A minimal partial realization algorithm for MIMO systems

Marc Van Barel
Adhemar Bultheel

Report TW 79, July 1986
Revised : April 1987

Department of Computer Science, K.U.Leuven

ABSTRACT

An algorithm for the solution of the minimal partial realization problem for MIMO systems is presented. It gives a solution in the form of a matrix fraction. The algorithm is a generalization of the Berlekamp-Massey algorithm for the scalar case. Our approach is mainly based on the work of Dickinson, Morf and Kailath who treated the same problem. Our main tool is the construction of a nested basis for the null spaces of block Hankel matrices. The connection with Padé approximation is also given.

The problem where some of the Markov parameters are only partially defined is also considered. We then try to give values to the unknown entries such that the appoximant is kept minimal.

The relation with the work of Graves-Morris on vector valued approximants is also commented on.

Finally we give an adaptation of the algorithm to match time moments instead of Markov parameters or to match both of these parameters.

Many examples are included to illustrate the algorithms.

1. Introduction

The minimal partial realization of matrix sequences is a problem which has received much attention in the literature. Many of the approaches to the problem are based on the use of Hankel matrices and they are most often formulated in a state space terminology. Dickinson, Morf and Kailath give a recursive procedure to solve the problem which generalizes the Berlekamp-Massey algorithm [9]. One of the purposes of this paper is to illucidate this approach and to give some generalizations. Conan [8] gives such an explanation in the scalar case, we want to do this for the matrix case too.

The main features of the Berlekamp-Massey algorithm are

a. Any new solution can be obtained as a linear combination of earlier solutions.

b. Only a subset of earlier solutions is needed in the computation of subsequent solutions.

These are exactly the features of the algorithm we shall formulate. The solutions mentioned are the columns of the denominator in a polynomial matrix fraction description of the transfer function. They can be characterized as elements in the kernels of block Hankel matrices (see also Shamir [16]) which are nested in a certain way. Therefore, a recursive description of the bases of these kernels will be our main tool. It will be shown how the resulting algorithm will simplify in the scalar case to the classical Berlekamp-Massey algorithm and how it can be interpreted as an algorithm to construct Padé approximants.

Graves-Morris [11] has given an algorithm for vector valued rational interpolation. He suggests that a MIMO (multi input-multi output) problem can be transformed into a SIMO (single input-multi output) problem to make it vector valued. It is shown in this paper that this transformation will not always give the minimal solution and that the MIMO problem is essentially different from the corresponding SIMO problem.

It will also be shown how we can solve a minimal partial realization problem if some of the Markov parameters are only partially known. We call this a minimal extension problem. Given a number of Markov parameters, the minimal partial realization problem constructs the transfer function of a system which has for its initial Markov parameters the given ones and it extends this sequence with Markov parameters such that the number of state variables is not increased. The problem considered above is an obvious extension of this one. The sequence of given Markov parameters is extended and on top of it we have to fill in the unknown elements in the Markov parameters that are only partially known so that the number of state variables is minimized. Unfortunately, our approach does not allow for a recursive procedure to solve this problem.

Usually the entries in the Hankel matrix are Markov parameters. It is however possible to match also a number of given time moments. Bistritz [5] gives a state space approach for matching both time moments and Markov parameters. The time moment problem is more difficult than the Markov parameter problem since one should take care of the extra condition that the denominator of the matrix fraction description, evaluated at $z = 0$ has to be nonsingular. In a state space terminology, this means that the state transition matrix has to be nonsingular because in that case the time moments don't exist. This can happen and then Bistritz's algorithm is not applicable.

The problem of minimal partial realization is also known to be closely related to the Padé approximation problem, at least in the scalar case [10]. The classical minimal partial realization problem treats the case where the matrix fraction description has to be strictly proper. In the scalar case this means that the numerator degree has to be one less than the denominator degree. In a Padé terminology this means that we are only interested in some diagonal of the Padé table. To give the connection with Padé approximation, we shall slightly generalize the classical problem by introducing a shift parameter. This parameter is -1 for the classical case and gives in general the difference between the denominator and numerator degree

for a scalar problem. In this way it is also possible to compute other diagonals in the Padé table.

When working on this paper, we became aware of the work of Antoulas [1] which seems in a number of aspects more powerful than our approach.

2. The minimal partial realization problem

We shall start by defining the problem of minimal partial realization. To give a connection with the Padé table later, we shall slightly generalize the classical concept to the concept of minimal partial k-realization.
We are given a power series in z^{-1} which specifies the transfer function of a linear system with m inputs and p outputs

$$T(z) = \sum_{k=1}^{\infty} M_k z^{-k} \in z^{-1}K^{p\times m}[[z^{-1}]].$$

The matrices $M_k \in K^{p\times m}$ with entries in an arbitrary field K are known as the Markov parameters of the system.
$K[z]$ is the ring of polynomials over K and $K[[z]]$ are the formal power series in z with coefficients in K.

Definition The rational matrix function $z^{-(k+1)}C(z)A(z)^{-1}$ with $k \in \mathbb{Z}$, $C(z) \in K^{p\times m}[z]$ and $A(z) \in K^{m\times m}[z]$ is called an *N-th order partial k-realization* of $T(z) \in z^{-1}K^{p\times m}[[z^{-1}]]$ iff

$$z^{-(k+1)}C(z)A(z)^{-1} = M_1z^{-1} + M_2z^{-2} + \cdots + M_{N-1}z^{-N+1} + M'_N z^{-N} + \cdots . \qquad (1)$$

I.e. the first $N-1$ Markov parameters of $T(z)$ and $z^{-(k+1)}C(z)A^{-1}(z)$ are the same. Note that for $k = -1$, we have the classical definition of partial realization. The parameter k will be called the *shift parameter*.

Definition The N-th order partial k-realization $z^{-(k+1)}C(z)A^{-1}(z)$ of $T(z)$ is *minimal* iff the degree of $\det A(z)$ is the smallest possible that can be obtained among all N-th order partial k-realizations of $T(z)$.
This means in the classical case of $k = -1$ that among all systems with a transfer function that is an N-th order partial -1 realization of $T(z)$, the minimal one will have the least possible dimension for its state space [12 p.439]. State space realizations for $C(z)A^{-1}(z)$ can be found in [12]. Note that analogous definitions are possible for the case $B^{-1}(z)D(z)$, but we restrict ourselves to the above definition.

3. Transformation of the problem

We are going to transform the problem in a more manageable form. Therefore we need some definitions.

Definition A square polynomial matrix is *unimodular* iff its determinant is a nonzero element of the field K.

Definition A polynomial matrix

$$A(z) = \left[\sum_{k=1}^{d_1} a_{1k} z^k, \sum_{k=1}^{d_2} a_{2k} z^k, \ldots, \sum_{k=1}^{d_m} a_{mk} z^k \right] \in K^{m\times m}[z]$$

is called *column reduced* iff the matrix of its highest degree coefficients is nonsingular. I.e. iff

$$\det [a_{1,d_1}, a_{2,d_2}, \ldots, a_{m,d_m}]$$

is a nonzero element of K. This is equivalent with $Q(0)$ is nonsingular if

$$Q(z) = A(z^{-1}) \operatorname{diag}(z^{d_1}, z^{d_2}, \ldots, z^{d_m}).$$

or $\{a_{k,d_k}\}_{k=1}^{m}$ are linearly independent over K, which implies that the columns of a column reduced polynomial matrix are linearly independent over $K[z]$.
In Wolovich [16] we find a theorem saying that any nonsingular polynomial matrix $A(z)$ can be transformed into a column reduced form by postmultiplying with a unimodular matrix $U(z)$. Hence, to find the solutions of the N-th order minimal partial k-realization problem, we need only find all the solutions of the form $z^{-(k+1)}C(z)A^{-1}(z)$ with $A(z)$ column reduced. Indeed, $z^{-(k+1)}(C(z)U(z))(A(z)U(z))^{-1} = z^{-(k+1)}C(z)A^{-1}(z)$, and the degree of $\det(A(z)U(z)) =$ degree $\det A(z)$ if $U(z)$ is unimodular. Thus we can find all solutions from those with $A(z)$ column reduced.

Note that for a column reduced polynomial matrix $A(z)$, $\det A(z)$ has degree $\sum_{i=1}^{m} d_i$ where d_i is the degree of the i-th column of $A(z)$. Thus to make an n-th order partial k-realization minimal, we have to take the least possible degrees for the columns of $A(z)$.
We shall now give a necessary and sufficient condition for $z^{-(k+1)}C(z)A^{-1}(z)$ to be an N-th order partial k-realization in terms of the columns of $A(z)$. Therefore we introduce another definition.

Definition A polynomial vector $a(z) \in K^m[z]$ of degree α is said to have a k-order n with respect to $T(z) \in z^{-1}K^{p\times m}[[z^{-1}]]$ iff $n > \alpha \geqslant -k-1$ and $z^{k+1}T(z)a(z) = c(z) + qz^{\alpha-n} + q'z^{\alpha-n-1} + \cdots$ with $0 \neq q \in K^p$ and $c(z)$ a polynomial vector : $c(z) \in K^p[z]$.

Theorem *$z^{-(k+1)}C(z)A^{-1}(z)$ with $A(z)$ column reduced is an N-th order partial k-realization of $T(z) \in z^{-1}K^{p\times m}[[z^{-1}]]$ iff $C(z)$ is the polynomial part of $z^{-(k+1)}T(z)A(z)$ and the columns of $A(z)$ have k-order at least $n = N-k-1$ with respect to $T(z)$.*

Proof

$\Rightarrow$ From (1) we get

$$z^{-(k+1)}C(z)A^{-1}(z) = [T(z)+z^{-N}(Q_0+Q_1z^{-1}+\cdots)]$$
$$(Q_i = T'_{N+i} - T_{N+i})$$

Denote the i th column of $A(z)$ and $C(z)$ by $a_i(z)$ and $c_i(z)$ respectively and suppose that the degree of $a_i(z)$ is α_i. Then

$$c_i(z) = z^{k+1}T(z)a_i(z) + q_i z^{-N+\alpha_i+k+1} + q_i' z^{-N+\alpha_i+k} + \cdots$$

with q_i not necessarily nonzero. Thus the k-order of $a_i(z)$ is at least $N-k-1$.

$\Leftarrow$ Because $A(z)$ is column reduced :

$$A^{-1}(z) = [b_1(z)b_2(z)\cdots b_m(z)]^T / \det A(z)$$

with $b_i(z) \in K^m[z]$ of degree $\sum_{j=1}^{m} \alpha_j - \alpha_i$ and degree $\det A(z) = \sum_{j=1}^{m} \alpha_j$. Since the columns of $A(z)$ have k-order at least $N-k-1$

$$z^{k+1}T(z)A(z) = C(z) + R(z)$$

with the i th column of $R(z)$ of the form

$$r_i(z) = r_i z^{\alpha_i-N+k+1} + r_i' z^{\alpha_i-N+k} + \cdots$$
(α_i is the degree of $a_i(z)$).

Multiplying with $z^{-(k+1)}A^{-1}(z)$ gives

$$T(z) = z^{-(k+1)}C(z)A^{-1}(z) + S(z)$$

with

$$S(z) = \sum_{j=1}^{m} r_j b_j^T(z) z^{-N+\alpha_j}/\det A(z) + \sum_{j=1}^{m} r_j' b_j^T(z) z^{-N+\alpha_j-1}/\det A(z) + \cdots$$

which is of the form $S(z) = z^{-N}(S_0+S_1z^{-1}+\cdots)$, $S_k \in K^{p\times m}$. This proves that $z^{-(k+1)}C(z)A^{-1}(z)$ is an N-th order partial k-realization of $T(z)$.

□

With these results we can reformulate our problem of minimal partial k-realization as follows : Find a column reduced polynomial matrix $A(z)$ with columns of least possible degree and of k-order at least $N-k-1$. The numerator $C(z)$ can be computed as the polynomial part of the product $z^{k+1}T(z)A(z)$.

4. Finding polynomial vectors of degree α and of k-order at least n

It is clear from the reformulation of our problem that we have to find polynomial vectors $a(z) = \sum_{i=0}^{\alpha} a_i z^i \in K^m[z]$ of degree α (i.e $a_\alpha \neq 0$) with k-order at least $n = N-k-1$.
Therefore, the coefficients of $a(z)$ have to satisfy the following linear block Hankel system based on the Markov parameters M_k (we suppose $M_i = 0$ for $i < 1$)

$$\begin{bmatrix} M_{k+2} & M_{k+3} & \cdots & M_{\alpha+k+2} \\ M_{k+3} & M_{k+4} & \cdots & \vdots \\ \vdots & \vdots & & \\ M_{n+k+1-\alpha} & & \cdots & M_{n+k+1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_\alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r \end{bmatrix}, \quad r \in K^p \text{ possibly} \neq 0. \tag{2}$$

The system (2) is referred to as the system $(n,\alpha)_k$. We also use the convention that if $a(z)$ is a polynomial then a will denote the vector of its coefficients. The vector space of all solutions of system $(n,\alpha)_k$ is denoted by $S(n,\alpha)_k$. This is the kernel or null space of the block Hankel matrix which is the coefficient matrix of the system $(n,\alpha)_k$ with the last block row deleted. If we denote the block Hankel coefficient matrix of the system $(n,\alpha)_k$ as $\mathbf{H}(n,\alpha)_k$, then the block Hankel matrix obtained by deleting the last block row in $\mathbf{H}(n,\alpha)_k$ is $\mathbf{H}(n-1,\alpha)_k$. Thus $S(n,\alpha)_k$ = the kernel of $\mathbf{H}(n-1,\alpha)_k$. r is called the *residual* of $a \in S(n,\alpha)_k$. The operator associating with $a \in S(n,\alpha)_k$ its residual is denoted by $R : Ra = r$. a_α is called the *highest degree coefficient* of a for obvious reasons. The operator associating with a its highest degree coefficient is denoted by $H : Ha = a_\alpha$. The operators R and H depend on n, k and/or α. It should be clear from the context what values for n and α are in order. The shift operator mapping $[a_0^T a_1^T \cdots a_\alpha^T]$ into $[0^T a_0^T \cdots a_\alpha^T]$ is denoted by z, again for obvious reasons.
Consider now the family of systems represented in table 1.

Notes

- The solution a of system $(l,\alpha)_k$ is a vector in $K^{m(\alpha+1)}$. I.e. $a(z) \in K^m[z]$ is a polynomial of degree at most α.
- A certain constant value of k is used in table 1 and in the following algorithm. Later on we shall see how we can compute solutions for a value $k+1$ when we know the solutions for the value k.
- The level refers to the k-order of approximation. On level n, solutions are found of k-order n.
- The arrows have to be understood as follows :
  $(l_1,\alpha_1)_k \rightarrow (l_2,\alpha_2)_k$ means that $S(l_1,\alpha_1)_k \subset S(l_2,\alpha_2)_k$: All solutions of $(l_1,\alpha_1)_k$ are also solutions of $(l_2,\alpha_2)_k$. If $\alpha_1 < \alpha_2$, then $(\alpha_2-\alpha_1)m$ zeros are supposed to be added at the bottom of the elements in $S(l_1,\alpha_1)_k$ to match the dimension of the elements in $S(l_2,\alpha_2)_k$.
  $(l_1,\alpha_1)_k \xrightarrow[z]{} (l_2,\alpha_2)_k$ means that $z\,S(l_1,\alpha_1)_k \subset S(l_2,\alpha_2)_k$.

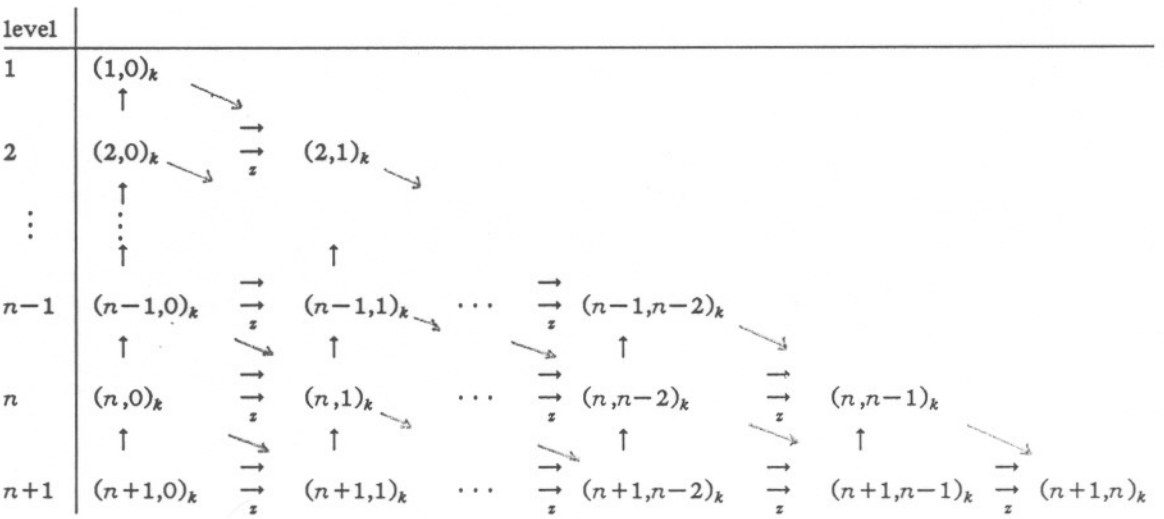


Table 1

The nesting of the solutions of the systems of table 1 form the basis for the construction of an algorithm to generate the columns (solution vectors) of a column reduced denominator $A(z)$ of the minimal partial k-realization problem of order $N = n+k+1$. Because $A(z)$ is column reduced and because of the minimality condition we find a certain structure in the set of solution vectors to be explained in the next section. Next we shall describe all possible solutions $S(n,\alpha)_k$ of an arbitrary system of table 1 by giving a basis for it. This will make it possible to construct solution vectors for order N from solution vectors for order $N-1$ and lower. For a practical implementation it will be shown next that we don't have to store the solution vectors of all previous orders, but only those of order $N-1$ and some of those of lower order (auxiliary vectors). This will finally give us an algorithm to construct the solution vectors for order N minimal partial k-realizations.

