

Induction step

Let k be the largest index such that in $\bar{K}(z)$

(i) $\mu_1 \geq \ldots \geq \mu_{k-1}$ (2.8 *a*)

(ii) $\mu_{k-1} \geq \mu_i \quad i = k, \ldots, p \quad \text{if } k - 1 < p$ (2.8 *b*)

The initialization step ensures that $k - 1 \geq 1$. If $k - 1 = p$, the desired ordering is achieved. If $k - 1 < p$, let $\mu_j \triangleq \max_{i=k,\ldots,p} \{n_i\}$. Then, necessarily, $\mu_{k-1} \geq \mu_j > \mu_k$. (If Fig. 1*a* is representative of $\bar{K}(z)$, then $k - 1 = 2$ and $j = 4$.) Permute the kth and jth rows of $\bar{K}(z)$, and denote $\tilde{K}(z)$ the new matrix, $\tilde{r}(k, l)$ its rows and $v_1, \ldots, v_p$ its Kronecker indices: see Fig.1*b*. We show that $v_k = \mu_j$. Since ant $\tilde{r}(i, v_i + 1) =$ ant $\bar{r}(i, \mu_i + 1)$ for $i = 1, \ldots, k - 1$, it follows that

$$v_i = \mu_i \quad i = 1, \ldots, k-1 \tag{2.9}$$

Since ant $\tilde{r}(k, n_j + 1) \subset$ ant $\bar{r}(j, n_j + 1)$, it follows that $v_k \geq \mu_j$, and since $\mu_j \triangleq \max_{i=k,\ldots,p} \{\mu_i\}$ it follows by Lemma 2.2 that $v_k = \mu_j$. Hence we now have

(i) $v_1 \geq \ldots \geq v_k$ (2.10 *a*)

(ii) $v_k \geq v_i \quad i = k + 1, \ldots, p \quad \text{if } k < p$ (2.10 *b*)

Comparing (2.8) and (2.10) we see that the induction step has increased the number of ordered Kronecker indices by at least one. Repeating this step a finite number of times therefore leads to the desired ordering.

Comment 2.1

Note that permuting two rows i and j may affect other indices than n_i and n_j.

Finally, we denote by $S(n)$ the set of all rational strictly proper $K(z)$ of McMillan degree n, and by $S_n(n_1, \ldots, n_p)$ the set of all such $K(z)$ whose ordered left Kronecker indices are $n_1, \ldots, n_p$ with $\sum_1^p n_i = n$. We then have the following result (see e.g. Deistler and Hannan (1981), Hazewinkel and Kalman (1976)).

Lemma 2.5

(i) The $S_n(n_1, \ldots, n_p)$ are disjoint subsets of $S(n)$ with $\cup S_n(n_1, \ldots, n_p) = S(n)$

(ii) $S_n(n_1, \ldots, n_p)$ can be mapped homeomorphically into an open and dense subset of $R^{d(n_1, \ldots, n_p)}$, where

$$d(n_1, \ldots, n_p) = n(m+1) + \sum_{i<j} \{\min(n_i, n_j) + \min(n_i, n_j+1)\} \quad (2.11)$$

$$\leqq n(m+p) \quad (2.12)$$

3. Canonical MFDs and their ARMA equivalent

The rows appearing on the right-hand side of (2.6) form a basis for the row space of $\mathscr{H}_{1,\infty}[K]$. It is clear from the structure of the Hankel matrix that the first m elements of these rows, together with the coefficients α_{ikl} of (2.6), completely determine the whole $\mathscr{H}_{1,\infty}[K]$, and therefore $K(z)$. They form a complete system of independent invariants (Guidorzi 1981):

$$\{\alpha_{ikl}, r_i(k, j); \quad i, k = 1, \ldots, p; \quad l = 1, \ldots, n_{ik}; \quad j = 1, \ldots, n_k\} \quad (3.1)$$

Here $r_i(k, j)$ denotes the ith element of the row $r(k, j)$. The number of invariants in (3.1) is precisely $d(n_1, \ldots, n_p)$: see (2.11).

Most canonical forms, whether in SS, MFD or ARMA form, can be derived from these $d(n_1, \ldots, n_p)$ invariants; they will most often have exactly $d(n_1, \ldots, n_p)$ parameters: see for example Rissanen (1974), Guidorzi (1981), Gevers and Wertz (1986).