5. Solution vectors

The columns of $A(z)$, being a choice for a column reduced denominator of an N-th order minimal partial k-realization ($N = n+k+1$) are called *solution vectors*.
All solution vectors of degree α are gathered in a set, which we denote by $[n,\alpha]_k$, $\alpha = 0,1,2,\ldots,n-1$. The degree of $\det A(z)$ is $\sum_{i=1}^{m} \alpha_i$ with α_i the degree of the i th solution vector. Since we have to minimize the degree of $\det A(z)$, we should minimize each α_i. In other words, we should put as many vectors in $[n,0]_k$ as possible. Because the elements in $[n,0]_k = H[n,0]_k$ must be linearly independent, otherwise $A(z)$ could not be column reduced, we take for $[n,0]_k$ a basis for $S(n,0)_k$. Next we take for $[n,1]_k$ as many elements of $S(n,1)_k$ as possible (degree 1 solutions). Their highest degree coefficients, together with the highest degree coefficients of $[n,0]_k$ must be a linearly independent set and generate $HS(n,0)_k \cup HS(n,1)_k = HS(n,1)_k$ ($zS(n,0)_k \subset S(n,1)_k$). In other words $H[n,0]_k \cup H[n,1]_k$ is a basis for $HS(n,1)_k$. In general, $\bigcup_{\alpha'=0}^{\alpha} H[n,\alpha']_k$ will be a basis for $\bigcup_{\alpha'=0}^{\alpha} HS(n,\alpha')_k = HS(n,\alpha)_k$. This means that the number of elements in $[n,\alpha]_k$ is $\dim \operatorname{span}_K H[n,\alpha]_k = \dim HS(n,\alpha)_k - \dim HS(n,\alpha-1)_k$. Hence the number of elements in $[n,\alpha]_k$ is fixed by the Markov parameters and independent of any particular choice made by some algorithm. A connection with controllability indices will be given later. From this description we can find all possible solutions to this problem. Indeed, suppose we have a particular choice for a column reduced

$A(z)$ with columns $a_i \in [n,\alpha_i]_k$, $i = 1,2,\ldots,m$. Then all column reduced solutions for $A(z)$ are found by replacing each a_i by any other vector in $S(n,\alpha_i)_k$ on condition that the highest degree coefficients remain linearly independent.

6. A basis for $S(n,\alpha)_k$

To describe all possible solutions of $(n,\alpha)_k$ we give a basis for the space $S(n,\alpha)_k$. It will turn out that such a basis can be given in terms of the solution vectors that are found in the triangle represented in table 2.

$$
\begin{array}{llll}
[n-\alpha,0]_k & & & \\
[n+1-\alpha,0]_k & [n+1-\alpha,1]_k & & \\
\vdots & \vdots & & \\
[n,0]_k & [n,1]_k & \cdots & [n,\alpha]_k
\end{array}
$$

Table 2

Theorem *A basis for* $S(n,\alpha)_k$ *is given by*

$$B(n,\alpha)_k = \bigcup_{\alpha'=0}^{\alpha} B'(n-\alpha',\alpha-\alpha')_k \qquad (3)$$

with $B'(l,j)_k = \bigcup_{i=0}^{j} z^i [l,j-i]_k$.

Proof The relations of table 1 show that all the vectors in $[n,\alpha]_k$, $z[n,\alpha-1]_k, \ldots, z^\alpha[n,0]_k$ are in $S(n,\alpha)_k$. They are linearly independent since their highest degree coefficients are. Hence their union, which is $B'(n,\alpha)_k$ is a subset of $B(n,\alpha)_k$. The remaining basis vectors must have highest degree coefficients that are linearly dependent of the highest degree coefficients of $B'(n,\alpha)_k$, otherwise, they would be in some $[n,\alpha-i]_k$. Suppose b is such a vector. $b \in B(n,\alpha)_k \backslash B'(n,\alpha)_k$. Then $Hb \in$ span $HB'(n,\alpha)_k$ i.e.

$$Hb = \sum_{h_i \in B'(n,\alpha)_k} \mu_i \, Hh_i \ .$$

Hence we may replace b in the basis by

$$b' = b - \sum_{h_i \in B'(n,\alpha)_k} \mu_i \, h_i \ .$$

$b'(z)$ is a polynomial of degree at most $\alpha-1$ since the coefficient of z^α is eliminated. Because b is an element of $S(n,\alpha)_k$, b' must be in $S(n-1,\alpha-1)_k$.
Conversely, every solution of $(n-1,\alpha-1)_k$ is a solution of $(n,\alpha)_k$ since $S(n-1,\alpha-1)_k \subset S(n,\alpha)_k$. Thus a basis for $S(n-1,\alpha-1)_k$ will fill up the remaining basis vectors in $B(n,\alpha)_k \backslash B'(n,\alpha)_k$. This shows that $B(n,\alpha)_k = B'(n,\alpha)_k \cup B(n-1,\alpha-1)_k$. Hence, by induction

$$B(n,\alpha)_k = B'(n,\alpha)_k \cup B'(n-1,\alpha-1)_k \cup \cdots \cup B'(n+1-\alpha,1)_k \cup B(n-\alpha,0)_k \ .$$

Now $B(n-\alpha,0)_k = B'(n-\alpha,0)_k = [n-\alpha,0]_k$. This proves the theorem.

□

The basis can be rearranged as

$$B(n,\alpha)_k = \bigcup_{\pi=0}^{\alpha} z^{\alpha-\pi} P_\pi(n)_k \tag{4}$$

with

$$P_\pi(n)_k = \bigcup_{j=0}^{\pi} [n-j,\pi-j]_k .$$

The splitting (3) has the advantage that we can split up $B(n,\alpha)_k$ as

$$B(n,\alpha)_k = B(n-1,\alpha-1)_k \cup B'(n,\alpha)_k$$

where $B'(n,\alpha)_k$ contains only solution vectors of level n itself and $B(n-1,\alpha-1)_k$ contains only vectors of a degree lower than α and a level lower than n. I.e. the basis vectors in $B(n-1,\alpha-1)_k$ do not contribute to the highest degree coefficients in $S(n,\alpha)_k$. The splitting of the basis vectors in $B(n,\alpha)_k$ in subsets $z^{\alpha-\pi} P_\pi(n)_k$ has the advantage that $P_\pi(n)_k$ is independent of α. $P_\alpha(n)_k$ contains all the solution vectors that appear in the basis $B(n,\alpha)_k$ but do not appear in the basis $B(n,\alpha-1)_k$. Indeed,

$$B(n,\alpha)_k = z\, B(n,\alpha-1)_k \cup P_\alpha(n)_k.$$

7. Connection with controllability indices

As we have said before, there is a connection between the degrees of the solution vectors and the controllability indices from system theory. We shall explain this connection for the value $k = -1$ of the shift parameter. Since this shift value is fixed in this section, we shall drop it from the notation.

Consider the solution spaces $S(n+1,\alpha)$ for level $n+1$ and degree α and the solution space $S(n,\alpha-1)$ for level n and degree $\alpha-1$.

The rank information for a block Hankel matrix is given by controllability indices. Consider the block Hankel matrix with $p\times m$ blocks :

$$T = \mathrm{H}(2n-1,n-1)_{-1} = \begin{bmatrix} M_1 & M_2 & M_3 & \cdots & & M_n \\ M_2 & M_3 & M_4 & & M_n & x \\ M_2 & M_4 & M_5 & \cdots & x & x \\ \vdots & & & & & \\ M_n & x & x & \cdots & x & x \end{bmatrix} .$$

The x-entries are chosen such that the linear dependencies of the columns are not changed. Let T be of rank r and let $J_r = \{j_1, j_2, \ldots, j_r\}$ be the indices of the first r linear independent columns of T. The structure of the block Hankel matrix now guarantees that if the j-th column is linearly dependent on the previous column, then also columns $j+m, j+2m, \ldots$ are linearly dependent on the previous columns, provided the x-elements are chosen appropriately [4]. Now select from each block column in T the i-th column, i.e. the column corresponding to the i-th input. Call μ_i the number of linear independent columns among these. These μ_i are called the *controllability indices*. There is thus a one-to-one correspondence between the set of indices J_r and the set $U = \{\mu_1, \mu_2, \ldots, \mu_m\}$. Also $\sum_1^m \mu_i = r$.

Because $S(n,\alpha-1) \subset S(n+1,\alpha)$, it is clear that the number of basis vectors for level $n+1$ of degree $\leqslant \alpha$ is equal to the number of basis vectors for level n of degree $\leqslant \alpha-1$ plus the number of linear dependent columns contained in the last block column of the coefficient matrix of the system $(n+1,\alpha)$ without the last block.

Thus the number of elements in $[n+1,\alpha]$ is equal to the number of linearly dependent columns of the last block column whose corresponding column in the previous block column is linearly independent of the previous columns. This means that the degrees of the solution vectors are equal to the controllability indices.

Thus we here have that the following numbers are equal

- The degree of the minimal partial -1 realization
- The sum of the degrees of the solution vectors
- The sum of the controllability indices
- The rank of the block Hankel matrix T.

8. Generating solution vectors

We shall give a method to find a set of solution vectors for a certain level if solution vectors for all previous levels are given.
The initial step for level 1 is trivial since we can choose for level 1 any set of polynomials of degree zero that are linearly independent. Indeed, the system $(1,0)_k$ imposes no condition on the solution vectors. For simplicity we choose the m unit vectors of K^m as solution vectors for $[1,0]_k$.

Suppose that we have made a choice for all solution vectors of levels $1,2,\ldots,n$. We shall describe now how to find solution vectors of an arbitrary degree α' for level $n+1$. First notice that

$$[n+1,\alpha']_k \subset S(n+1,\alpha')_k \subset S(n,\alpha')_k \quad \text{(see table 1)}.$$

This means that if $a^{n+1} \in [n+1,\alpha']_k$ then it can be written as

$$a^{n+1} = \sum_{h \in B(n,\alpha')_k} \mu_h h. \tag{5}$$

Such an element in $S(n,\alpha')_k$ will produce a residual Ra^{n+1}. If it has to be in $S(n+1,\alpha')_k$ too, then we should choose the linear combination in (5) such that this residual is zero.

When generating all solution vectors a_i^{n+1}, $i = 1,2,\ldots,m$ of level $n+1$, two other points should be kept in mind :

1. The highest degree coefficients Ha_i^{n+1}, $i = 1,2,\ldots,m$ should be linearly independent (column reduced $A(z)$).
2. The degrees of the a_i^{n+1} should be as low as possible. I.e. we should choose $a_i^{n+1} \in [n+1,\alpha']_k$ with α' as low as possible (minimality).

9. Linear independence of highest degree coefficients

To generate solution vectors of level $n+1$ with linear independent highest degree coefficients, we shall generate them as follows : Suppose a_i^n, $i = 1,2,\ldots,m$ are the solution vectors of level n with $a_i^n \in [n,\alpha_i]_k$ with $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_m$. First generate $a_1^{n+1} \in [n+1,\alpha_1']_k$ of the form

$$a_1^{n+1} = z^{\alpha_1'-\alpha_1} a_1^n + \sum_{h \in B(n-1,\alpha_1'-1)_k} \mu_h h$$

such that its residual is zero and with $\alpha_1' \geqslant \alpha_1$ as low as possible. The second term in the right hand side does not contribute to the highest degree coefficient of a_1^{n+1}. Thus

$$Ha_1^{n+1} = Ha_1^n.$$

We call a_1^{n+1} the solution vector of level $n+1$ *based on* a_1^n or *with highest degree component* Ha_1^n or *generated by* a_1^n. Next we generate a_2^{n+1} of the form

$$a_2^{n+1} = z^{\alpha_2'-\alpha_2} a_2^n + \nu_1 z^{\alpha_2'-\alpha_1} a_1^n + \sum_{h \in B(n-1,\alpha_2'-1)_k} \mu_h h$$

such that its residual is zero and with $\alpha_2' \geqslant \alpha_2$ as low as possible. Its highest degree coefficient is

$$Ha_2^{n+1} = Ha_2^n + \nu_1 Ha_1^n.$$

This is the solution vector of level $n+1$ *based on* a_2^n or with *highest degree component* Ha_2^n. In general we shall generate the solution vector a_i^{n+1} *based on* a_i^n of the form

$$a_i^{n+1} = z^{\alpha_i' - \alpha_i} a_i^n + \sum_{j=1}^{i-1} \nu_j z^{\alpha_i' - \alpha_j} a_j^n + \sum_{h \in B(n-1, \alpha_i' - 1)_k} \mu_h h \tag{6}$$

such that its residual vanishes and such that $\alpha_i' \geqslant \alpha_i$ is as low as possible. Its highest degree coefficient is

$$Ha_i^{n+1} = Ha_i^n + \sum_{j=1}^{i-1} \nu_j \, Ha_j^n.$$

Therefore we also say that it has *highest degree component* Ha_i^n.
Because the Ha_i^n, $i = 1,2,\ldots,m$ are linearly independent, since they are solution vectors of the previous level, also the highest degree coefficients Ha_i^{n+1}, $i = 1,2,\ldots,m$ will be linearly independent.

10. Solution vectors of minimal degree (potential degree)

The basis $B(n,\alpha')_k$ could be partitioned in the sets

$$z^{\alpha' - \pi} P_\pi(n)_k, \quad \pi = 0,1,\ldots,\alpha'.$$

The solution vectors in $P_\pi(n)_k$, i.e.

$$P_\pi(n)_k = \bigcup_{j=0}^{\pi} [n-j, \pi - j]_k$$

are said to have *potential degree* π *with respect to level* n. The reason for this is that the solution vectors in $P_\pi(n)_k$ appear only in the linear combination (6) if we are looking for a solution vector a_i^{n+1} of degree $\alpha_i' \geqslant \pi$ and they don't feature in it if we are looking for a solution vector a_i^{n+1} of degree $\alpha_i' = \pi - 1$ or lower. For solution vectors of level n their degree and their potential degree are the same. The potential degree depends on the level n. It is increased by one if we go from level n to level $n+1$. In table 2 the potential degree of solution vectors on a downward sloping diagonal is constant. It starts with potential degree 0 for $P_0(n)_k = [n,0]_k$ and it goes up to a potential degree $n-1$ for the diagonal $P_{n-1}(n)_k = [1,0]_k \cup [2,1]_k \cup \cdots \cup [n,n-1]_k$. Usually it will be clear to what level we are referring and we shall speak of the potential degree of a solution vector without further specification.

To make the potential degree explicit in (6), suppose that the last sum in (6) contains r terms

$$\sum_{h \in B(n-1, \alpha_i' - 1)_k} \mu_h h = \sum_{j=1}^{t} \mu_j h_j$$

and that h_j is in $z^{\alpha_i' - \pi_j} P_{\pi_j}(n)_k$, i.e. it is of the form

$$h_j = z^{\alpha_i' - \pi_j} x_j \quad \text{with} \quad x_j \in P_{\pi_j}(n)_k, \ \alpha_i' \geqslant \pi_j.$$

Set further $x_{r+j} = a_j$ and $\pi_{r+j} = \alpha_j$ for $j = 1,2,\ldots,i-1$. Thus we can rewrite (6) as

$$a_i^{n+1} = z^{\alpha_i' - \alpha_i} a_i^n + \sum_{j=1}^{t} \mu_j z^{\alpha_i' - \pi_j} x_j \tag{7}$$

where

$$X = \{x_j\}_{j=1}^t = \bigcup_{\pi=1}^{\alpha_i' - 1} P_\pi(n-1)_k \cup \{a_j^n\}_{j=1}^{i-1}, \quad t = r+i-1$$

and with x_j of potential degree π_j (with respect to level n). Note that $x \in P_\pi(n-1)_k$ has

potential degree π with respect to level $n-1$ and potential degree $\pi+1$ with respect to level n. The coefficients μ_j are chosen such that the residual of a_i^{n+1} (as a vector in $S(n,\alpha_i')_k$) vanishes.
To keep α_i' as low as possible, i.e. $\alpha_i' = \alpha_i$, it is clear that we should first try to make the residual vanish, using only x_j that have a potential degree less than $\alpha_i = \pi_i$. This means that they are in the set

$$X_0 = \bigcup_{\pi=1}^{\pi_i-1} P_\pi(n-1)_k \cup \{a_j^n\}_{j=1}^{i-1}.$$

If we succeed in making the residual zero, then $a_i^{n+1} \in [n+1,\pi_i]_k$. If we don't, we try $\alpha_i' = \alpha_i+1 = \pi_i+1$. I.e. we enlarge the set X_0 with $P_{\pi_i}(n-1)_k$ to give

$$X_1 = X_0 \cup P_{\pi_i}(n-1)_k$$

and try again to make the residual zero. If we succeed, $a_i^{n+1} \in [n+1,\pi_i+1]_k$. If we don't we again enlarge X_1 with $P_{\pi_i+1}(n-1)_k$ etc... .