The most commonly used canonical MFD is the Guidorzi form (Guidorzi 1975, 1981), also called the echelon form in time series analysis (see e.g. Deistler 1986). Let n_1, $\ldots, n_p$ be the ordered left Kronecker indices of $K(z)$. Then $K(z)$ is uniquely described by $P^{-1}(z)Q(z)$, where

$$p_{ii}(z) = z^{n_i} - \alpha_{iin_i} z^{n_i-1} - \ldots - \alpha_{ii1} \quad (3.2\ a)$$

$$p_{ij}(z) = -\alpha_{ijn_{ij}} z^{n_{ij}-1} - \ldots - \alpha_{ij1} \quad \text{for } i \neq j \quad (3.2\ b)$$

$$q_{ij}(z) = \beta_{ijn_i} z^{n_i-1} + \ldots + \beta_{ij1} \quad (3.2\ c)$$

The coefficients α_{ijk} are the invariants α_{ijk} of (3.1) and (2.6). The β_{ijk} are bilinear functions of the α_{ijk} and $r_i(j, k)$ of (3.1): see Guidorzi (1981) or Gevers and Wertz (1986). We notice that this canonical form has the following properties which uniquely define its structure:

(i) The polynomials on the main diagonal of $P(z)$ are monic with

$$\deg p_{ii} = n_i \quad (3.3\ a)$$

(ii) $\deg p_{ij} \leqq \deg p_{ii}, \quad j \leqq i; \quad \deg p_{ij} < \deg p_{ii}, \quad j > i \quad (3.3\ b)$

$$\deg p_{ji} < \deg p_{ii}, \quad j \neq i \quad (3.3\ c)$$

(iii) $\deg q_{ij} < \deg p_{ii}$ and $P(z)$, $Q(z)$ are left coprime. (3.3 d)

With the notation of (1.5), it is clear from (3.3 *b, c*) that

(i) $P_{hc} = I_p$ (3.4 *a*)

(ii) P_{hr} is lower triangular with unit diagonal elements (3.4 *b*)

Hence $P(z)$ is both row-reduced and column-reduced. In an identification context, once the left Kronecker indices n_i have been estimated, the structure of $P(z)$ and $Q(z)$ is completely specified by either (3.2) or (3.3). The α_{ijk} and β_{ijk} are parameters to be estimated; their number is exactly $d(n_1, \ldots, n_p)$ as given by (2.11).

Example 3.1

Consider a 2×2 matrix $K(z)$ of McMillan degree 3 with left Kronecker indices (2, 1). Then $d(2, 1) = 12$, $n_{12} = 1$, $n_{21} = 2$ and

$$P(z) = \begin{bmatrix} z^2 - \alpha_{112}z - \alpha_{111} & -\alpha_{121} \\ -\alpha_{212}z - \alpha_{211} & z - \alpha_{221} \end{bmatrix}, \quad Q(z) = \begin{bmatrix} \beta_{112}z + \beta_{111} & \beta_{122}z + \beta_{121} \\ \beta_{211} & \beta_{221} \end{bmatrix} \tag{3.5}$$

Note that

$$P_{hr} = \begin{bmatrix} 1 & 0 \\ -\alpha_{212} & 1 \end{bmatrix} \tag{3.6}$$

Comment 3.1

In the control engineering literature another uniquely defined MFD is called the canonical echelon MFD (see Forney (1975), Kailath (1980)). It is obtained from the Guidorzi canonical form by a permutation of the rows of $P(z)$ (and correspondingly of $Q(z)$: see the proof of Lemma 2.3) such that in the transformed $\bar{P}(z)$:

(i) the row degrees are arranged in increasing order;

(ii) if in $P(z)$ $n_i = n_j$ with $i < j$, then the ith row of $P(z)$ is above the jth row of $P(z)$ in $\bar{P}(z)$.

Finally, with $P(z)$ and $Q(z)$ defined by (1.3), we notice that the relations (2.6) can be rewritten as follows:

$$[P_0 \quad P_1 \quad \cdots \quad P_u]\mathscr{H}_{1,\infty}[K] = 0 \tag{3.7}$$

This yields (3.2 *a, b*) with $u = \rho \triangleq \max \{n_i\}$. The form of $Q(z)$ in (3.2 *c*) follows from the requirement that $P^{-1}(z)Q(z)$ must be strictly proper. It follows that $v = \rho - 1$.

We now turn to canonical ARMA forms. Although there are other ways (see Comment 1.1), the most obvious way to construct a canonical ARMA form is first to construct a row-reduced canonical MFD $P^{-1}(z)Q(z)$ with $P(z)$ having row degrees $n_1, \ldots, n_p$, and to apply the transformation (1.6 *a*). If we apply this transformation to the Guidorzi canonical MFD, we get $\bar{K}(D) = A^{-1}(D)B(D) = K_1 D + K_2 D^2 + \ldots +$ with $A(D)$ and $B(D)$ as follows:

$$a_{ii}(D) = 1 - \alpha_{iin_i}D - \ldots - \alpha_{ii1}D^{n_i} \tag{3.8 a}$$

$$a_{ij}(D) = -\alpha_{ijn_{ij}}D^{n_i - n_{ij} + 1} - \ldots - \alpha_{ij1}D^{n_i} \quad \text{for } i \neq j \tag{3.8 b}$$

$$b_{ij}(D) = \beta_{ijn_i}D + \ldots + \beta_{ij1}D^{n_i} \tag{3.8 c}$$

We can also write

$$A(D) = A_0 + A_1 D + \dots + A_\rho D^\rho, \quad B(D) = B_1 D + \dots + B_\rho D^\rho \tag{3.9}$$

where $A_0 = P_{hr}$ (see (3.4 *a*)). Note that A_ρ or B_ρ could be zero, but not both. The left Kronecker indices of $K(z)$ completely determine the structure of $A(D)$ and $B(D)$ through (3.8), but notice that, unlike the canonical MFD, the row degrees and column degrees of $A(D)$ are not necessarily equal to $(n_1, \dots, n_p)$. Indeed some of the α_{ijk} can be zero: think of a moving-average model, for example.

The following properties hold:

(i) $\partial r_i[A(D) \vdots B(D)] = n_i$ (3.10 *a*)

(ii) $\partial c_i[A(D)] \leq \rho, \quad \partial c_i[B(D)] \leq \rho$ (3.10 *b*)

(iii) $A(D)$ and $B(D)$ are left coprime (3.10 *c*)

This canonical ARMA form is sometimes called the reversed echelon form (see Deistler, 1986).

Example 3.2

For the matrix $K(z)$ of Example 3.1 we get

$$A(D) = \begin{bmatrix} 1 & 0 \\ -\alpha_{212} & 1 \end{bmatrix} + \begin{bmatrix} -\alpha_{112} & 0 \\ -\alpha_{211} & -\alpha_{211} \end{bmatrix} D + \begin{bmatrix} -\alpha_{111} & -\alpha_{121} \\ 0 & 0 \end{bmatrix} D^2 \tag{3.11 a}$$

$$B(D) = \begin{bmatrix} \beta_{112} & \beta_{122} \\ \beta_{211} & \beta_{221} \end{bmatrix} D + \begin{bmatrix} \beta_{111} & \beta_{121} \\ 0 & 0 \end{bmatrix} D^2 \tag{3.11 b}$$

It is immediately clear that our canonical reversed echelon form does not have the desired property that $A_0 = I$. Since $A_0 = P_{hr}$, it will be lower triangular with unit diagonal elements: see (3.4 *b*). On the other hand, the number of free parameters in $A(D)$, $B(D)$ is $d(n_1, \dots, n_p)$.

4. Canonical monic ARMA models

As we said above, it would be interesting to work with monic ARMA models, i.e. to have $A_0 = I$. We shall show here that it is in general impossible to construct canonical monic ARMA models such that the row or column degrees are specified by the left Kronecker indices of $K(z)$. In addition we shall show that, generically, canonical monic ARMA models can only represent systems whose McMillan degree is a multiple of $p = \dim y$. By generically we mean that if a monic ARMA model with fixed lag lengths is chosen, and if its parameters are randomly generated, then almost surely the McMillan degree of the corresponding system will be a multiple of p.

First notice that one obvious way to obtain a monic ARMA form from the canonical reversed echelon form is to redefine

$$\bar{A}(D) = A_0^{-1} A(D), \quad \bar{B}(D) = A_0^{-1} B(D) \tag{4.1}$$

However, this will in general make all row degrees of $\bar{A}(D)$ and $\bar{B}(D)$ equal to ρ.

Example 4.1

From Example 3.2 we get

$$\bar{A}(D)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\begin{bmatrix}-\alpha_{112} & 0\\ -\alpha_{211}-\alpha_{212}\alpha_{112} & -\alpha_{211}\end{bmatrix}D+\begin{bmatrix}-\alpha_{111} & -\alpha_{121}\\ -\alpha_{212}\alpha_{111} & -\alpha_{212}\alpha_{121}\end{bmatrix}D^2 \tag{4.1 a}$$

$$\bar{B}(D)=\begin{bmatrix}\beta_{112} & \beta_{122}\\ \beta_{211}+\alpha_{212}\beta_{112} & \beta_{221}+\alpha_{212}\beta_{122}\end{bmatrix}D+\begin{bmatrix}\beta_{111} & \beta_{121}\\ \alpha_{212}\beta_{111} & \alpha_{212}\beta_{121}\end{bmatrix}D^2 \tag{4.1 b}$$

Notice that the row degrees of $\bar{A}(D)$ and $\bar{B}(D)$ are now (2, 2). The number of 'free' parameters has increased from $d(2, 1)=12$ to 15, but of course these are not all independent. The next two lemmas show that this problem is a generic one with monic ARMA models.