11. Making the residuals vanish (auxiliary vectors)

Our remaining problem is to make the residuals vanish. What can be said about the residual of a vector like (7) where $a_i^n \in [n,\alpha_i]_k$ and $x_j \in [n-q_j,\pi_j-q_j]_k$ with $q_j \geqslant 0$? Suppose Ra_i^n respresents the residual of a_i^n as a solution of the system $(n,\alpha_i)_k$ and let Rx_j represent the residual of x_j as a solution of the system $(n-q_j,\pi_j-q_j)_k$. The shifts $z^{\alpha_i'-\alpha_i}$ and $z^{\alpha_i'-\pi_j}$ are then such that the residual of a_i^{n+1} as a solution of the system $(n,\alpha_i')_k$ is given by

$$Ra_i^{n+1} = Ra_i^n + \sum_{j=1}^{t} \mu_j \, Rx_j . \tag{8}$$

Thus we have to find the μ_j such that (8) becomes zero. In other words, we have to write Ra_i^n as a linear combination of the $\{Rx_j\}_{j=1}^t$. The subspace of K^p, generated by the residuals of all the solution vectors that were found at a certain stage of the algorithm is called the *residual space*. It is clear that we only have to keep a set of vectors x_j $j = 1,2,\ldots,s \leqslant t$ in (7) whose residuals Rx_j form a basis for the residual space. These are called *auxiliary vectors*. In general there are many possible choices for the auxiliary vectors, but to keep α_i' in (7) as low as possible, we should choose as auxiliary vectors those solution vectors with a potential degree that is as low as possible. Suppose at a certain stage of the algorithm we are looking for a_i^{n+1} which is based on $a_i^n \in [n,\alpha_i]_k$ and that we have a set of auxiliary vectors $X = \{x_j\}_{j=1}^s$ whose residuals $RX = \{Rx_j\}_{j=1}^s$ form a basis for the current residual space which is defined to be the space generated by the residual vectors

$$RB(n-1,n-2)_k \cup \{Ra_j^n\}_{j=1}^{i-1}$$

if we are in a stage where we are looking for a solution vector based on a_i^n. Suppose further that the potential degrees (with respect to level n) of the auxiliary vectors x_j are π_j, $j = 1,2,\ldots,s$ with $\pi_1 \leqslant \pi_2 \leqslant \cdots \leqslant \pi_s$ and that these potential degrees are as low as possible.
Thus to find a_i^{n+1} as (7) with α_i' as low as possible and such that (8) vanishes we should find μ_j such that

$$Ra_i^n = -\sum_{j=1}^{r} \mu_j Rx_j \tag{9}$$

with r $(1 \leqslant r \leqslant s)$ as low as possible, i.e. $\mu_r \neq 0$. There are two possibilities. Either Ra_i^n is in the residual space span RX we had so far or it is not.

(1) If $Ra_i^n \notin$ span RX, we can not make the residual vanish. In that case we can choose any polynomial of degree n, based on a_i^n which is then in $S(n+1,n)_k$. Indeed the system

$(n+1,n)_k$ imposes no condition on the solution vector. We could take e.g. $a_i^{n+1} = z^{n-\pi_i} a_i^n$.
Before we proceed with a_{i+1}^n, we should add a_i^n of (potential) degree α_i to the set of auxiliary vectors (i.e. $s \to s+1$, $X \to X \cup \{a_i^n\}$).

(2) If $Ra_i^n \in$ span RX, we shall find coefficients $\mu_1, \mu_2, \ldots, \mu_r$ satisfying (9) with $r \leqslant s$ as low as possible. Hence a_i^{n+1} is given by

$$a_i^{n+1} = z^{\alpha_i' - \alpha_i} a_i^n + \sum_{j=1}^{r} \mu_j z^{\alpha_i' - \pi_j} x_j \in [n+1, \alpha_i']_k$$

where $\alpha_i' = \max\{\alpha_i, \pi_1, \pi_2, \ldots, \pi_r\}$. The residual space remains the same. However its basis which is the set of auxiliary vectors with lowest possible potential degrees may need some updating. There are three possibilities

(2a) $\alpha_i < \pi_r$
This means that a_i^n has a lower (potential) degree α_i than x_r which has potential degree π_r. We should then remove x_r from X and replace it by a_i^n at the appropriate place since this new set of auxiliary vectors generates the same residual space span RX, but with auxiliary vectors of lower potential degree.

(2b) $\alpha_i > \pi_r$
In this case we leave X unchanged.

(2c) $\alpha_i = \pi_r$
Now we can do as in (2a) or (2b) because the residual space and the potential degrees of the auxiliary vectors are left unaltered. This is a degree of freedom that is left by the algorithm.

12. Final algorithm

All what has been said before can be accumulated in the algorithm A that is formulated below.
It finds the denominator of a column reduced n-th order minimal k-realization of $T(z) = \sum_{i=1}^{\infty} M_i z^{-i}$, $M_i \in K^{p \times m}$ for $n = 1,2, \cdots$.

A.1 {Initialization}

A.1.1 Choose $[1,0]_k = \{e_1, \ldots, e_m\}$ the m unit vectors of K^m.

A.1.2 The set of auxiliary vectors X and the corresponding set of residuals RX is empty ($s=0$).

A.1.3 The set of solution vectors $A = [1,0]_k = \{e_1, \ldots, e_m\}$ all of (potential) degree $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$.
The corresponding set of residuals $RA = M_{k+2}$ (the m columns of M_{k+2}).

A.2 for $n = 1,2,\ldots$ {loop to find the solution vectors of level $n+1$}
Suppose we know

- $X = [x_1, \ldots, x_s]$ the set of auxiliary vectors with potential degrees $\pi_1 \leqslant \pi_2 \leqslant \cdots \leqslant \pi_s$ (with respect to level n)
- $RX = [Rx_1, \ldots, Rx_s]$ the set of corresponding residuals (= basis for the residual space)
- $A = [a_1^n, \ldots, a_m^n]$ the set of solution vectors for level n having (potential) degrees $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_m$.
- A' is the set of solution vectors for level $n+1$ which is empty at the beginning of this block.

A.2.1 **for** $i = 1,2, \ldots, m$ {loop to find the solution vector of level $n+1$ based on a_i^n}

A.2.1.1 $Ra_i^n = [M_{n+k-\alpha_i+1}, \ldots, M_{n+k+1}]a_i^n$
{i.e. compute the residual Ra_i^n of $a_i^n \in [n,\alpha_i]_k$}

A.2.1.2 Try to write Ra_i^n as a linear combination of $Rx_1, \ldots, Rx_r$:

$$Ra_i^n = -\sum_{j=1}^{r} \mu_j Rx_j$$

with r as low as possible ($\mu_r \neq 0$).
If $Ra_i^n = 0$ set $r = 0$ and $\pi_0 = 0$.

A.2.1.3 {Find a^{n+1}, the solution vector of level $n+1$ based on a_i^n and update the set X of auxiliary vectors}

A.2.1.3.1 **If** $r \leqslant s$ does not exist **then**

A.2.1.3.1.1 $a^{n+1} = z^{n-\alpha_i} a_i^n$

A.2.1.3.1.2 $\alpha' = n$

A.2.1.3.1.3 Add a_i^n of potential degree α_i to X at the correct place to keep nondecreasing potential degree ordering in X ($s \rightarrow s+1$).

A.2.1.3.1.4 Introduce the residual Ra_i^n at the corresponding plaxe in RX.

A.2.1.3.2 **If** $r \leqslant s$ does exist and $\alpha_i < \pi_r$ **then**

A.2.1.3.2.1 $a^{n+1} = z^{\pi_r - \alpha_i} a_i^n + \sum_{j=1}^{r} \mu_j z^{\pi_r - \pi_j} x_j$

A.2.1.3.2.2 $\alpha' = \pi_r$

A.2.1.3.2.3 Replace x_r in X by a_i^n of (potential) degree α_i at the correct place to keep nondecreasing potential degree ordering in X.

A.2.1.3.2.4 Replace Rx_r by the residual Ra_i^n at the corresponding position in RX.

A.2.1.3.3 **If** $r \leqslant s$ does exist and $\alpha_i \geqslant \pi_r$ **then**

A.2.1.3.3.1 $a^{n+1} = a_i^n + \sum_{j=1}^{r} \mu_j z^{\alpha_i - \pi_j} x_j$

A.2.1.3.3.2 $\alpha' = \alpha_i$

A.2.1.3.4 Add a^{n+1} of degree α' to the set A' at the correct place such that nondecreasing degree ordering in A' is maintained.

A.2.2 {All solution vectors of level $n+1$ are found. Before we go to the next level do the following}

A.2.2.1 Replace A by A'

A.2.2.2 Increase the potential degrees of the vectors in X with 1

Notes

1. The choice $[1,0]_k = [e_1, \ldots, e_m]$ is arbitrary. We could take any set of m linearly independent vectors forming a basis for K^m.
2. We have chosen to replace an auxiliary vector only when the potential degree could actually be decreased.

13. Proof of the algorithm

By construction, the solution vectors will have at each level the appropriate order of approximation and the highest degree coefficients will be linearly independent. There may be some doubt about the minimality of the solution. Therefore we prove the following theorem.

Theorem *Algorithm A finds solution vectors of least possible degree.*

Proof It is clear that all possible solutions of degree 0 are found as elements of $[n+1,0]_k$. Indeed, a basis for $S(n+1,0)_k$ is given by $B(n+1,0)_k = [n,0]_k$. Thus if $[n,0]_k$ is empty, then $[n+1,0]_k$ will be empty too. If $[n,0]_k$ contains the solution vectors $a_1^n, a_2^n, \ldots, a_r^n$, then the algorithm will always try to find linear combinations of these vectors (and of auxiliarly vectors) of lowest possible degree, which is in this case first of degree zero. Thus if there are vectors in $[n+1,0]_k$, then the algorithm will find them all.

Suppose now that all solution vectors of level $n+1$ of degree $< \alpha'$ are found by the algorithm. E.g. $a_1^{n+1}, a_2^{n+1}, \ldots, a_r^{n+1}$. They must be based on some $a_{j_1}^n, a_{j_2}^n, \ldots, a_{j_r}^n$ all solution vectors from $[n,0]_k \cup [n,1]_k \cup \cdots \cup [n,\alpha'-1]_k$. The algorithm will also find solution vectors in $[n+1,\alpha']_k$ viz. $a_{r+1}^{n+1}, \ldots, a_s^{n+1}$ based on $a_{j_{r+1}}^n, \ldots, a_{j_s}^n$, solution vectors of $[n,0]_k \cup \cdots \cup [n,\alpha']_k$. Let us suppose that there exists some a^{n+1} also of degree α', that was not found by the algorithm. Because

$$a^{n+1} \in [n+1,\alpha']_k \subset S(n,\alpha')_k$$

it can be written as

$$a^{n+1} = \sum_{h \in B(n,\alpha)_k} \mu_h h$$

only the vectors in $B'(n,\alpha')_k$ will contribute to the highest degree coefficient of a^{n+1}. I.e.

$$Ha^{n+1} = \sum_{j=1}^{t} \mu_j H a_j^n$$

where $a_1^n, a_2^n, \ldots, a_t^n$ are the vectors in $[n,0]_k \cup \cdots \cup [n,\alpha']_k$. Any μ_j that goes with one of the a_j^n that is not among $a_{j_1}^n, \ldots, a_{j_s}^n$ must be zero. Indeed if e.g. $\mu_{j_{s+1}}$ were nonzero, then a^{n+1} would have a highest degree component $Ha_{j_{s+1}}^n$ and then it would have been found by the algorithm. Hence

$$Ha^{n+1} = \sum_{l=1}^{s} \mu_{j_l} H a_{j_l}^n.$$

Because the $a_1^{n+1}, \ldots, a_s^{n+1}$ have highest degree components $Ha_{j_1}^n, \ldots, Ha_{j_s}^n$ we can write a linear combination of $\{Ha_{j_l}^n\}_{l=1}^s$ as a linear combination of the $\{Ha_l^{n+1}\}_{l=1}^s$.
Hence

$$Ha^{n+1} = \sum_{l=1}^{s} \nu_l H a_l^{n+1}.$$

But this means that the highest degree coefficient of a^{n+1} is linearly dependent on the highest degree coefficients of previous solution vectors. This is impossible if $A(z)$ has to be column reduced.

□

14. Examples

Our notation for the examples is the following : By

$$M_n = \begin{matrix}[\ x_1(z) & \cdots & x_s(z) & | & a_1^n(z) & \cdots & a_m^n(z)\ ], \\ & (\alpha_1') & \cdots & (\alpha_s') & (\alpha_1) & \cdots & (\alpha_m)\end{matrix}$$

$$RM_n = [Rx_1 \cdots Rx_s \mid Ra_1^n \cdots Ra_m^n] \rightarrow A_{n+1} = [a_1^{n+1}(z) \cdots a_m^{n+1}(z)],$$

we mean that $x_1(z), \ldots, x_s(z)$ are auxiliary vectors for level n and that $a_1^n(z), \ldots, a_m^n(z)$ are the solution vectors for that level (they make up $A_n(z)$). RM_n contains the corresponding residuals. The numbers (α_i') and (α_i) indicate the potential degrees (with respect to level n) of the vectors immediately above them. The arrow indicates that $a_i^{n+1}(z)$ is the solution vector of level $n+1$ based on a_i^n. These will again appear in the second part of M_{n+1} possibly after reordering.

Example 1

First we shall give an example to show that we should reorder the vectors as described in steps A.2.1.3.1.3 and A.2.1.3.2.3 of the algorithm (i.e. keep the potential degree of the auxiliary vectors as low as possible).
Consider the example

$$T(z) = \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} z^{-1} + \begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix} z^{-2} + \begin{bmatrix}2 & 3\\ 0 & 0\end{bmatrix} z^{-3} + \cdots .$$

The initialization of algorithm A gives for $k = -1$:

$$M_1 = \begin{matrix}\left[\begin{array}{c|cc} & 1 & 0\\ & 0 & 1\end{array}\right] \\ (0)\ \ (0)\end{matrix}, \quad RM_1 = \left[\begin{array}{c|cc} & 1 & 1\\ & 0 & 0\end{array}\right] \rightarrow A_2 = \begin{bmatrix}-1 & z\\ 1 & 0\end{bmatrix}$$

Performing step A.2 for $n=1$ gives

$$M_2 = \begin{matrix}\left[\begin{array}{cc|cc} 1 & & -1 & z\\ 0 & & 1 & 0\end{array}\right] \\ (1)\quad (0)\ (1)\end{matrix}, \quad RM_2 = \left[\begin{array}{cc|cc} 1 & & 0 & 1\\ 0 & & 0 & 0\end{array}\right] \rightarrow A_2 = \begin{bmatrix}-1 & z-1\\ 1 & 0\end{bmatrix}.$$

If we repeat this step for $n=2$ and disregard the reordering of steps A.2.1.3.1.3 and A.2.1.3.2.3, we would obtain

$$M_3' = \begin{matrix}\left[\begin{array}{cc|cc} 1 & & -1 & z-1\\ 0 & & 1 & 0\end{array}\right] \\ (2)\quad (0)\ (1)\end{matrix}, \quad RM_3' = \left[\begin{array}{cc|cc} 1 & & 1 & 1\\ 0 & & 0 & 0\end{array}\right].$$

In the next step we would obtain $[-1-z^2 \ \ z^2]^T$ for the solution vector based on $[-1 \ \ 1]^T$ and $[-1-z+z^2 \ \ o]^T$ for the solution vector based on $[z-1 \ \ 0]^T$, which gives

$$A_4' = \begin{bmatrix}-z^2-1 & -1-z+z^2\\ z^2 & 0\end{bmatrix}.$$

However, if we do reorder the columns as prescribed by steps A.2.1.3.1.3 and A.2.1.3.2.3 of the algorithm we get

$$M_3 = \begin{matrix}\left[\begin{array}{cc|cc} -1 & & -1 & z-1\\ 1 & & 1 & 0\end{array}\right] \\ (0)\quad (0)\ (1)\end{matrix}, \quad RM_3 = \left[\begin{array}{cc|cc} 1 & & 1 & 1\\ 0 & & 0 & 0\end{array}\right]$$

i.e. we replace the auxiliary vector $[1 \ \ 0]^T$ of potential degree 2 by the vector $[-1 \ \ 1]^T$ which

has a lower potential degree and which generates the same residual space. In this case, the solution vector based on $[-1\ \ 1]^T$ would be as before, but the solution vector based on $[z-1\ \ 1]^T$ is now $[2z-1\ \ -z]^T$ which is of degree 1 instead of 2. So

$$A_4 = \begin{bmatrix} -z^2-1 & 2z-1 \\ z^2 & -z \end{bmatrix}$$

which is indeed minimal while A'_4 was not.

Example 2

As a next example we take the one given by Dickinson, Morf and Kailath [9].

$$T(z) = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} z^{-2} + \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} z^{-3} + \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} z^{-4} + \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix} z^{-5} + \begin{bmatrix} 2 & 7 \\ 1 & 2 \end{bmatrix} z^{-6} + \cdots .$$

The initialization gives again for $k = -1$.

$$M_1 = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right], \quad RM_1 = \left[\begin{array}{cc} -1 & 1 \\ 0 & 0 \end{array}\right] \rightarrow A_2 = \begin{bmatrix} z & 1 \\ 0 & 1 \end{bmatrix}$$
$$(0)\ (0)$$

and the next steps give

$$M_2 = \left[\begin{array}{cc|cc} 1 & & 1 & z \\ 0 & & 1 & 0 \end{array}\right], \quad RM_2 = \left[\begin{array}{cc|cc} -1 & & 1 & 0 \\ 0 & & 0 & 0 \end{array}\right] \rightarrow A_3 = \begin{bmatrix} z+1 & z \\ z & 0 \end{bmatrix}$$
$$(1)\quad (0)\ (1)$$

$$M_3 = \left[\begin{array}{cc|cc} 1 & & z+1 & z \\ 1 & & z & 0 \end{array}\right], \quad RM_3 = \left[\begin{array}{cc|cc} 1 & & 2 & 1 \\ 0 & & 0 & -1 \end{array}\right] \rightarrow A_4 = \begin{bmatrix} z-1 & z^3 \\ z-2 & 0 \end{bmatrix}.$$
$$(1)\quad (1)\ (1)$$

Note that the auxiliary vector $[1\ \ 0]^T$ of potential degree 1 is replaced by $[1\ \ 1]^T$ which is of potential degree 0 and generates the same residual space.

$$M_4 = \left[\begin{array}{cc|cc} 1 & z & z-1 & z^3 \\ 1 & 0 & z-2 & 0 \end{array}\right]$$
$$(2)\ (2)\quad (1)\ (3)$$

$$RM_4 = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 \end{array}\right] \rightarrow A_5 = \begin{bmatrix} z-1 & z^3-z \\ z-2 & -z \end{bmatrix}.$$

Note that X is now two-dimensional because $Ra_2 = [1\ \ -1]^T$ is not in the space generated by $Rx_1 = [1\ \ 0]^T$. The next steps give:

$$M_5 = \left[\begin{array}{cc|cc} 1 & z & z-1 & z^3-z \\ 1 & 0 & z-2 & -z \end{array}\right],$$
$$(3)\ (3)\quad (1)\ \ (3)$$

$$RM_5 = \left[\begin{array}{cc|cc} 1 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \end{array}\right] \rightarrow A_6 = \begin{bmatrix} z-1 & z^3 \\ z-2 & -z \end{bmatrix}$$

$$M_6 = \left[\begin{array}{cc|cc} 1 & z & z-1 & z^3 \\ 1 & 0 & z-2 & -z \end{array}\right].$$
$$(4)\ (4)\quad (1)\ (3)$$

$$RM_6 = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array}\right] \rightarrow A_7 = \begin{bmatrix} z-1 & z^3 \\ z-2 & -z \end{bmatrix}$$

etc... .

We can order the solution vectors in table 3 giving the sets $[n,\alpha]_{-1}$ for this example. All missing entries are empty.

n \ α	0	1	2	3
1	$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$			
2	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}z\\0\end{bmatrix}$		
3		$\begin{bmatrix}z\\0\end{bmatrix}$ $\begin{bmatrix}z+1\\z\end{bmatrix}$		
4		$\begin{bmatrix}z-1\\z-2\end{bmatrix}$		$\begin{bmatrix}z^3\\0\end{bmatrix}$
5		$\begin{bmatrix}z-1\\z-2\end{bmatrix}$		$\begin{bmatrix}z^3-z\\-z\end{bmatrix}$
6		$\begin{bmatrix}z-1\\z-2\end{bmatrix}$		$\begin{bmatrix}z^3\\-z\end{bmatrix}$
7		$\begin{bmatrix}z-1\\z-2\end{bmatrix}$		$\begin{bmatrix}z^3\\-z\end{bmatrix}$

Table 3: Table of $[n,\alpha]_k$.

The solutions given by Dickinson-Morf-Kailath [9] are different. They find :

$$A_1' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ , \ A_2' = \begin{bmatrix} -z & 1 \\ 0 & 1 \end{bmatrix} \ , \ A_3' = \begin{bmatrix} -z & z+1 \\ 0 & z \end{bmatrix} \ , \ A_4' = \begin{bmatrix} -z^3-z^2 & z-1 \\ -z^2 & z-2 \end{bmatrix}$$

$$A_5' = \begin{bmatrix} -z^3-z^2+3z & z-1 \\ -z^2+3z & z-2 \end{bmatrix} \ , \ A_6' = \begin{bmatrix} -z^3-z^2+z & z-1 \\ -z^2+3z & z-2 \end{bmatrix} \ , \ A_7' = \begin{bmatrix} -z^3-z^2-z & z-1 \\ -z^2+3z & z-2 \end{bmatrix} .$$

This can be explained as follows : Algorithm A produces only one specific minimal n-th order parital k-realization. We have however enough information to find all minimal solutions with $A_n(z)$ column reduced, hence all minimal solutions. Indeed, it was explained before that all reduced $A(z)$ that solve the problem can be found by replacing the i-th column of $A_{n+1}(z)$, $a_i^{n+1}(z) \in [n+1,\alpha]_k$ say, by a linear combination of the form

$$\sum_{h_j(z) \in B(n+1,\alpha)_k} \mu_j h_j(z)$$

on condition that the highest degree coefficients remain linearly independent. In our case is $k = -1$ and

$$\begin{aligned} B(7,1)_k &= [7,1]_k \cup z[7,0]_k \cup [6,0]_k = [7,1]_k \\ &= \left\{ \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \right\} \end{aligned}$$

and

$$\begin{aligned} B(7,3)_k &= [7,3]_k \cup z[7,2]_k \cup z^2[7,1]_k \cup z^3[7,0]_k \\ &\quad \cup [6,2]_k \cup z[6,1]_k \cup z^2[6,0]_k \\ &\quad \cup [5,1]_k \cup z[5,0]_k \\ &\quad \cup [4,0]_k \\ &= [7,3]_k \cup z^2[7,1]_k \cup z[6,1]_k \cup [5,1]_k \\ &= \left\{ \begin{bmatrix} z^3 \\ -z \end{bmatrix} \ , \ z^2 \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \ , \ z \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \ , \ \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \right\} . \end{aligned}$$

Thus all minimal solutions are given by

$$A_7(z) = [a_1^7(z) \ \ a_2^7(z)]$$

with

$$a_1^7(z) = \mu \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \ , \ \mu \neq 0$$

$$a_2^7(z) = \nu_1 \begin{bmatrix} z^3 \\ -z \end{bmatrix} + \nu_2 z^2 \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} + \nu_3 z \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} + \nu_4 \begin{bmatrix} z-1 \\ z-2 \end{bmatrix} \ , \ \nu_1 \neq 0$$

and the solution A_7' is a special case of this for

$$\mu = 1 \text{ and } \nu_1 = \nu_3 = -1, \ \nu_2 = \nu_4 = 0.$$

Example 3

One final example (see Kalman [13,p.399], Graves-Morris [11])

$$T(z) = \begin{bmatrix}1\\1\end{bmatrix} z^{-1} + \begin{bmatrix}1\\2\end{bmatrix} z^{-2} + \begin{bmatrix}1\\3\end{bmatrix} z^{-3} + \begin{bmatrix}2\\4\end{bmatrix} z^{-4} + \begin{bmatrix}1\\5\end{bmatrix} z^{-5} + \begin{bmatrix}3\\6\end{bmatrix} z^{-6} + \cdots .$$

For $k = -1$ we get as an initialization :

$$M_1 = \underset{(0)}{[\,|\ 1\ ]},$$

$$RM_1 = \begin{bmatrix}|\ 1\\|\ 1\end{bmatrix} \rightarrow A_2 = [z].$$

And further

$$M_2 = \underset{(1)\ \ (1)}{[\ 1\ |\ z\ ]}$$

$$RM_2 = \begin{bmatrix}1 & | & 1\\1 & | & 2\end{bmatrix} \rightarrow A_3 = [z^2]$$

$$M_3 = \underset{(2)\ (2)\ \ (2)}{[\ 1\ \ z\ |\ z^2\ ]},$$

$$RM_3 = \begin{bmatrix}1 & 1 & | & 1\\1 & 2 & | & 3\end{bmatrix} \rightarrow A_4 = [z^2-2z+1]$$

$$M_4 = \underset{(3)\ (3)\ \ \ (2)}{[\ 1\ \ z\ |\ z^2-2z+1\ ]},$$

$$RM_4 = \begin{bmatrix}1 & 1 & | & 1\\1 & 2 & | & 0\end{bmatrix} \rightarrow A_5 = [z^3-2z^2+2z-2]$$

$$M_5 = \underset{(3)\ \ \ (4)\ \ \ \ \ (3)}{[\ z^2-2z+1\ \ 1\ |\ z^3-2z^2+2z-2\ ]},$$

$$RM_5 = \begin{bmatrix}1 & 1 & | & -3\\0 & 1 & | & -1\end{bmatrix} \rightarrow A_6 = [z^4+z^3-4z^2+2z+1]$$

$$M_6 = \underset{(4)\ \ \ \ \ (4)}{[\ z^2-2z+1\ \ z^3-2z+2z-2\ |\ z^4+z^3-4z^2+2z+1\ ]},$$

$$RM_6 = \begin{bmatrix}1 & -3 & | & -3\\0 & -1 & | & -1\end{bmatrix} \rightarrow A_7 = [z^4-2z^2+1].$$

So that at this step we have a solution

$$C_7(z)A_7(z)^{-1} = \begin{bmatrix} z^3+z^2-z \\ z^3+2z^2+z \end{bmatrix} (z^4-2z^2+1)^{-1} = \begin{bmatrix} \frac{z^3+z^2-z}{(z^2-1)^2} \\ \frac{z}{(z-1)^2} \end{bmatrix} .$$

15. Changing the structure of the data

Graves-Morris [11] develops an algorithm for vector valued interpolation. It can also be used for the case where all interpolation points are at ∞. But then the problem has to be tranformed in a vector valued one, i.e. in a SIMO (single input-multi output) problem like in example 3 in the previous section. We illustrate this by another example. Take again example 2 of the previous section but now rearranged in the form.

Example 4

$$T(z) = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} z^{-2} + \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} z^{-3} + \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} z^{-4} + \begin{bmatrix} 1 \\ 4 \\ 1 \\ 1 \end{bmatrix} z^{-5} + \begin{bmatrix} 2 \\ 7 \\ 1 \\ 2 \end{bmatrix} z^{-6} + \cdots .$$

We successively find the following solutions :

$$M_1 = \underset{(0)}{[\,|\ 1\ ]},$$

$$RM_1 = \begin{bmatrix} | & -1 \\ | & 1 \\ | & 0 \\ | & 0 \end{bmatrix} \rightarrow A_2 = z$$

$$M_2 = \underset{(1)\quad(1)}{[\ 1\ |\ z\ ]},$$

$$RM_2 = \begin{bmatrix} -1 & | & 0 \\ 1 & | & 1 \\ 0 & | & 0 \\ 0 & | & 0 \end{bmatrix} \rightarrow A_3 = z^2$$

$$M_3 = \underset{(2)\ (2)\quad(2)}{[\ 1\quad z\ |\ z^2\ ]},$$

$$RM_3 = \begin{bmatrix} -1 & 0 & | & 1 \\ 1 & 1 & | & 1 \\ 0 & 0 & | & -1 \\ 0 & 0 & | & 1 \end{bmatrix} \rightarrow A_4 = z^3$$

$$M_4 = \begin{bmatrix} 1 & z & z^2 & | & z^3 \\ (3) & (3) & (3) & & (3) \end{bmatrix},$$

$$RM_4 = \left[\begin{array}{cccc|c} -1 & 0 & 1 & & 1 \\ 1 & 1 & 1 & & 2 \\ 0 & 0 & -1 & & 0 \\ 0 & 0 & 1 & & 1 \end{array}\right] \rightarrow A_5 = z^4$$

$$M_5 = \begin{bmatrix} 1 & z & z^2 & z^3 & | & z^4 \\ (4) & (4) & (4) & (4) & & (4) \end{bmatrix},$$

$$RM_5 = \left[\begin{array}{cccc|c} -1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 & 4 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right] \rightarrow A_6 = z^4-2z^3+z^2-z$$

$$M_6 = \begin{bmatrix} 1 & z & z^2 & z^3 & | & z^4-2z^3+z^2-z \\ (5) & (5) & (5) & (5) & & (4) \end{bmatrix},$$

$$RM_6 = \left[\begin{array}{cccc|c} -1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \rightarrow A_7 = z^4-2z^3+z^2-z.$$

The solution at this moment is

$$C_7(z)A_7(z)^{-1} = \begin{bmatrix} -z^3+2z^2 \\ -z+2 \\ z^3-z^2 \\ z-1 \end{bmatrix} (z^4-2z^3+z^2-z)^{-1}$$

which can be rearranged as

$$\begin{bmatrix} -z^3+2z^2 & z^3-z^2 \\ -z+2 & z-1 \end{bmatrix} (z^4-2z^3+z^2-z)^{-1}.$$

As you can easily verify, this is the same solution as we obtained before. The same solution was also found by Graves-Morris.

This example may suggest that we can change the structure of the problem from MIMO to SIMO or MISO at free will but this is not true in general. We illustrate this with a counterexample.

Example 5

$$T(z) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} z^{-2} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} z^{-3} + \cdots$$

The solutions are constructed as follows :

$$M_1 = \left[\begin{array}{c|cc} & 1 & 0 \\ & 0 & 1 \end{array}\right], \qquad \begin{matrix} \\ (0)\ (0) \end{matrix}$$

$$RM_1 = \left[\begin{array}{c|cc} & 1 & 0 \\ & 0 & 0 \end{array}\right] \rightarrow A_2 = \begin{bmatrix} 0 & z \\ 1 & 0 \end{bmatrix}$$

$$M_2 = \left[\begin{array}{cc|cc} 1 & & 0 & z \\ 0 & & 1 & 0 \end{array}\right], \qquad \begin{matrix} (1) & (0)\ (1) \end{matrix}$$

$$RM_2 = \left[\begin{array}{c|cc} 1 & 1 & 1 \\ 0 & 1 & 1 \end{array}\right] \rightarrow A_3 = \begin{bmatrix} z & 0 \\ -z & z^2 \end{bmatrix}$$

$$M_3 = \left[\begin{array}{cc|cc} 0 & 1 & z & 0 \\ 1 & 0 & -z & z^2 \end{array}\right], \qquad \begin{matrix} (1)\ (2) & (1)\ (2) \end{matrix}$$

$$RM_3 = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right] \rightarrow A_4 = \begin{bmatrix} z^2-1 & z^2 \\ -z^2+z & 0 \end{bmatrix}.$$

The numerator for level 4 is $C_4(z) = \begin{bmatrix} z & z+1 \\ 0 & 1 \end{bmatrix}$, so that we find

$$C_4(z)A_4(z)^{-1} = \begin{bmatrix} z^3-z & z^2-z-1 \\ z^2-z & z^2-1 \end{bmatrix} (z^4-z^3)^{-1}.$$

All solutions can be reconstructed from table 4. Of course several solutions are possible for level 4, but all of them can be realized with 4 states (= the degree of det $A_4(z)$). Now let us rearrange this example into a MISO form :

$$T(z) = [1\ 0\ 0\ 0]z^{-1} + [1\ 1\ 1\ 1]z^{-2} + [0\ 0\ 0\ 1]z^{-3} + \cdots$$

The solutions are then constructed as follows :

$$M_1 = \left[\begin{array}{c|cccc} & 1 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 \\ & 0 & 0 & 1 & 0 \\ & 0 & 0 & 0 & 1 \end{array}\right], \qquad \begin{matrix} \\ \\ \\ (0)\ (0)\ (0)\ (0) \end{matrix}$$

degree → / level ↓	0	1	2
1	$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$		
2	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}z\\0\end{bmatrix}$	
3		$\begin{bmatrix}z\\-z\end{bmatrix}$	$\begin{bmatrix}0\\z^2\end{bmatrix}$
4			$\begin{bmatrix}z^2-1\\-z^2+z\end{bmatrix}$ $\begin{bmatrix}z^2\\0\end{bmatrix}$

Table 4

$$RM_1 = [|1\ 0\ 0\ 0] \rightarrow A_2 = \begin{bmatrix} 0 & 0 & 0 & z \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

$$M_2 = \left[\begin{array}{c|cccc} 1 & 0 & 0 & 0 & z \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right], \qquad (1)\quad (0)\ (0)\ (0)\ (1)$$

$$RM_2 = [1 \mid 1\ 1\ 1\ 1] \rightarrow A_3 = \begin{bmatrix} 0 & 0 & -1 & z \\ -1 & -1 & z & -z \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$M_3 = \left[\begin{array}{c|cccc} 0 & 0 & 0 & -1 & z \\ 1 & -1 & -1 & z & -z \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right], \qquad (1)\quad (0)\ (0)\ (1)\ (1)$$

$$RM_3 = [1 \mid 0\ 1\ -1\ 0] \rightarrow A_4 = \begin{bmatrix} 0 & 0 & -1 & z \\ -1 & -z-1 & 0 & -z \\ 1 & 0 & 0 & 0 \\ 0 & z & z & 0 \end{bmatrix}$$

and the solution is now

$$C_4(z)A_4(z)^{-1} = [-z^2-z \quad -z \quad -z \quad -z-1]\,(-z^3)^{-1}.$$

All solutions can be realized with 3 states (instead of 4). Still another variant of the same example could be to rearrange it in a SIMO form :

$$T(z) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} z^{-2} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} z^{-3} + \cdots .$$

And now the solutions are constructed as :

$$M_1 = \underset{(0)}{[\mid\ 1\ ]},$$

$$RM_1 = \begin{bmatrix} \mid 1 \\ \mid 0 \\ \mid 0 \\ \mid 0 \end{bmatrix} \rightarrow A_2 = [z]$$

$$M_2 = \underset{(1)\quad(1)}{[\ 1\ \mid\ z\ ]},$$

$$RM_2 = \begin{bmatrix} 1 \mid 1 \\ 0 \mid 1 \\ 0 \mid 1 \\ 0 \mid 1 \end{bmatrix} \rightarrow A_3 = [z^2]$$

$$M_3 = \underset{(2)\ (2)\quad(2)}{[\ 1\quad z\ \mid\ z^2\ ]},$$

$$RM_3 = \begin{bmatrix} 1\ 1 \mid 0 \\ 0\ 1 \mid 0 \\ 0\ 1 \mid 0 \\ 0\ 1 \mid 1 \end{bmatrix} \rightarrow A_4 = [z^3].$$

The solution is now :

$$C_4(z)A_4(z)^{-1} = \begin{bmatrix} z^2+z \\ z \\ z \\ z+1 \end{bmatrix} z^{-3} \qquad (10)$$

which is the same solution as before. Again all solutions can be realized with 3 states.