Lemma 4.1

Let $K(z)$ be a strictly proper $p \times m$ rational transfer matrix with left Kronecker indices $n_1, \ldots, n_p$. Then it is in general impossible to construct a monic ARMA model $A(D)$, $B(D)$ for $K(z)$ such that the row degrees of $[A(D) \vdots B(D)]$ are $n_1, \ldots, n_p$.

Proof

Let $P^{-1}(z)Q(z)$ be the Guidorzi canonical form for $K(z)$. By Lemma 2.4 we know that it is always possible to arrange the rows of $K(z)$ such that the n_i are in decreasing order, but that it is not always possible to arrange them so that the n_i are increasing. Therefore, the generic situation is that $n_1 \geqq \ldots \geqq n_p$. Now suppose that for at least one $i \varepsilon \{1, \ldots, p-1\}: n_i > n_{i+1}$. Then by (2.7) $n_{i+1,i}=n_{i+1}+1$. Hence in the Guidorzi canonical form the $(i+1, i)$ element has degree n_{i+1}, which introduces a non-zero element in that same position in P_{hr}. Since the monic diagonal element in the ith column of $P(z)$ has degree $n_i > n_{i+1}$, it is impossible to remove that $(i+1, i)$ element of $P(z)$ by elementary row operations. It is therefore impossible to find a unimodular $U(z)$ such that $U(z)P(z)=R(z)$ with $R_{hr}=I$. Now any monic ARMA model $A(D)$, $B(D)$ with $\partial r_i[A(D) \vdots B(D)]=n_i$ is equivalent with a MFD $P(z)$, $Q(z)$ such that $\partial r_i[P(z)]=n_i$ and $P_{hr}=I$. Therefore it is in general impossible to construct a monic ARMA model $A(D)$, $B(D)$ with row degrees of $[A(D) \vdots B(D)]$ equal to $n_1, \ldots, n_p$.

Comment 4.1

The proof relies on the assumption that there exists at least one i such that $n_i > n_{i+1}$. In fact the same proof goes through if there exists i, j with $i<j$ such that $n_i > n_j$. The Guidorzi canonical form will have $P_{hr}=I$ if any only if $n_1 \leqq \ldots \leqq n_p$. However, such situation is unlikely: see Lemma 2.4. In fact, generically, a system will have its left Kronecker indices such that $n_1=\ldots=n_k=n_{k+1}+1=\ldots=n_p+1$ for some k, so that the only generic system that can be modelled by a monic ARMA model with row degrees of $[A(D) \vdots B(D)]$ equal to the left Kronecker indices is when $n_1=n_2=\ldots=n_p$. But this is the case only when the McMillan degree is a multiple of p, the number of outputs of the system. We state this as a corollary.

Corollary 4.1

Let $K(z)$ be a strictly proper $p \times m$ rational transfer matrix with left Kronecker indices $n_1, \ldots, n_p$ and let $K(z)$ be generic, i.e. $n_1 \geqq \ldots \geqq n_p$. Then $K(z)$ can be represented by a monic ARMA model $A(D)$, $B(D)$ wirh $\partial r_i[A(D) \vdots B(D)] = n_i$, $i = 1, \ldots, p$, only if $n_1 = \ldots = n_p$, which requires in particular that $n = p \times k$ for some integer k.

In Bokor and Keviczky (1982) it was assumed that the rows of $K(z)$ could be permuted so that the ordered Kronecker indices were increasing ($n_1 \leqq \ldots \leqq n_p$). This would have led to a Guidorzi canonical form with $P_{hr} = I$ and hence a monic canonical ARMA form with minimal row degrees and therefore a minimal number of free parameters. However, we have seen that generically this cannot be done.

A consequence of Corollary 4.1 is that, except when $n_1 = \ldots = n_p$, any monic ARMA model will have more than the minimal number of free parameters, i.e. more than $d(n_1, \ldots, n_p)$, as Example 4.1 illustrates.