Some insight can be gained by considering the 4 scalar series involved in this example separately. I.e. we look at the series :

(1) $t_1(z) = 1.z^{-1} + 1.z^{-2} + 0.z^{-3} + \cdots$

(2),(3) $t_2(z) = t_3(z) = 0.z^{-1} + 1.z^{-2} + 0.z^{-3} + \cdots$

(4) $t_4(z) = 0.z^{-1} + 1.z^{-2} + 1.z^{-3} + \cdots$

Let us solve the problem for each of these separately :

(1) $t_1(z) = 1.z^{-1} + 1.z^{-2} + 0.z^{-3} + \cdots$

The solution vectors are generated as follows :

$M_1 = [\,|\ 1\ ]$,
(0)

$RM_1 = [|1] \rightarrow A_2 = [z]$

$M_2 = [\ 1\ |\ z\ ]$,
(1) (1)

$RM_2 = [1\,|\,1] \rightarrow A_3 = [z-1]$

$M_3 = [\ 1\ |\ z-1\ ]$,
(2) (1)

$RM_3 = [1\,|\,-1] \rightarrow A_4 = [z^2-z+1]$.

The solutions are summarized in table 5.

degree → / level ↓	0	1	2
1	1		
2		z	
3		$z-1$	
4			z^2-z+1

Table 5

From this table we see that all solutions for level 4 of degree 2 can be written as

$\alpha(z^2-z+1) + \beta(z-1)$ with $\alpha,\beta \in K$, $\alpha \neq 0$

and every polynomial of degree 3,4,5, ... is a solution.

(2),(3) $t_2(z) = t_3(z) = 0.z^{-1} + 1.z^{-2} + 0.z^{-3} + \cdots$
Now the solutions are found as :

$M_1 = \underset{(0)}{[\,|\ 1\ ]}$,

$RM_1 = [|0] \rightarrow A_2 = [1]$

$M_2 = \underset{(0)}{[\,|\ 1\ ]}$,

$RM_2 = [|1] \rightarrow A_3 = [z^2]$

$M_3 = \underset{(1)\quad(2)}{[\ 1\ |\ z^2\ ]}$,

$RM_3 = [1 \mid 0] \rightarrow A_4 = [z^2]$.

This is summarized in table 6.

degree → / level ↓	0	1	2
1	1		
2	1		
3			z^2
4			z^2

Table 6

Thus all solutions for level 4 of degree 2 can be written as

$\alpha z^2 + \beta.1$ with $\alpha,\beta \in K$, $\alpha \neq 0$

and every polynomial of degree $\geqslant 3$ is a solution.

(4) Finally consider $t_4 = 0.z^{-1} + 1.z^{-2} + 1.z^{-3} + \cdots$
The solutions are found as :

$M_1 = \underset{(0)}{[\,|\ 1\ ]}$,

$RM_1 = [|0] \rightarrow A_2 = 1$

$$M_2 = \underset{(0)}{[|\ 1\ ]},$$

$RM_2 = [|1] \rightarrow A_3 = z^2$

$$M_3 = \underset{(1)\quad(2)}{[\ 1\ |\ z^2\ ]},$$

$RM_3 = [1 \mid 1] \rightarrow A_4 = z^2 - z.$

degree → / level ↓	0	1	2
1	1		
2	1		
3			z^2
4			z^2-z

Table 7

And all solutions for level 4 of degree 2 can be written as

$\alpha(z^2-z) + \beta 1 \quad , \quad \alpha,\beta \in K, \ \alpha \neq 0$

and any polynomial of degree ≥ 3 is a solution.

Now we try to reassemble these solutions to get a solution for the SIMO case considered above. Each of the scalar series allows for a degree 2 denominator. However the SIMO case will have a degree 2 denominator solution only if there is a common solution to the 4 scalar problems which is of degree 2. Now, since this is not the case, the minimal degree solution for the SIMO case will be 3, which is exactly what we have found before.
The MISO case is not different from the SIMO case. Every solution of the SIMO case can be written in MISO form and conversely. The number of states in the minimal realization will thus be the same (see Kailath [12,p.381]).

Now consider our original MIMO formulation which is here a 2×2 problem. To find now the 2×2 denominator polynomial of a minimal solution, much more conditions have to be fulfilled. To write the 2×2 solution $C_4(z)A_4(z)^{-1}$ in the form $\tilde{C}_4(z)d(z)^{-1}$ with $d(z) = \det A_4(z)$, we should construct a polynomial $d(z)$ such that in general

(1) it is of minimal degree

(2) it is a common element of the solutions constructed for the scalar problems for the level n considered

(3) it is the determinant of a $m \times m$ polynomial matrix which is column reduced. Hence the degree of $d(z)$ is the sum of the degrees of the columns of $A_n(z) \in K^{m \times m}[z]$. The columns of $A_n(z)$ must have an order at least n.

For our example, these conditions can only be met for a degree of $d(z)$ which is 4.

Remark

For SIMO systems, conditions (1) and (2) imply (3). For the general case, conditions (1) and (3) imply condition (2). Thus the minimal number of states for MISO and SIMO systems are the same and smaller than or equal to the degree of the other MIMO structured possibilities.

To illustrate the previous conditions. Suppose we start from the SIMO solution (10) and restructure it as :

$$C_4(z)A_4(z)^{-1} = \begin{bmatrix} z^2+z & z \\ z & z+1 \end{bmatrix} z^{-3} = \begin{bmatrix} z^2+z & z \\ z & z+1 \end{bmatrix} \begin{bmatrix} z^3 & 0 \\ 0 & z^3 \end{bmatrix}^{-1}$$

A greatest common right divisor is $\begin{bmatrix} z & z+1 \\ 0 & 1 \end{bmatrix}$, so we can rewrite it (Kailath [12,p.377]) as

$$\begin{bmatrix} z+1 & -z^2-z-1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z^2 & -z^3-z^2 \\ 0 & z^3 \end{bmatrix}^{-1}$$

which has a colomn reduced denominator.
The degree of the determinant of the denominator is 5. Thus the realization has 5 states. The direct solution of the MIMO formulated problem needed only 4 states which shows that the previous solution is not minimal.

16. The minimal extension problem

Suppose we have a problem where the Markov parameters are only partially known from a certain point on. We replace the unknown entries by free parameters. The minimal extension problem is then to solve a partial realization problem and give values to the free parameters such that the degree of the realization is kept as low as possible. The basis we have constructed to give all the solution vectors will be useful in solving this problem. The general principle can be illustrated by some examples.

Example 6 (compare with example 3)

Given :

$$T(z) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} z^{-1} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} z^{-2} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} z^{-3} + \begin{bmatrix} 2 \\ \alpha \end{bmatrix} z^{-4} + \begin{bmatrix} 1 \\ \beta \end{bmatrix} z^{-5} + \begin{bmatrix} 3 \\ \gamma \end{bmatrix} z^{-6} + \cdots .$$

The problem is to choose α, β and γ such that the degree of the minimal realization (shift parameter $k=-1$) is as small as possible.
The first part of the computations runs as in example 3. We find $A_1 = 1$, $A_2 = z$, $A_3 = z^2-2z+1$. From then on we have

$$M_4 = \begin{bmatrix} 1 & z-1 & | & z^2-2z+1 \\ (3) & (3) & & (2) \end{bmatrix},$$

$$RM_4 = \begin{bmatrix} 1 & 0 & | & 1 \\ 1 & 1 & | & -4+\alpha \end{bmatrix} \rightarrow A_5 = z^3-2z^2+(6-\alpha)z+(-6+\alpha)$$

$$M_5 = \underset{(3)}{[\ z^2-2z+1} \quad \underset{(4)}{z-1} \mid A_5\ ],$$

$$RM_5 = \left[\begin{array}{cc|c} 1 & 0 & -3 \\ -4+\alpha & 1 & -3\alpha+\beta+6 \end{array}\right].$$

We can keep the degree of A_6 down to 3 if we set $\beta=6$ and we then find

$$A_6(z) = z^3+z^2-\alpha z+(\alpha-3)$$

$$M_6 = \underset{(4)}{[\ z^2-2z+1} \quad \underset{(5)}{z-1} \mid \underset{(3)}{A_6}\ ],$$

$$RM_6 = \left[\begin{array}{cc|c} 1 & 0 & 1-\alpha \\ -4+\alpha & 1 & \gamma+3\alpha-\alpha^2-3 \end{array}\right].$$

Also $A_7(z)$ can be kept of degree 3 if we set $\alpha=1$ and $\gamma=1$. We find

$$A_7(z) = z^3+z^2-z-2.$$

If we had chosen $\beta\neq 6$ or $\alpha\neq 1$ or $\gamma\neq 1$, the degree of $A_7(z)$ would have been greater than 3 (see example 3 : $\alpha=3$, $\beta=5$, $\gamma=6$, and $A_7(z)$ is of degree 4).

This example is a bit misleading because if many suggest that the parameters could be chosen at each level such that the number of states for that level is minimal and then continue with that parameter as it was chosen. In general, this is not true. In general we have to consider the final solution with all parameters in it and then make an optimal choice as the following example illustrates.

Example 7

Given :

$$T(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z^{-2} + \begin{bmatrix} \alpha & 0 \\ 0 & 1 \end{bmatrix} z^{-3} + \begin{bmatrix} \beta & 1 \\ 1 & 0 \end{bmatrix} z^{-4} + \cdots$$

α and β have to be chosen such that the minimal partial realization $C_5(z)A_5(z)^{-1}$ has the minimal number of states ($k = -1$).
The computations go as follows :

$$M_1 = \underset{(0)\ (0)}{\left[\begin{array}{c|cc} & 1 & 0 \\ & 0 & 1 \end{array}\right]},$$

$$RM_1 = \left[\begin{array}{c|cc} & 1 & 0 \\ & 0 & 1 \end{array}\right] \rightarrow A_2 = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix}$$

$$M_2 = \underset{(1)\ (1)\quad (1)\ (1)}{\left[\begin{array}{cc|cc} 1 & 0 & z & 0 \\ 0 & 1 & 0 & z \end{array}\right]},$$

$$RM_2 = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array}\right] \rightarrow A_3 = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix}$$

$$M_3 = \begin{array}{c} \left[\begin{array}{cc|cc} 1 & 0 & z & 0 \\ 0 & 1 & 0 & z \end{array}\right] \\ (2)\ (2)\quad (1)\ (1) \end{array},$$

$$RM_3 = \left[\begin{array}{cc|cc} 1 & 0 & \alpha & 0 \\ 0 & 1 & 0 & 1 \end{array}\right].$$

Now we have several possibilities :

(a) $\alpha = 0$.

$$\rightarrow A_4 = \begin{bmatrix} z & 0 \\ 0 & z^2-1 \end{bmatrix}$$

$$M_4 = \begin{array}{c} \left[\begin{array}{cc|cc} 0 & 1 & z & 0 \\ z & 0 & 0 & z^2-1 \end{array}\right] \\ (2)\ (3) \end{array},$$

$$RM_4 = \left[\begin{array}{cc|cc} 0 & 1 & \beta & 1 \\ 1 & 0 & 1 & 0 \end{array}\right].$$

Again we have two possibilities

(1) $\beta = 0$

then $A_5 = \begin{bmatrix} z^2 & -1 \\ -z & z^3-z \end{bmatrix}$

(2) $\beta \neq 0$

then $A_5 = \begin{bmatrix} z^3-\beta & z^2 \\ -z^2 & -\beta z^2-z+\beta \end{bmatrix}$.

(b) $\alpha \neq 0$

$$\rightarrow A_4 = \begin{bmatrix} z^2-\alpha & 0 \\ 0 & z^2-1 \end{bmatrix}$$

$$M_4 = \begin{array}{c} \left[\begin{array}{cc|cc} z & 0 & z^2-\alpha & 0 \\ 0 & z & 0 & z^2-1 \end{array}\right] \\ (2)\ (2)\quad (2)\quad (2) \end{array},$$

$$RM_4 = \left[\begin{array}{cc|cc} \alpha & 0 & \beta & 1 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

$$\rightarrow A_5 = \begin{bmatrix} \alpha z^2-(\alpha+\beta)z-\alpha^2 & -z \\ 0 & \alpha z^2-\alpha \end{bmatrix} .$$

Note that $C_4(z)A_4(z)^{-1}$ from (a) can be realized with 3 states and $C_4(z)A_4(z)^{-1}$ from (b) with 4 states, but $C_5(z)A_5(z)^{-1}$ from (a) needs 5 states while in case (b) 4 states are sufficient. A similar observation is also true for SIMO systems as the following example illustrates

Example 8

Given :

$$T(z) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} z^{-2} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} z^{-3} + \begin{bmatrix} \alpha \\ 0 \end{bmatrix} z^{-4} + \begin{bmatrix} \beta \\ 0 \end{bmatrix} z^{-5} + \cdots$$

Find the values α and β such that the degree of A_6 is minimal ($k=-1$). The computations are

$$M_1 = \underset{(0)}{[\,|\ 1\ ]} ,$$

$$RM_1 = \begin{bmatrix} | & 1 \\ | & 0 \end{bmatrix} \rightarrow A_2 = z$$

$$M_2 = \underset{(1)\ \ (1)}{[\ 1\ |\ z\ ]},$$

$$RM_2 = \begin{bmatrix} 1 & | & 0 \\ 0 & | & 1 \end{bmatrix} \rightarrow A_3 = z^2$$

$$M_3 = \underset{(2)\ (2)\ \ (2)}{[\ 1\ \ z\ |\ z^2\ ]},$$

$$RM_3 = \begin{bmatrix} 1 & 0 & | & 0 \\ 0 & 1 & | & 0 \end{bmatrix} \rightarrow A_4 = z^2$$

$$M_4 = \underset{(3)\ (3)\ \ (2)}{[\ 1\ \ z\ |\ z^2\ ]},$$

$$RM_4 = \begin{bmatrix} 1 & 0 & | & \alpha \\ 0 & 1 & | & 0 \end{bmatrix} .$$

(a) $\alpha=0$

$\rightarrow A_5 = z^2$,

$$M_5 = \underset{(4)}{[\ 1} \quad \underset{(4)}{z} \ \mid \ \underset{(2)}{z^2}\ ],$$

$$RM_5 = \begin{bmatrix} 1 & 0 & \mid & \beta \\ 0 & 1 & \mid & 1 \end{bmatrix} \rightarrow A_6 = z^4 - z - \beta$$

(b) $\alpha \neq 0$

$$\rightarrow A_5 = z^3 - \alpha$$

$$M_5 = \underset{(3)}{[\ z^2} \quad \underset{(4)}{z} \ \mid \ \underset{(3)}{z^3-\alpha}\ ],$$

$$RM_5 = \begin{bmatrix} \alpha & 0 & \mid & \beta \\ 0 & 1 & \mid & 1-\alpha \end{bmatrix}$$

(1) $\alpha = 1 \quad \rightarrow A_6 = \alpha z^3 - \beta z^2 - \alpha^2 = z^3 - \beta z^2 - 1$

(2) $\alpha \neq 1, \neq 0 \rightarrow A_6 = \alpha z^4 - \beta z^3 + (-\alpha^2 + \alpha - 1)z$.

In case (a) A_5 has degree 2 and A_6 degree 4.
In case (b) A_5 has degree 3 and A_6 degree 4 as before if $\alpha \neq 1$ and $\neq 0$. Moreover in case (b1), A_6 has only degree 3 which is the optimal one.

17. Computing solution vectors for another value of the shift parameter

In the previous section we have given an algorithm to compute the solution vectors $[n,\alpha]_k$ for a fixed value of the shift parameter k. In this section we shall show how to compute the solution vectors for level n with shift parameter $k+1$ from the solution vectors of level $n+1$ for shift parameter k. In analogy with the operators H and z we used earlier, we shall now introduce the operators L and z^{-1}.

Definition

Let $a(z) = \sum_{i=0}^{\alpha} a_i z^i \in K^m[z]$ be a polynomial and as before, a is the vector of its coefficients. Then L is the operator associating with a the coefficient $a_0 : La = a_0$ and z^{-1} is defined for all a such that $a(z) \in zK^m[z]$ (i.e. with $a(0) = 0$) by $z^{-1}a = a'$ with a' the coefficient vector for the polynomial $a'(z) = a(z)/z$.
It can be simply verified by inspection that the set $zS(n,\alpha)_{k+1}$ of solutions for the system $(n,\alpha)_{k+1}$ is exactly the subset $S_0(n+1,\alpha+1)_k$ of all solutions a for the system $(n+1,\alpha+1)_k$ for which $La = 0$. I.e.

$$zS(n,\alpha)_{k+1} = S_0(n+1,\alpha+1)_k := \{a \in S(n+1,\alpha+1)_k \mid La = 0\}.$$

It is equally simple to verify that all solutions of $(n+1,\alpha)_k$ are also solutions of $(n,\alpha)_{k+1}$. I.e.

$$S(n+1,\alpha)_k \subset S(n,\alpha)_{k+1}.$$

Hence, in addition to all the relations of table 1, we also have relations like in table 8. We shall now try to generate the solution vectors b_i^n for level n and shift parameter $k+1$ from the solution vectors a_i^{n+1} for level $n+1$ and shift parameter k.

— To keep the highest degree coefficients linearly independent we shall generate a b_i^n based on a_i^{n+1}, i.e. the highest degree coefficient of b_i^n should be given by

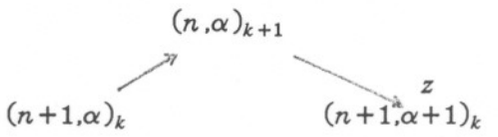


Table 8

$$Hb_i^n = Ha_i^{n+1} + \sum_{j=1}^{i-1} \mu_j \, Ha_j^{n+1}.$$

— If $b_i^n \in [n,\alpha]_{k+1} \subset S(n,\alpha)_{k+1} = z^{-1} S_0(n+1,\alpha+1)_k$ then it is clear that we should look for vectors in $S(n+1,\alpha+1)_k$ for which the lowest degree coefficient vanishes. Suppose that the a_i^{n+1}, $i = 1,2,\ldots,m$ have (potential) degrees $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_m$, then a vector in $S(n+1,\alpha_i)_k$ based on a_i^{n+1} can be written as (in analogy with (6))

$$a_i^{n+1} + \sum_{j=1}^{i-1} \mu_j \, z^{\alpha_i - \alpha_j} \, a_j^{n+1} + \sum_{h_j \in B(n,\alpha_i-1)_k} \nu_j \, h_j . \tag{11}$$

— The problem now is *not* to make the residual of this expression vanish, but to make the lowest degree coefficient of this expression zero. Hence of all the vectors appearing in the summations of (11) it is sufficient to keep only a set of *auxiliary vectors of the second kind*, y_j say, whose lowest degree coefficients form a basis for the space generated by the lowest degree coefficients of all these vectors. Thus (11) can be replaced by

$$a_i^{n+1} + \sum_{j=1}^{r} \mu_j \, y_j \tag{12}$$

where, of course, the potential degree of the y_j in the summation must be bounded by α_i. The μ_j are defined by the condition

$$L\, a_i^{n+1} = - \sum_{j=1}^{r} \mu_j \, L\, y_j , \quad \mu_r \neq 0. \tag{13}$$

— Suppose we start our computation of the b_i^n, $i = 1,2,\ldots,m$ with the solution vectors a_i^{n+1}, $i = 1,2,\ldots,m$ and a set Y of auxiliary vectors y_i, $i = 1,2,\ldots,m$. The degrees of the a_i^{n+1} are $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_m$ and the (potential) degrees of the y_i are $\sigma_1 \leqslant \sigma_2 \leqslant \cdots \leqslant \sigma_m$. Note that we have m auxiliary vectors of the second kind since the vector space generated by the lowest degree coefficients is K^m from the very beginning for $n=1$ since $\text{span}_K \{a_i^1\} = K^m$ (we have chosen $a_i^1 = e_i$). There are two possibilities

1. (13) is satisfied with $\sigma_r \leqslant \alpha_i$. Then we have found a vector (12) which is based on a_i^{n+1} of degree α_i and which has a vanishing lowest degree coefficient. Hence it is in $S_0(n+1,\alpha_i)_k$. Hence

$$b_i^n = z^{-1} (a_i^{n+1} + \sum_{j=1}^{r} \mu_j \, y_j) \in [n,\alpha_i - 1]_{k+1}.$$

2. (13) can only be satisfied with $\sigma_r > \alpha_i$. In that case, zb_i^n can not be in $[n,\alpha_i]_{k+1}$. Thus zb_i^n has a degree which is at least α_i+1, i.e. it must be in $S_0(n+1,\alpha_i+1)_k$. Now, a vector in $S_0(n+1,\alpha_i+1)_k$ which is based on a_i^{n+1} can simply be chosen to be za_i^{n+1}, so that in that case

$$b_i^n = a_i^{n+1} \in [n,\alpha_i]_{k+1}.$$

These observations show that it is important to keep the degrees of the auxiliary vectors as low as possible in order to make the degree of b_i^n as low as possible. Thus if (13) is satisfied with $\sigma_r > \alpha_i$, we should replace y_r by a_i^{n+1}.

Thus we can formulate the following algorithm.

B.1 The set of auxiliary vectors Y for level 1 is $\{e_1,e_2,...,e_m\}$, all of degree 0.

B.2 Compute for $n = 1,2,3, \ldots$ the solution vectors b_i^n of level n for the shift parameter $k+1$; given

- $Y = [y_1, \ldots, y_m]$, the set of auxiliary vectors for this level with potential degrees $\sigma_1 \leqslant \sigma_2 \leqslant \cdots \leqslant \sigma_m$ (with respect to level $n+1$),
- $A = [a_1^{n+1}, \ldots, a_m^{n+1}]$ the set of solution vectors for level $n+1$ and shift parameter k, with degrees $\alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_m$.

B.2.1 For $i = 1,2, \ldots, m$
look for the solution b_i^n of level n for shift parameter $k+1$, which is based on a_i^{n+1} as follows.

B.2.1.1 Write $L\, a_i^{n+1}$ as a linear combination of $L\, y_1, \ldots, L\, y_r$

$$L\, a_i^{n+1} = -\sum_{j=1}^{r} \mu_j\, L\, y_j$$

with $r \leqslant m$ as low as possible ($\mu_r \neq 0$).
If $L\, a_i^{n+1} = 0$ set $r=0$ and $\sigma_0=0$.

B.2.1.1.1 **If** $\alpha_i < \sigma_r$ **then**

B.2.1.1.1.1 Set $b_i^n = a_i^{n+1}$ ($\in [n,\alpha_i]_{k+1}$).

B.2.1.1.1.2 Replace y_r by a_i^{n+1} at the correct place in Y.

B.2.1.1.2 **If** $\alpha_i \geqslant \sigma_r$ **then**

B.2.1.1.2.1 Set

$$b_i^n = z^{-1}(a_i^{n+1} + \sum_{j=1}^{r} \mu_j y_j)\ (\in [n,\alpha_i-1]_{k+1})$$

Note

We have only computed the solution vectors with shift parameter $k+1$ from the solution vectors with shift parameter k. The other direction, i.e. computing the solution vectors with shift parameter $k-1$ from the solution vectors with shift parameter k has computationally no advantage.

18. Connection with Padé approximation

We shall now give a reformulation of the minimal realization problem which is a matrix version of the scalar Padé approximation problem. Because our partial realization problem was subject to a minimality condition, we conjecture that it is possible to give a definition for a Padé approximation for the matrix case which is minimal in a certain sense. What this definition should be is not clear for the moment and it is still under investigation. The shift parameter k which was introduced in the minimal partial k-realization problem was needed to find the different elements in a Padé table for the scalar case, each k referring to a certain (descending) staircase in the Padé table.

We shall have to start with an investigation of the table of solution vectors for the partial realization problem. Consider a general table of solution vectors like table 3 (see example 2 of section 14). The horizontal axis indicates the degree and the vertical axis indicates the level. The (n,α) entry of the table contains the set $[n,\alpha]_k$ of solution vectors for level n of degree α

for shift parameter k. Concerning this table we have the following property.

Theorem *Let $\{a_i^1\}_{i=1}^m$ be the solution vectors (of degree 0) in $[1,0]_k$. Then by construction, for each level n we will find exactly one solution vector with highest degree component Ha_i^1. If you trace the solution vectors for increasing levels that are based on a_i^1, then it can not move to the left in this table. In other words, the degree of the solution vectors with highest degree component a_i^1 is non decreasing with the level n.*

Proof This result is certainly true for $n=2$ since the solution vectors are all of degree 0 for $n=1$. In general the result follows by induction from the construction of these solution vectors by algorithm A. Suppose a_i^n is a solution vector for level n with degree α'. The solution vector for level $n+1$ with highest degree component Ha_i^n is found as a linear combination of the form (6) with α as low as possible but $\alpha \geqslant \alpha'$. Hence the degree does not decrease.

□

Corollary *Let a_i^n $i = 1,\ldots,m$ be the solution vectors for level n with $n = 1,2,\ldots,N$. Suppose the potential degree of a_i^n (with respect to level N) is α_i^n (which is the lowest possible one). Then $\alpha_i^{n+1} \geqslant \alpha_i^n - 1$.*

Proof This is trivial since the potential degree is constant on a down sloping diagonal and it increases from left to right.
Since solution vectors can not move to the left by previous theorem, we have the relation stated.

□

We shall now give an equivalent formulation of the minimal partial realization problem. We shall use the notation $O_+(z^n)$ for any formal series of the form

$$T_n z^n + T_{n+1} z^{n+1} + \cdots \in z^n K^{p\times m}[[z]].$$

If $T(z) = T_1 z^{-1} + T_2 z^{-2} + \cdots \in z^{-1} K^{p\times m}[[z^{-1}]]$ is given and $C(z)A(z)^{-1}$ is a solution of the order N minimal partial k-realization problem, then

$$z^{-(k+1)} C(z) A(z)^{-1} = T(z) + O_-(z^{-N}).$$

Since $C(z)$ is the polynomial part of $z^{k+1}T(z)A(z)$, the degree of the i-th column of $C(z)$ must be less than or equal to k plus the degree of the i-th column of $A(z)$. Hence

$$z^{k+1} C(z^{-1})[A(z^{-1})]^{-1} = T(z^{-1}) + O_+(z^N)$$

implies

$$[C(z^{-1}) \operatorname{diag}(z^{\alpha_1+k}, \ldots, z^{\alpha_m+k})][A(z^{-1}) \operatorname{diag}(z^{\alpha_1}, \ldots, z^{\alpha_m})]^{-1} = z^{-1}T(z^{-1}) + O_+(z^{N-1})$$

if α_i is the degree of the i-th column of $A(z)$. Now set

$$P(z) := C(z^{-1}) \operatorname{diag}(z^{\alpha_1+k}, \ldots, z^{\alpha_m+k}) \tag{14}$$

which is a polynomial matrix with i-th column $p_i(z)$ of degree at most α_i and set

$$Q(z) = A(z^{-1}) \operatorname{diag}(z^{\alpha_1}, \ldots, z^{\alpha_m}) \tag{15}$$

which is a polynomial matrix with i-th column $q_i(z)$ of degree at most α_i.

Lemma *With $P(z)$ and $Q(z)$ as defined above, we have at least one of both for $i = 1,2,\ldots,m$*

1. *the degree of the i-th column of $P(z)$ is exactly $\alpha_i + k$*

2. *the degree of the i-th column of $Q(z)$ is exactly α_i*

where α_i is the degree of the i-th column of $A(z)$.

Proof As we explained before, α_i+k is an upper bound for the degree of the i-th column $p_i(z)$ of $P(z)$ and α_i is an upper bound for the degree of the i-th column $q_i(z)$ of $Q(z)$. We have to show that at least one of these bounds is obtained.
Suppose that $q_i(z)$ is not of strict degree α_i and suppose $a_i(z)$ is the i-th column of $A(z)$. Then the constant term of $a_i(z)$, viz. $a_i(0)$, is zero. Suppose that also $p_i(z)$ is not of strict degree α_i+k. Then the constant term of $c_i(z)$ (the i-th column of $C(z)$) is also zero. This means that

$$\tilde{a}_i := z^{-1}a_i \in S(n,\alpha_i-1)_k \quad \text{with} \quad n = N-k-1.$$

Hence $\tilde{a}_i(z) = z^{-1}a_i(z)$ is a polynomial of degree α_i-1 with highest degree coefficient Ha_i which is linearly independent of all the highest degree coefficients of the solution vectors, i.e. columns of $A(z)$. Thus we could replace $a_i(z)$ by $\tilde{a}_i(z)$ which is of a lower degree and this contradicts the fact that $A(z)$ is minimal.

□

It follows from this lemma that our original problem (N-th order minimal partial k-realization with column reduced denominator) is equivalent with the following one :
Given $F(z) = F_0 + F_1z + F_2z^2 + \cdots \in K^{p\times m}[[z]]$ and $k \in \mathbb{Z}$ find polynomials $P(z) \in K^{p\times m}[z]$ and $Q(z) \in K^{m\times m}[z]$ such that $Q(0)$ is nonsingular and

1. $P(z)Q(z)^{-1} = F_0 + F_1z + \cdots + F_{N-2}z^{N-2} + F'_{N-1}z^{N-1} + \cdots$
2. $\sum_{i=1}^{m} \max\{\partial p_i-k, \partial q_i\}$ is minimal with $p_i(z)$ the i-th column of $P(z)$ and $q_i(z)$ the i-th column of $Q(z)$ and ∂p_i and ∂q_i indicate their degrees.

Note

$\alpha_i = \max\{\partial p_i-k, \partial q_i\}$ is the degree of the i-th column of $A(z)$ in the original problem and $\sum_{i=1}^{m} \alpha_i = \partial \det A(z)$.

We shall now give the connection with the scalar Padé approximation problem. The reformulated problem becomes for the scalar case :
Given $F(z) = \sum_{k=0}^{\infty} F_k z^k \in K[[z]]$, a scalar formal power series, and $k \in \mathbb{Z}$.
Find a polynomial $P(z) \in K[z]$ and a polynomial $Q(z) \in K[z]$ with $Q(0) \neq 0$ such that

1. $F(z)-P(z)/Q(z) = O_+(z^{N-1})$
2. $\max\{\partial P-k, \partial Q\}$ is minimal.

We shall call $\max\{\partial P-k, \partial Q\}$ the *k-degree* of the rational approximant $P(z)/Q(z)$ for $F(z)$. Thus we have to find an approximant of order $N-1$ (condition 1) which is of minimal k-degree (condition 2). The classical definition of Padé approximation is the following :
Given $F(z)$ as above, and some nonnegative integers t and s, find polynomials $P(z) \in K[z]$ and $Q(z) \in K[z]$ with $Q(0) \neq 0$, such that

1. $F(z) - P(z)/Q(z) = O_+(z^{N-1})$ with $N=t+s+2$
2. $\partial P \leqslant t$ and $\partial Q \leqslant s$

The rational function $P(z)/Q(z)$ is called a *(t/s) Padé approximant* and it is denoted as $[t/s]$. The table with as (t,s) entry the Padé approximant $[t/s]$ is called the *Padé table*. A Padé table is said to be *normal* if all its entries exist. If $F(z) - P(z)/Q(z) = qz^M + O_+(z^{M+1})$ with $q \neq 0$, M is called the *order* of the approximation. All Padé approximants with an equal order of approximation are found on an antidiagonal in the Padé table (see table 9a). Suppose that the Padé table is normal and $k \geq 0$. The table of k-degrees of the Padé approximants looks as in table 9b.

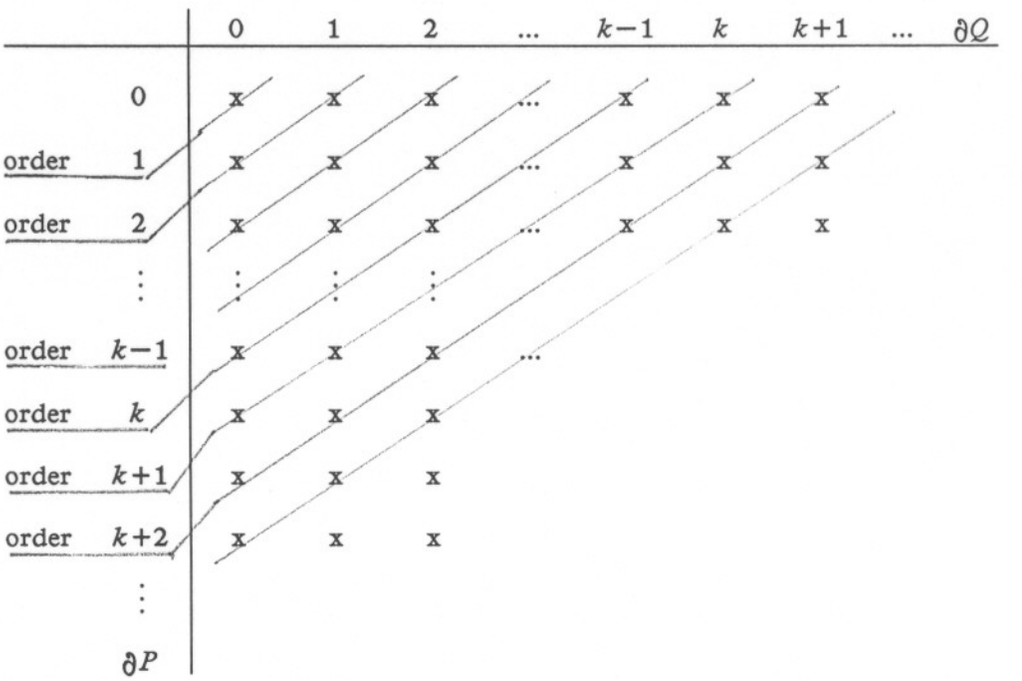


Table 9a

There are $k+1$ elements of order ≥ 1 that have a minimal k degree 0 viz. $[0/0], \ldots, [k/0]$. There are k elements of order ≥ 2 that have a minimal k-degree 0 viz. $[1/0], \ldots, [k/0]$ etc. There is only one element, viz. $[k/0]$ which has an order $\geq k+1$ and has a minimal k-degree zero. From then on, the rule is that for an order of approximation $\geq k+2n$, there are three different elements of minimal k-degree n viz. $[k+n/n-1]$, $[k+n/n]$ and $[k+n-1/n]$ (see table 10) and for an approximant of order $\geq k+2n+1$, there is only one possible choice for it to have a minimal k-degree n viz. $[k+n/n]$ (see table 10). To get away with the freedom that is left we could pick a unique approximant if we choose e.g. the one with the lowest possible numerator degree. In that case we are left with the elements on a staircase : $[0/0]$, $[1/0]$, ..., $[k/0]$, $[k/1]$, $[k+1/1]$, $[k+1/2]$, $[k+2/2]$, $\cdots$. Or we could choose the ones with lowest possible denominator degree and if this still leaves us some freedom (in the beginning) we pick among these the one with the lowest possible numerator degree. Then we keep the elements of another staircase : $[0/0]$, $[1/0]$, ..., $[k/0]$, $[k+1/0]$, $[k+1/1]$, $[k+2/1]$, $[k+2/2]$, $\cdots$. The successive k-degrees of the elements are 0,0,...,0,1,1,2,2,3,3,... in both cases.