Alternative characterizations of identifiable monic ARMA models have been proposed in the time series literature. One is to prescribe the column degrees of $[A(D) \vdots B(D)]$ and to impose that $A(D)$, $B(D)$ are left coprime and that the $p \times (p+m)$ matrix of coefficients of highest column degrees in each column of $[A(D) \vdots B(D)]$ be full rank (see e.g. Deistler (1985), Hannan and Kavalieris (1984)). The disadvantage is that $p+m$ integers have to be prescribed rather than p. Another identifiable parametrization is to prescribe only the highest degrees q and r in $A(D)$ and $B(D)$ respectively, and to impose that $A(D)$ and $B(D)$ be left coprime and that the matrix $[A_q \vdots B_r]$ have full rank. The matrices A_i, $1 \leqq i \leqq q$ and B_i, $1 \leqq i \leqq r$ are then fully parametrized (see e.g. Hannan (1976), Stoica (1982)). The advantage of these fully parametrized forms is that only two integers need to be specified, but a disadvantage is that the rank condition automatically implies that the McMillan degree of $K(z)$ is $p \times \max(q, r)$. This will be shown by Theorem 5.1. The implication is that only systems whose McMillan degree is a multiple of p can be represented with this parametrization. Finally, another identifiable canonical ARMA model can be derived from a basis of the constructibility matrix: see Comment 1.1. However, this parametrization is limited to systems having no poles at the origin; in particular, this excludes moving-average models.

5. The McMillan degree of ARMA models

Consider a monic ARMA (q, r) model (1.1 a, b, d) and assume that $A(D)$ and $B(D)$ are left coprime. We now express the McMillan degree $\delta[K(z)]$ of $K(z) = A^{-1}(D)B(D)$ as a function of the A_i, B_i.

Theorem 5.1

Let $A^{-1}(D)B(D)$ be a monic left coprime ARMA model for $K(z)$ with $A(D)$ and $B(D)$ as in (1.1 b, d), and let $u \triangleq \max(q, r)$. Then

$$\delta[K(z)] = \text{rank } M_u \tag{5.1}$$

where

$$\underset{pu\times(p+m)u}{M_u} \triangleq \begin{bmatrix} A_u & B_u & 0 & 0 & \cdots & 0 & 0 \\ A_{u-1} & B_{u-1} & A_u & B_u & \ddots & \vdots & \vdots \\ \vdots & \ddots & & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & & \ddots & 0 & 0 \\ A_1 & B_1 & \cdots & A_{u-1} & B_{u-1} & A_u & B_u \end{bmatrix} \tag{5.2}$$

Proof

Since $A(D)$, $B(D)$ is left coprime, we have

$$\text{rank } [A(D) \vdots B(D)] = p \quad \text{for all } D \tag{5.3}$$

Now define

$$[P_1(z) \vdots Q_1(z)] = z^u [A(D) \vdots B(D)] \tag{5.4}$$

Then by (5.3)

$$\text{rank } [P_1(z) \vdots Q_1(z)] = p \quad \text{for all } z \neq 0 \tag{5.5}$$

If the rank condition (5.5) held for all z, including $z=0$, then $P_1(z)$, $Q_1(z)$ would be left coprime and since $P_1(z)$ is row-reduced with row degrees all equal to u, the McMillan degree of $P_1^{-1}(z)Q_1(z) = A^{-1}(D)B(D)$ would be pu. To compute the actual McMillan degree, we thus have to substract from pu the number of common zeros at $z=0$ introduced by (5.4). We shall need the following notation:

$$\underset{pi\times(p+m)i}{M_i} \triangleq \begin{bmatrix} A_u & B_u & 0 & 0 & \cdots & 0 & 0 \\ A_{u-1} & B_{u-1} & A_u & B_u & \ddots & \vdots & \vdots \\ \vdots & \ddots & & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & & \ddots & 0 & 0 \\ A_{u-i+1} & B_{u-i+1} & \cdots & A_{u-1} & B_{u-1} & A_u & B_u \end{bmatrix}, \quad i = 1, \ldots, u \tag{5.6}$$

We shall call 'defect of M_i', denoted def M_i, the following quantity:

$$\text{def } M_i = pi - \text{rank } M_i \tag{5.7}$$

When M_i is full rank, def $M_i = 0$. Now consider

$$[P_1(z) \vdots Q_1(z)]_{z=0} = [A_u \vdots B_u] = M_1 \tag{5.8}$$

If rank $M_1 = p$, then $P_1(z)$ and $Q_1(z)$ are left coprime, and the McMillan degree of $K(z)$ is pu as stated above. Suppose def $M_1 = t_1$. Then there exists a $p \times p$ real non-singular matrix T_1 such that

$$T_1[P_1(z) \vdots Q_1(z)]_{z=0} = \begin{bmatrix} 0 & 0 \\ \bar{A}_u & \bar{B}_u \end{bmatrix} \begin{matrix} \}t_1 \\ \}p-t_1 \end{matrix} = T_1[A_u \vdots B_u] \tag{5.9}$$