A Padé table can be nonnormal, in which case it contains *singular blocks*. These are square blocks in the Padé table where up and above the main antidiagonal all elements are equal to the left top element and below this antidiagonal, the Padé approximants do not exist. What happens then is illustrated by an example for $k = -1$ in table 11 which shows the table of k-degrees of the Padé approximants. An x is placed where the Padé approximants don't exist. Approximants for increasing order and minimal k-degree are found by algorithm A (after a

	0	1	2	...	∂Q
0	0	1	2		
1	0	1	2		
2	0	1	2		
	⋮	⋮	⋮		
$k-1$	0	1	2		
k	0	1	2		
$k+1$	1	1	2		
$k+2$	2	2	2		
⋮					
∂P					

Table 9b

transformation).

	$n-1$	n	$n+1$	∂Q
	⋮	⋮	⋮	order $k+2n$
	$n-1$	n	$n+1$	order $k+2n+1$
$k+n-1$	$n-1$	n	$n+1$	
$k+n$	n	n	$n+1$	
$k+n+1$	$n+1$	$n+1$	$n+1$	
∂P				

Table 10

For our example the successive solutions would be as in table 12. Whenever there is more than one solution for the denominator, any linear combination of these that has a nonzero constant term will also be the denominator of a solution. Note however that the approximant thus obtained may *not* be Padé approximant!

19. Connection with the Berlekamp-Massey algorithm

For a scalar problem (i.e. $p = m = 1$) and for a shift parameter $k = -1$, algorithm A can be restated as follows : Let $T(z) = T_1 z^{-1} + T_2 z^{-2} + \cdots \in z^{-1} K[[z^{-1}]]$ be given and suppose $T_1 = T_2 = \cdots = T_{s-1} = 0$, $T_s \neq 0$ and suppose A_n is the solution (vector) for level

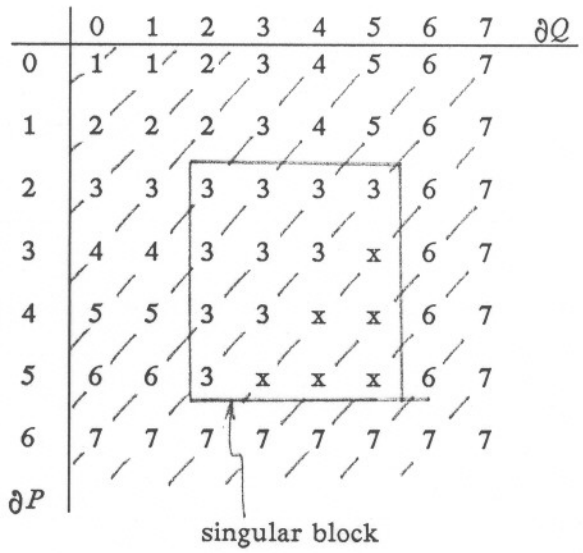


Table 11

n of degree α_n.

level 1

$$M_1 = \underset{(0)}{[\,|\ 1\ ]},$$

$$RM_1 = [|0] \rightarrow A_2 = 1$$

$\vdots$

level s-1

$$M_{s-1} = \underset{(0)}{[\,|\ 1\ ]},$$

$$RM_{s-1} = [|0] \rightarrow A_s = 1$$

level s

$$M_s = \underset{(0)}{[\,|\ 1\ ]},$$

$$RM_s = [|T_s] \rightarrow A_{s+1} = z^s$$

level s+1

$$M_{s+1} = \underset{(1)\quad(1)}{[\ 1\ |\ z^s\ ]},$$

order	k-degree (=minimal)	possible solutions
1	1	[0/0] or [0/1]
2	1	[0/1]
3	2	[1/1], [1/2] or [0/2]
4	2	[1/2]
5	3	[1/3] or [2/2] or any other element from the singular block
6	3	[2/2] or any other element from the singular block
7	3	[2/2] or any other element from the singular block
8	3	[2/2] or any other element from the singular block
9	6	[2/6], [3/6], [4/6] or [5/6]
10	6	[3/6], [4/6] or [5/6]
11	6	[4/6] or [5/6]
12	6	[5/6]
13	7	[6/6], [6/7] or [5/7]
⋮	⋮	⋮

Table 12

$RM_s = [T_s \mid T_{s+1}]$.

The choice we made for $A_{s+1}(z) = z^s$ is a particular choice since *any* polynomial of degree s is a solution. This particular choice makes the degree of $z^s\ A_{s+1}(z^{-1}) = 1$ minimal.
From then on, the general situation applies : Suppose at level n we have a solution vector $A_n(z)$ and an auxiliary vector which is the solution vector of level m. The corresponding residuals are R_n and R_m. Hence we have the following situation at the end of level n :

$$M_n = \begin{matrix} [& A_m(z) & \mid & A_n(z) &] \\ & (\alpha_m + n - m) & & (\alpha_n) & \end{matrix} ,$$

$RM_n = [R_m \mid R_n]$.

To generate the solution of level $n+1$ we have to see if R_n can be written as a linear combination of R_m. This is always possible since these are both scalars. Next we should see if we have to replace the auxiliary vector or not. Three situations can occur

(1) $R_n = 0$. In this case, the auxiliary vector does not appear actively and we don't replace it. Hence $A_{n+1}(z) = A_n(z)$ and

$$M_{n+1} = \begin{bmatrix} A_m(z) & | & A_n(z) \\ (\alpha_m+n-m+1) & & (\alpha_n) \end{bmatrix},$$

$$RM_{n+1} = [R_m \mid R_{n+1} = R_n].$$

(2) $R_n \neq 0$ and $\alpha_n \geq \alpha_m+n-m$.
If $R_n \neq 0$, then $A_m(z)$ appears actively in the linear combination

$$A_{n+1}(z) = A_n(z) - (R_n/R_m)A_m(z)\, z^{\alpha_n-(\alpha_m+n-m)}$$

but its potential degree is not greater than the (potential) degree of $A_n(z)$. Hence the auxiliary vector is not replaced.

$$M_{n+1} = \begin{bmatrix} A_m(z) & | & A_{n+1}(z) \\ (\alpha_n+n-m+1) & & (\alpha_{n+1}=\alpha_n) \end{bmatrix},$$

$$RM_{n+1} = [R_m \mid R_{n+1}].$$

(3) $R_n \neq 0$ and $\alpha_n < \alpha_m+n-m$.
Now $A_m(z)$ is active in the linear combination

$$A_{n+1}(z) = z^{(\alpha_m+n-m)-\alpha_n}A_n(z) - (R_n/R_m)A_m(z)$$

but its potential degree is greater than the (potential) degree of A_n. Hence we replace the auxiliary vector by A_n.

$$M_{n+1} = \begin{bmatrix} A_n(z) & | & A_{n+1}(z) \\ (\alpha_n) & & (\alpha_{n+1}=\alpha_m+n-m) \end{bmatrix},$$

$$RM_{n+1} = [R_n \mid R_{n+1}].$$

We can now proceed to level $n+2$ etc.

Note that if we set $Q_n(z) = z^{\alpha_n}A_n(z^{-1})$, then in case (2) and (3), the updating formula becomes

$$Q_{n+1}(z) = Q_n(z) - \frac{R_m}{R_n}Q_m(z)z^{n-m}.$$

We can now give a theorem that relates the degrees of the solution vectors for different levels.

Theorem *Suppose $T(z)$ is as described above and suppose the degree of the solution vector of level n is denoted as α_n, $n = 1,2,\ldots$. Suppose also that the auxiliary vector at the end of level n was the solution vector of level m. Thus having a potential degree α_m+n-m. Then*

$$\alpha_n = \alpha_{m+1} = m-\alpha_m. \tag{16}$$

Proof Let s have the meaning of the index of the first nonzero coefficient in $T(z)$ as before. Then for level $s+1$, we have $\alpha_{s+1} = s-\alpha_s = s$, since $\alpha_s = 0$. Hence (16) is valid.
For the further levels we know that as long as case (1) or case (2) above occur, α_{n+1} will be equal to α_n and the auxiliary vector is left untouched. I.e. m and α_m remain unchanged. Hence (16) is still true for the next level. We only have to check what happens when the auxiliary vector is replaced, which is in case (3).
Suppose case (3) applies for level m and the previous time it was applicable was at level $m' < m$. Suppose for the induction step that $\alpha_m = \alpha_{m'+1} = m'-\alpha_{m'}$. We have to prove that

$\alpha_{m+1} = m-\alpha_m$. By construction of the algorithm, $\alpha_{m+1} = \alpha_{m'} + m-m'$, which is, by the induction hypothesis, equal to $\alpha_{m+1} = m-\alpha_{m'+1} = m-\alpha_m$.

□

Corollary *The test whether we are in case (2) or case (3), i.e.*

$\alpha_n \geqslant (<) \alpha_m+n-m$

is equivalent with the simpler one :

$2\alpha_n \geqslant (<) n.$

□

If we transfer this scalar algorithm into a Padé language, we find the following :
Given $F_0 + F_1z + F_2z^2 + \cdots \in K[[z]]$ with $F_0 = F_1 = \cdots = F_{s-2} = 0$, $F_{s-1} \neq 0$. Set $Q_n(z) = z^{\alpha_n} A_n(z^{-1}) = \sum_{k=0}^{\alpha_n} q_{kn} z^k$, with $A_n(z)$ as before, then the polynomials $Q_n(z)$ are generated by the following algorithm

C.1 $m = 0, Q_m(z) = 0, \alpha_m = 0, R_m = 1$

C.2 $n = 1, Q_n(z) = 1, \alpha_n = 0$

C.3 **for** $n = 1,2,3,...$

C.3.1 $R_n = \sum_{k=0}^{\alpha_n} F_{n-1-k}\ q_{kn}$

C.3.2 **if** $R_n = 0$ **then**

C.3.2.1 $Q_{n+1} = Q_n ; \alpha_{n+1} = \alpha_n$

else

C.3.2.2.1 $Q_{n+1} = Q_n - \frac{R_n}{R_m} z^{n-m} Q_m$

C.3.2.2.2 **if** $2\alpha_n \geqslant n$ **then**

C.3.2.2.2.1 $\alpha_{n+1} = \alpha_n$

else

C.3.2.2.2.2 $\alpha_{n+1} = \alpha_m+n-m = n-\alpha_n ; m = n$

This is almost the algorithm as given in the paper of Massey [14] the only difference is the initialization $Q_0 = 0$. This results in $Q_{s+1}(z) = 1$ as in the previous algorithm. Massey uses the initialization $Q_0 = 1$ which gives $Q_{s+1}(z) = 1 - F_{s-1}z^s$ but, because $A_{s+1}(z)$ is an arbitrary polynomial of degree s, both give a correct result.
$Q_{n+1}(z)$ is the denominator of an $O_+(z^n)$ approximant for the series $F(z)$. The corresponding numerator can be found by taking the first $\alpha_{n+1}-1$ terms in the product $F(z)Q_{n+1(z)}$. Another way to compute it is by the same updating formulas as for $Q_n(z)$ with initial conditions $P_0 = -z^{-1}$ and $P_1 = 0$.
To make clear that the algorithm need not produce a Padé approximant, take the simple case $F(z) = F_0 + F_1z + \cdots$ with $F_0 \neq 0$ and shift value $k = 0$. The algorithm gives with inital condition $Q_0 = 0$ as a result for level 1 : $Q_1 = 1$ and $P_1 = 0$ and for level 2 : $Q_2 = 1$, $\alpha_2 = 1$

and $P_2 = F_0$ which is a [0/0] Padé approximant of order 1, but with initial condition $Q_0 = 1$, we find $Q_2 = 1 - F_0 z$ and $P_2 = F_0$ which gives an approximant of order 1 which may not be a Padé approximant if $F_1 \neq (F_0)^2$ because $\partial Q_2 + \partial P_2 + 1 = 2 >$ the order 1.
If we want to construct by the algorithm elements that are in the Padé table, we should introduce some extra conditions. The only thing we can say is that whenever case (3) applies, i.e. $R_n \neq 0$ and $\alpha_n < \alpha_m + n - m$, we have a Padé approximant.

Theorem *If* $2\alpha_n < n$ *and* $n \geqslant s$ *in algorithm C, then* $P_n(z)/Q_n(z)$ *is a Padé approximant of* $F(z)$.

Proof Because $\partial Q_n(z) \leqslant \alpha_n$ and $\partial P(z) \leqslant \alpha_n - 1$ we have

$$\partial Q_n + \partial P_n + 1 \leqslant 2\alpha_n \leqslant n-1 \leqslant \text{the order of approximation} = n.$$

□

A formulation of algorithm C for general shift parameter k and with built in restrictions to generate certain Padé approximants is given in [7].

20. The matching of time moments

It is well known that the initial Markov parameters of a system give most adequate information about the transient behaviour of the system, while time moments will be better to model the steady state behaviour. If $T(z)$ is the transfer function of a continuous time system, then its time moments are defined as

$$m_k = (-1)^k \left. \frac{d^k T(z)}{dz^k} \right|_{z=0}, \quad k = 0,1,2,\ldots$$

and if $T(z)$ is the transfer function of a discrete time system, then its time moments are

$$m_k = (-1)^k \left. \left(z \frac{d}{dz} \right)^k T(z) \right|_{z=1}, \quad k = 0,1,2,\ldots .$$

Thus matching the first N time moments is equivalent with matching the first N coefficients T_k in

$$T(z) = \sum_{k=0}^{\infty} T_k z^k \quad \text{for continuous time systems}$$

and with matching the first N coefficients T_k in

$$T(z) = \sum_{k=0}^{\infty} T_k (z-1)^k \quad \text{for discrete time systems.}$$

We only give our problem formulation for the continuous time case, since the discrete time case can be brought in the same form with a transformation $\zeta = z - 1$. We shall call it the *minimal partial k-realization problem for* $T(z) \in K^{p \times m}[[z]]$ or the *minimal partial k-realization problem around* $z = 0$.

Definition The rational matrix function $D(z)B(z)^{-1}$ with $D(z) \in K^{p \times m}[z]$ and $B(z) \in K^{m \times m}[z]$ is called an *N-th order partial k-realization* of $T(z) \in K^{p \times m}[[z]]$ iff

1. $D(z)B(z)^{-1} = T_0 + T_1 z + \cdots + T_{N-1} z^{N-1} + T'_N z^N + \cdots$
2. $D(z)B(z)^{-1}$ expanded in decreasing powers of z is $O_-(z^k)$

for any $k \in \mathbb{Z}$.

It is called *minimal* if $\det B(z)$ has the smallest possible degree.

Notes :

1. Again the classical situation is where $k = -1$, i.e. where $D(z)B(z)^{-1}$ is strictly proper.
2. As for the minimal partial k-realization problem around $z = \infty$ where Markov parameters are matched, it is sufficient to construct a solution with a column reduced $B(z)$, since all other solutions can be derived from it with unimodular right multiplications.
3. If $b_i(z)$ denotes the i-th column of $B(z)$ and $d_i(z)$ denotes the i-th column of $D(z)$, then the minimality condition can be expressed as :

   $$\sum_{i=1}^{m} \partial b_i(z) \quad \text{minimal}$$

   and condition 2 in the definition above corresponds to

   $$\partial d_i(z) \leqslant k + \partial b_i(z), \quad i = 1,2, \ldots, m$$

   if $B(z)$ is column reduced.

We have the following theorem

Theorem *If $D(z)B(z)^{-1}$ is an N-th order minimal partial k-realization of $T(z) = \sum_{i=0}^{\infty} T_i z^i \in K^{p\times m}[[z]]$, then $B(0)$ is nonsingular.*

Proof If $N = 0$ then $B(z) \equiv B(0)$ and therefore it should be invertible, (i.e. nonsingular) otherwise the definition would make no sense. If $N > 0$, then $T_0 B(0) = D(0)$. Suppose there is a solution with $B_0 = B(0)$ singular. Then there exists a regular matrix R such that $B_0 R = [B_0' 0]$ and thus also $D(0)R = T_0 B_0 R = [D_0' 0]$ with $B_0' \in K^{m\times r}$ and $D_0' \in K^{p\times r}$ with $r < m$. Thus because $\begin{bmatrix} B(0) \\ D(0) \end{bmatrix} R = \begin{bmatrix} B_0' & 0 \\ D_0' & 0 \end{bmatrix}$ has a rank at most $r < m$. Thus $\begin{bmatrix} B(z) \\ D(z) \end{bmatrix}$ is not of full rank for all z and $B(z)$ and $D(z)$ have therefore a common right divisor. But this contradicts the fact that deg det $B(z)$ is minimal.

□

As for the minimal partial k-realization around $z = \infty$, we shall now reformulate our problem around $z = 0$ in terms of the columns of $B(z)$. Therefore we need the following

Definition A polynomial vector $b(z) \in K^m[z]$ of degree β is said to have a k-order n with respect to $T(z) \in K^{p\times m}[[z]]$ iff $n \geqslant -k$, $n \geqslant \beta+1$ and

$$T(z)b(z) = d(z) + qz^{k+n} + q'z^{k+n+1} + \cdots \quad \text{with } 0 \neq q \in K^p$$

and $d(z)$ a polynomial vector $\in K^p[z]$ with $\partial d(z) \leqslant k+\beta$. In analogy with the theorem of section 3 for the problem around $z = \infty$, we now have

Theorem *$D(z)B(z)^{-1}$ with $B(z)$ column reduced and $B(0)$ regular is an N-th order partial k-realization of $T(z) \in K^{p\times m}[[z]]$ iff the columns $b_i(z)$ of $B(z)$ have k-order at least $n = N-k$ with respect to $T(z)$ and the i-th column of $D(z)$ is $d_i(z)$ defined by*

$$T(z)b_i(z) = d_i(z) + O_+(z^{k+n})$$

i.e.

$$d_i(z) = T(z)b_i(z) (\text{mod } z^{k+n})$$

Proof The proof is analogous to the one for the case around $z = \infty$.