This shows that $P_1(z)$ and $Q_1(z)$ have at least t_1 common zeroes at $z=0$. We extract these common zeros by defining

$$[P_2(z) \vdots Q_2(z)] = \text{diag}\,\{\underbrace{z^{-1}, \ldots, z^{-1}}_{t_1}, \underbrace{1, \ldots, 1}_{p-t_1}\}T_1[P_1(z) \vdots Q_1(z)] \tag{5.10}$$

Notice that $[P_2(z) \vdots Q_2(z)]$ is polynomial with $P_2^{-1}(z)Q_2(z) = A^{-1}(D)B(D)$, it has full rank for all $z \neq 0$ and $P_{2_{hr}} = I$. We now repeat the same procedure with $[P_2(z) \vdots Q_2(z)]$:

$$[P_2(z) \vdots Q_2(z)]_{z=0} = \left| \begin{matrix} \bar{A}_{u-1} & \bar{B}_{u-1} \\ \bar{A}_u & \bar{B}_u \end{matrix} \right| \begin{matrix} \}t_1 \\ \}p-t_1 \end{matrix} \triangleq S_2 \tag{5.11}$$

where $[\bar{A}_{u-1} \vdots \bar{B}_{u-1}]$ are the first t_i rows of $T_1[A_{u-1} \vdots B_{u-1}]$ and $[\bar{A}_u \vdots \bar{B}_u]$ are the last $p-t_1$ rows of $T_1[A_u \vdots B_u]$ (see (5.9)). If rank $S_2 = p$, then $P_2(z)$, $Q_2(z)$ are left coprime, and since $P_2(z)$ is row-reduced with its first t_1 row degrees equal to $u-1$ and its last $p-t_1$ row degrees equal to u, the McMillan degree of $K(z)$ is then $pu - t_1$. Denote def $S_2 \triangleq p - \text{rank } S_2$ and let def $S_2 = t_2$. Note that $t_2 \leqq t_1$ since the last $p-t_1$ rows of S_2 are linearly independent. Now

$$t_1 + t_2 = \text{def } M_2 \tag{5.12}$$

This follows from the fact that

$$\begin{bmatrix} T_1 & 0 \\ 0 & T_1 \end{bmatrix} \left[\begin{array}{cc|cc} A_u & B_u & 0 & 0 \\ \hline A_{u-1} & B_{u-1} & A_u & B_u \end{array} \right] = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ \bar{A}_u & \bar{B}_u & 0 & 0 \\ \hline \bar{A}_{u-1} & \bar{B}_{u-1} & 0 & 0 \\ \tilde{A}_{u-1} & \tilde{B}_{u-1} & \bar{A}_u & \bar{B}_u \end{array} \right] \tag{5.13}$$

From (5.13) it is clear that

$$\text{rank } M_2 \triangleq 2p - \text{def } M_2 = \text{rank } M_1 + \text{rank } S_2 \tag{5.14}$$

We can now apply a left non-singular transformation T_2 to $[P_2(z) \vdots Q_2(z)]$ to transform the top t_2 rows of S_2 to zeros, and then eliminate these t_2 common zeros in $P_2(z)$ and $Q_2(z)$ by defining

$$[P_3(z) \vdots Q_3(z)] = \text{diag}\,\{\underbrace{z^{-1}, \ldots, z^{-1}}_{t_2}, \underbrace{1, \ldots, 1}_{p-t_2}\}T_2[P_2(z) \vdots Q_2(z)] \tag{5.15}$$

Defining $[P_3(z) \vdots Q_3(z)]_{z=0} \triangleq S_3$, and letting def $S_3 = t_3$, we notice that necessarily $t_3 \leqq t_2 \leqq t_1$ and that $t_1 + t_2 + t_3 = \text{def } M_3$. Iterating this procedure, we find that $t_1 + t_2 + \ldots + t_u = \text{def } M_u =$ number of common zeros at $z = 0$ between $P_1(z)$ and $Q_1(z)$. It follows by (5.5) that the McMillan degree of $K(z)$ is $pu - \text{def } M_u = \text{rank } M_u$.

Comment 5.1

It is important to notice that the result of Theorem 5.1 holds only if $A(D)$ and $B(D)$ are coprime. This can be easily seen as follows. Let $A(D)$ and $B(D)$ be left coprime with $u = \max(q, r)$ so that $\delta[K(z)] = \text{rank } M_u$. Now define

$$\bar{A}(D) = (I + PD)A(D), \quad \bar{B}(D) = (I + PD)B(D) \tag{5.16}$$

for some $p \times p$ non-singular constant matrix P. Then for $\bar{A}(D)$ and $\bar{B}(D)$ we have

$$M_{u+1} \triangleq \begin{bmatrix} \bar{A}_{u+1} & \bar{B}_{u+1} & 0 & 0 & \cdots & 0 & 0 \\ \bar{A}_u & \bar{B}_u & \bar{A}_{u+1} & \bar{B}_{u+1} & \ddots & \vdots & \vdots \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & \ddots & 0 & 0 \\ \bar{A}_1 & \bar{B}_1 & & & \cdots & \bar{A}_{u+1} & \bar{B}_{u+1} \end{bmatrix}$$

Now it is easy to see that

$$M_{u+1} = \begin{bmatrix} P & 0 & \cdots & \cdots & \cdots & 0 \\ I & P & \ddots & & & \vdots \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & I & P \end{bmatrix} \bar{M}_u \tag{5.17}$$

where

$$\bar{M}_u = \left[\begin{array}{cccc|cc} & & & & 0 & 0 \\ & M_u & & & \vdots & \vdots \\ & & & & 0 & 0 \\ \hline I & 0 & \cdots & 0 & A_u & B_u \end{array}\right] \tag{5.18}$$

It follows that

$$\text{rank } M_{u+1} = \text{rank } \bar{M}_u = \text{rank } M_u + \text{rank } [A_u \quad B_u] \neq \text{rank } M_u$$

Hence $\delta[K(z)] \neq \text{rank } M_{u+1}$.

Relationships between the McMillan degree of $K(z)$ and the rank of generalized Bezoutian and Sylvester matrices have been obtained in Anderson and Jury (1976). These relationships do not require that $A(D)$ and $B(D)$ be left coprime as is required in Theorem 5.1. However, the Bezoutian matrix contains the coefficient matrices of both left and right factorizations of $K(z)$ whereas only a left factorization is required here. As for the test based on Sylvester matrices, it requires the computation of the ranks of matrices of much higher dimension and rank than M_u. Our condition is a much simpler one, at the expense of a coprimeness requirement on $A(D)$, $B(D)$.

Corollary 5.1

Consider the AR model

$$y(t) + A_1 y(t-1) + \dots + A_p y(t-p) = u(t) \tag{5.19}$$

Then the McMillan degree of the system represented by (5.19) is given by

$$\text{rank} \begin{bmatrix} A_0 & 0 & \cdots & 0 \\ A_{p-1} & A_p & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A_1 & \cdots & A_{p-1} & A_p \end{bmatrix} \tag{5.20}$$

Although Corollary 5.1 is a special case of Theorem 5.1, a much simpler proof can be given in this case. The system (5.19) can be realized in state-space form as

$$x(t+1)=\begin{bmatrix} -A_1 & I_s & 0_s & \dots & 0_s \\ -A_2 & 0_s & \ddots & & \vdots \\ \vdots & & & \ddots & 0_s \\ \vdots & & & & I_s \\ -A_p & 0_s & \dots & \dots & 0_s \end{bmatrix} x(t)+\begin{bmatrix} -A_1 \\ \vdots \\ \vdots \\ \vdots \\ -A_p \end{bmatrix} u(t) \tag{5.21 a}$$

$$y(t)=[I_s \quad 0_s \quad \dots \quad 0_s]x(t)+u(t) \tag{5.21 b}$$

It is trivial to see that the observability matrix has full rank. As for the controllability matrix, it can be written as

$$\mathscr{C}=\begin{bmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ \vdots & & & & 0 \\ A_{p-1} & & & & \vdots \\ A_p & 0 & \dots & \dots & 0 \end{bmatrix}\begin{bmatrix} -I_s & -A_1 & -A_2 & \dots & -A_{p-1} \\ 0 & -I_s & -A_1 & & \vdots \\ \vdots & & \ddots & \ddots & -A_1 \\ 0 & \dots & 0 & & -I_s \end{bmatrix}^{-1} \tag{5.22}$$

The result follows immediately using the Sylvester inequality.

It is easy to modify Theorem 5.1 to apply to ARMA models with $A_0 \neq I$ or $B_0 \neq 0$, or to MA models. For example, for the MA(q) model $y(t)=u(t)+\sum_1^q A_i u(t-i)$, the McMillan degree is given by (5.20). This last result is of course well known, since (5.20) is then the Hankel matrix of the impulse response, written upside down.

Assuming for simplicity that $m \geq p$, then Theorem 5.1 shows that if monic ARMA (q, r) models are used in system identification, one will almost always estimate models of McMillan degree pu with $u = \max(q, r)$ since M_u will almost always be full-rank when estimated A_i and B_i are used. It was shown in Dunsmuir and Hannan (1976) that the set of monic ARMA (q, r) models satisfying the full rank condition on $[A_q \vdots B_r]$ form an analytic manifold of dimension $p(pq+mr)$. Using arguments from Hannan (1976), Stoica argued that systems that cannot be described by such manifolds are 'pathological' and unlikely to be encountered in practice: see Stoica (1982). It is true that if one models a system with a fully parametrized monic ARMA (q, r) model, then any system that does not satisfy the rank condition on $[A_q \vdots B_r]$ lies in a submanifold of lower dimension. But to argue that such a system is pathological is to believe that certain McMillan degrees are more likely than others. If a system has a McMillan degree that is not a multiple of p, and this fully parametrized form is used, then asymptotically $[A_q \vdots B_r]$ will have less than full rank. However, with finite data this matrix will have full rank because the true system lies on a thin submanifold; and therefore the order will be overestimated and the number of free parameters will be unduly large.

These points are well illustrated by Example 3.2. If $K(z)$ has McMillan degree 3 with Kronecker indices (2, 1), then the non-monic ARMA model (3.11) has structural zeros in the second row of A_2 and B_2. If a monic ARMA (2, 2) model is used, then it is clear from (4.1) that rank $[A_2 \vdots B_2]=1$. But if this fully parametrized model is used in parameter estimation, then the two rows of $[A_2 \vdots B_2]$ will not be exactly linearly dependent and a fourth-order system will be estimated.

6. Conclusions

It is commonly assumed in the control engineering community that MFDs and monic ARMA models are equivalent and can therefore be used interchangeably.

We have highlighted a number of difficulties with the use of monic ARMA models as opposed to MFDs in system identification. These models can only represent systems whose McMillan degree is a multiple of the number of outputs. In all other cases they will tend to produce estimated models of higher order than the true system. We have also produced a new rank test for the determination of the McMillan degree of a left coprime ARMA decomposition. It is clear from our result that the control over the McMillan degree of a system is much more difficult when ARMA models are used than when MFDs are used.

ACKNOWLEDGMENTS

The author would like to thank Michael Green, Ted Hannan, Manfred Deistler, Joseph Bokor and Laszlo Keviczky for their comments on an earlier version of this paper.

REFERENCES

AKAIKE, H., 1974, *Ann. Inst. Statist. Math.*, **26,** 363.
ANDERSON, B. D. O., and JURY, E. I., 1976, *I.E.E.E. Trans. autom. Control*, **21,** 551.
BOKOR, J., and KEVICZKY, L., 1982, *Int. J. Control*, **36,** 461; 1985, ARMA canonical forms obtained from constructibility invariants. Submitted for publication.
CLARK, J. M. C., 1976, 'The consistent selection of parametrizations in system identification', *JACC*, Purdue University.
DEISTLER, M., 1985, *Handbook of Statistics*, Vol. 5, edited by E. J. Hannan, P. R., Krishnaiah and M. M. Rao (Amsterdam: North-Holland) pp. 257–277; 1986, *Adv. Statist. Analysis and Statist. Computing*, **1,** 51.
DEISTLER, M., and HANNAN, E. J., 1981, *J. multivariate Analysis*, **11,** 474.
DUNSMUIR, W., and HANNAN, E. J., 1976, *Adv. appl. Probab.*, **8,** 339.
FORNEY, G. D., Jr., 1975, *SIAM J. Control*, **13,** 493.
GEVERS, M., and WERTZ, V., 1986, Techniques for the selection of identifiable parametrizations for multivariable linear systems. *Control and Dynamic Systems*, Vol. XXIV (Orlando, FL: Academic Press), to be published.
GUIDORZI, R. P., 1975, *Automatica*, **11,** 361; 1981, *Automatica*, **17,** 117.
HANNAN, E. J., 1976, *Econometrica*, **44,** 713.
HANNAN, E. J., and KAVALIERIS, L., 1984, *Adv. appl. Probab.*, **16,** 492.
HAZEWINKEL, M., and KALMAN, R. E., 1976, in *Lecture notes Econ.-Math. System Theory*, Vol. 131 (Berlin: Springer-Verlag), pp. 48–60.
KAILATH, T., 1980, *Linear Systems* (Englewood Cliffs, NJ: Prentice-Hall).
POPOV, V. M., 1972, *SIAM J. Control*, **10,** 252.
RISSANEN, J., 1974, *Automatica*, **10,** 175.
STOICA, P., 1982, *Bull. Inst. Politech. Buc., ser. Electro*, 99.
WOLOWICH, W. A., 1974, 'Linear multivariable systems', *Applied Math. Sciences*, **11** (Berlin: Springer-Verlag).
WOLOWICH, W. A., and ELLIOTT, H., 1983, *Int. J. Control*, **38,** 337.