□

This theorem shows that we should construct polynomial vectors $b(z) = \sum_{i=0}^{\beta} b_i z^i \in K^m[z]$, of degree β ($b_\beta \neq 0$) with k-order at least $n = N-k$. Such a vector should satisfy the block Hankel system

$$\begin{bmatrix} T_{k+1} & \cdots & T_{k+\beta+1} \\ \vdots & & \vdots \\ T_{k+n-\beta} & \cdots & T_{k+n} \end{bmatrix} \begin{bmatrix} b_\beta \\ \vdots \\ b_0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r \end{bmatrix} \quad \text{with } r \in K^p \text{ possibly nonzero.} \tag{17}$$

Formally this system is the same as for the problem around $z = \infty$. Indeed (17) is transformed into (2) if we write M_{i+1} instead of T_i and $a(z)$ instead of $z^\beta b(z^{-1})$. Let us denote the system (17) as $(n,\beta)_k^+$. A superscript + is added, as it will be in other notation, to refer to the fact that $T(z)$ is now a power series in nonnegative powers of z. It is easily verified that table 1 can be translated into table 13.

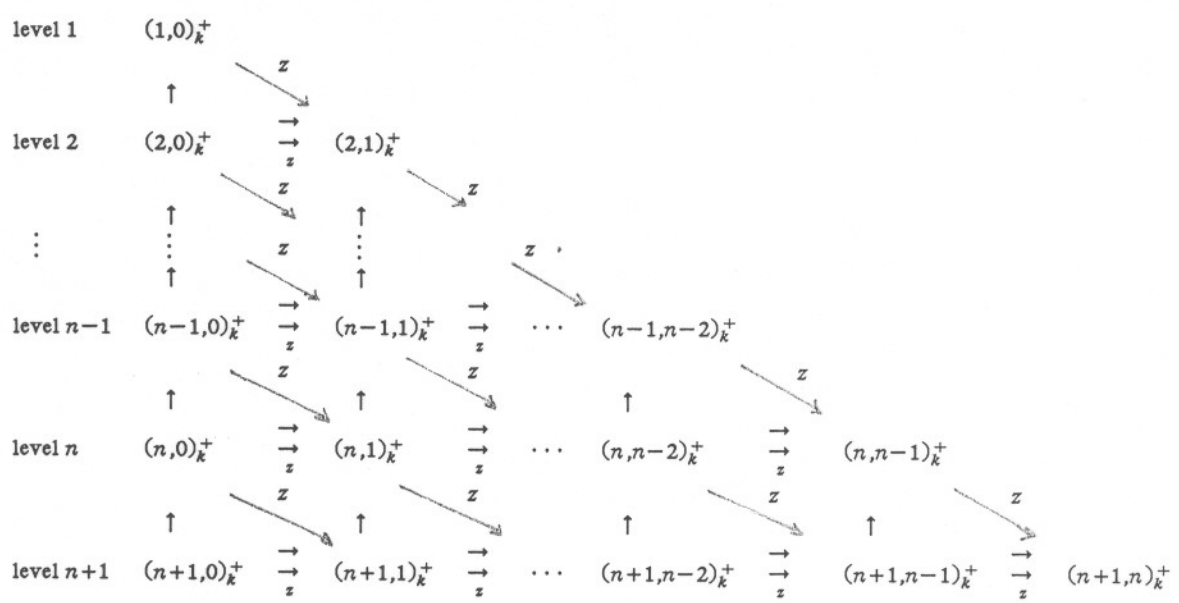


Table 13

Because this table is similar to table 1, we can easily make an adaptation of algorithm A to generate vectors $b_1^*(z), \ldots, b_m^*(z)$ of minimal degree that solve the N-th order partial k-realization of $T(z) \in K^{p\times m}[[z]]$ such that $B(0) = [b_1^*(0) \cdots b_m^*(0)]$ is nonsingular. However because the $b_i^*(z)$, are essentially found as $b_i^*(z) = z^{\alpha_i} a_i(1/z)$ if $a_i(z) \in [n,\alpha_i]_k$ by algorithm A for the data $T_i = M_{i+1}$, $b_i^*(z)$ may not be of strict degree α_i, and hence, $B(z) = [b_1^*(z), \ldots, b_m^*(z)]$ will not be column reduced and therefore nothing can be said about the minimality of the solution. Thus we have to build in an extra step in the algorithm to keep linear independent highest degree coefficients.
The systems we have to consider for level n are

$$\begin{bmatrix} T_{k+1} & \cdots & T_{k+\beta_i+1} \\ \vdots & & \vdots \\ & & \\ T_{k+n-\beta_i} & \vdots & T_{k+n} \end{bmatrix} \begin{bmatrix} b^n_{i,\beta_i} \\ \vdots \\ \\ b^n_{i,0} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r \end{bmatrix}, \quad i = 1,2,\ldots,m \text{ and } \beta_i \text{ as low as possible} \quad (18)$$

$B = [b^n_1 b^n_2 \cdots b^n_m]$. The highest degree coefficients $HB = [b^n_{1,\beta_1} \cdots b^n_{m,\beta_m}]$ are linearly independent and so are the lowest degree coefficients $LB = [b^n_{10} b^n_{20} \cdots b^n_{m0}]$.
Algorithm A can be used to construct solutions of the systems

$$\begin{bmatrix} T_{k+1} & \cdots & T_{k+\alpha_i+1} \\ \vdots & & \vdots \\ & & \\ T_{k+n+\alpha_i} & \cdots & T_{k+n} \end{bmatrix} \begin{bmatrix} a^n_{i0} \\ \vdots \\ \\ a^n_{i,\alpha_i} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r \end{bmatrix}, \quad i = 1,2,\ldots,m \text{ and } \alpha_i \text{ as low as possible .} \quad (19)$$

$A = [a^n_1 a^n_2 \cdots a^n_m]$. The highest degree coefficients $HA = [a^n_{1\alpha_1} \cdots a^n_{m\alpha_m}]$ are linearly independent. Now consider the set

$$B^* = [b^{n*}_1 \cdots b^{n*}_m]$$

with $b^{n*}_i = z^{\alpha_i} a^n_i(1/z)$. These solve systems of the form (18), but their highest degree coefficients

$$HB^* = [b^{n*}_{1\alpha_1} \cdots b^{n*}_{m\alpha_m}] = [a^n_{10} \cdots a^n_{m0}]$$

are not linearly independent. To make them linearly independent we shall add some *generating vectors* g to the b^{n*}_i. These generating vectors will be solutions for a system of the form (18) and will have linear independent highest degree coefficients. So if Hb^{n*}_i is not linearly independent of the previous Hb^n_j, $j < i$, we replace it by

$$b^n_i = b^{n*}_i + z^{\gamma_r} g_r . \quad (20)$$

The degree of b^n_i, which is $\beta_i \geqslant \alpha_i$, should be as low as possible. Hence we shall use some g_r with a potential degree which is as low as possible. If the g_j have potential degree $\sigma_1 \leqslant \sigma_2 \leqslant \cdots \leqslant \sigma_m$, then we should take r as low as possible. If we don't want to affect the linear independency of the Lb^{n*}_i, we should take $\gamma_r > 0$. Therefore we use for the generating vectors the relations of table 13 indicated by $\overset{z}{\rightarrow}$. We shall now see how we can construct the solution vectors of level $n+1$, i.e. $B_{n+1} = [b^{n+1}_1 \cdots b^{n+1}_m]$ of degrees $\beta^{n+1}_1, \beta^{n+1}_2, \ldots, \beta^{n+1}_m$, given the sets :
The A-set for level $n+1$:
$A_{n+1} = [a^{n+1}_1 \cdots a^{n+1}_m]$ of degrees $\alpha^{n+1}_1 \leqslant \alpha^{n+1}_2 \leqslant \cdots \leqslant \alpha^{n+1}_m$,
The G-set of generating vectors for level n :
$G_n = [g^n_1 \cdots g^n_m]$ of potential degrees (with respect to level n) $\sigma^n_1 \leqslant \sigma^n_2 \leqslant \cdots \leqslant \sigma^n_m$ and of strict degrees $\pi_1, \pi_2, \ldots, \pi_m$.
Initially (i.e. for the first level) the G set will be the same as the A set, i.e. $G = [e_1, e_2, \ldots, e_m]$ the unit vectors of K^m, all of degree zero.
We now construct for $i = 1,2,\ldots,m$ the vectors b^{n+1}_i as follows : At the start B_{n+1} is empty. If Hb^{n+1*}_i is linearly independent of the previous Hb^{n+1}_j, $j < i$ then add $b^{n+1}_i = b^{n+1*}_i$ to the set B_{n+1} of degree $\beta^{n+1}_i = \alpha^{n+1}_i$. If Hb^{n+1*}_i is linearly dependent of Hb^{n+1}_j, $j < i$, (This may be the case for $i=1$ if $Hb^{n+1*}_i = 0$.) then look for the smallest r such that Hg_r is linearly independent of Hb^{n+1}_j, $j < i$. There are two possibilities :

1. $\sigma_r + 1 > \alpha^{n+1}_i$. Then set

$b_i^{n+1} = b_i^{n+1*} + z^{\sigma_r+1-\pi_r} g_r^n$ of degree $\beta_i^{n+1} = \sigma_r + 1$.

The exponent $\gamma_r = \sigma_r + 1 - \pi_r > 0$ is the smallest possible one such that $z^{\gamma_r} g_r$ is a solution for a system on level $n+1$. Hence, the degree of b_i^{n+1} will be minimal among all solutions of level $n+1$ with Hb_i^{n+1} linearly independent of Hb_j^{n+1}, $j < i$.
Furthermore $\gamma_r > 0$ and hence $z^{\gamma_r} g_r$ will not contribute to the lowest degree coefficient of b_i^{n+1} and the linear independency of the lowest degree coefficients of the b_i^{n+1*} will be maintained.

2. $\sigma_r + 1 \leqslant \alpha_i^{n+1}$. Then set

$b_i^{n+1} = b_i^{n+1*} + z^{\alpha_i^{n+1}-\pi_r} g_r^n$ of degree $\beta_{i+1} = \alpha_i^{n+1}$.

Since $\pi_r \leqslant \sigma_r$ and $\sigma_r + 1 < \alpha_i^{n+1}$, we shall have that $\gamma_r = \alpha_i^{n+1} - \pi_r > 0$. The same justification as in the previous case can be applied.

Thus we get the following adaptation of algorithm A.

Add to the initialization step

A.1.4 Set $B := [1,0]_k^+ = A$. The vectors of B are of degree $\beta_i = \alpha_i = 0$, $i = 1,2,\ldots,m$.

A.1.5 The set G of generating vectors can be initialized as : $G := [1,0]_k^+$ with potential degrees $\sigma_i = 0$ and strict degrees $\pi_i = 0$, $i = 1,2,\ldots,m$.

We have to add after step A.2.1 and before step A.2.2 the following operations :

A'.2.1 **for** $i = 1,2,\ldots,m$
{B is the set of B-vectors which is empty at the beginning of this block and will contain the m B-vectors for level $n+1$ at the end}

A'.2.1.1 $b_i^{n+1*} = z^{\alpha_i} a_i^{n+1}(1/z)$ where a_i^{n+1} is the i-th vector in A' of degree α_i.

A'.2.1.2 **If** $Hb_i^{n+1*} \in$ span $\{Hb_j^{n+1}, j = 1,\ldots,i-1\}$ **then**

A'.2.1.2.1 Look for the first r such that $Hg_r^n \notin$ span $\{Hb_j^{n+1}, j = 1,\ldots,i-1\}$. It will always exist and r will be $\leqslant i$.

A'.2.1.2.2 **If** $\sigma_r + 1 > \alpha_i$ **then**

$$b_i^{n+1} = b_i^{n+1*} + z^{\sigma_r+1-\pi_r} g_r^n$$

of degree $\beta = \sigma_r + 1$.

A'.2.1.2.3 **If** $\sigma_r + 1 \leqslant \alpha_i$ **then**

$$b_i^{n+1} = b_i^{n+1*} + z^{\alpha_i-\pi_r} g_r^n$$

of degree $\beta = \alpha_i$.

A'.2.1.3 **If** $Hb_i^{n+1*} \notin$ span $\{Hb_j^{n+1}, j = 1,2,\ldots,i-1\}$ **then**

$b_i^{n+1} = b_i^{n+1*}$ of degree $\beta = \alpha_i$

A'2.1.4 Add b_i^{n+1} to the set B of degree β.

Step A.2.2 becomes :

A.2.2.1 Replace A by A'

A.2.2.2 Increase the potential degrees of the vectors in X and G with 1.

A.2.2.3 The new vectors g_i^{n+1} of G are the m vectors of the sets G and B^* of minimal potential degree such that $Hg_i^{n+1} \notin$ span$\{Hg_j^{n+1}, j = 1,\ldots,i-1\}$. These vectors are

ordered from low to high potential degree.

Example 9

We illustrate the previous modifications of the algorithm with the following example

$$T(z) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^0 + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z^1 + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 2 & 3 \\ 0 & 0 \end{bmatrix} z^3 + \cdots$$

We solve the problem for $k = -1$ (the classical case).

The initialization gives $B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and

$$G_1 = \begin{array}{c} \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\right] \\ (0)\ \ (0) \end{array}, \quad M_1 = \begin{array}{c} \left[\begin{array}{c|cc} & 1 & 0 \\ & 0 & 1 \end{array}\right] \\ \qquad (0)\ \ (0) \end{array},$$

$$RM_1 = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 1 \end{array}\right] \rightarrow A_2 = \begin{bmatrix} -1 & z \\ 1 & 0 \end{bmatrix}$$

$$B_2^* = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow B_2 = \begin{bmatrix} -1 & z+1 \\ 1 & 0 \end{bmatrix}, \quad G_2 = \begin{array}{c} \left[\begin{array}{cc} -1 & 1 \\ 1 & 0 \end{array}\right] \\ (0)\ \ (1) \end{array}.$$

For $n=2$ we get

$$M_2 = \begin{array}{c} \left[\begin{array}{c|cc} 1 & -1 & z \\ 0 & 1 & 0 \end{array}\right] \\ (1)\quad (0)\ (1) \end{array},$$

$$RM_2 = \left[\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right] \rightarrow A_3 = \begin{bmatrix} -1 & z \\ 1 & 0 \end{bmatrix}$$

$$B_3^* = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow B_3 = \begin{bmatrix} -1 & z^2+1 \\ 1 & 0 \end{bmatrix}, \quad G_3 = \begin{array}{c} \left[\begin{array}{cc} -1 & 1 \\ 1 & 0 \end{array}\right] \\ (0)\ \ (2) \end{array}.$$

For $n=3$:

$$M_3 = \begin{array}{c} \left[\begin{array}{c|cc} 1 & -1 & z \\ 0 & 1 & 0 \end{array}\right] \\ (2)\quad (0)\ (1) \end{array},$$

$$RM_3 = \left[\begin{array}{c|cc} 0 & 0 & 1 \\ 1 & 0 & 1 \end{array}\right] \rightarrow A_4 = \begin{bmatrix} -1 & z^3 \\ 1 & 0 \end{bmatrix}$$

$$B_4^* = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow B_4 = \begin{bmatrix} -1 & z^3+1 \\ 1 & 0 \end{bmatrix}, \; G_4 = \begin{matrix} \left[\begin{matrix} -1 & 1 \\ 1 & 0 \end{matrix}\right] \\ (0)\ (3) \end{matrix}.$$

For $n=4$:

$$M_4 = \begin{matrix} \left[\begin{matrix} z & 1 \\ 1 & 0 \end{matrix} \middle| \begin{matrix} -1 & z^3 \\ 1 & 0 \end{matrix}\right] \\ (2)\ (3) \quad (0)\ (3) \end{matrix},$$

$$RM_4 = \left[\begin{array}{cc|cc} 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 \end{array}\right] \rightarrow A_5 = \begin{bmatrix} 1-z^2-z^3 & 3z^3 \\ z^3 & -2z^3 \end{bmatrix}$$

$$B_5^* = \begin{bmatrix} z^3-z-1 & 3 \\ 1 & -2 \end{bmatrix} \rightarrow B_5 = \begin{bmatrix} z^3-z-1 & 3-z^3 \\ 1 & -2+z^3 \end{bmatrix}, \; G_5 = \begin{matrix} \left[\begin{matrix} -1 & z^3-z-1 \\ 1 & 1 \end{matrix}\right] \\ (1) \quad (3) \end{matrix}.$$

The following table 14 gives the B^* sets on the first row and the B sets below it for the first 5 levels.

degree → / level ↓	0	1	2	3
1	$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}0\\1\end{bmatrix}$			
2	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}z+1\\0\end{bmatrix}$		
3	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}z^2+1\\0\end{bmatrix}$	
4	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$			$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}z^3+1\\0\end{bmatrix}$
5				$\begin{bmatrix}z^3-z-1\\1\end{bmatrix}$ $\begin{bmatrix}3\\-2\end{bmatrix}$ $\begin{bmatrix}z^3-z-1\\1\end{bmatrix}$ $\begin{bmatrix}3-z^3\\-2+z^3\end{bmatrix}$

Table 14

The solutions are then

$$n=1: \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1}$$

$$n=2: \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & z+1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$n=3: \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & z^2+1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$n=4: \begin{bmatrix} 0 & z^2 \\ 0 & z^2+1 \end{bmatrix} \begin{bmatrix} -1 & z^3+1 \\ 1 & 0 \end{bmatrix}^{-1}$$

$$n=5: \begin{bmatrix} 0 & z^2 \\ -z & z^2+1 \end{bmatrix} \begin{bmatrix} z^3-z-1 & 3-z^3 \\ 1 & -2+z^3 \end{bmatrix}^{-1}$$

To compute solutions for other values of the shift parameter k, we can use as before algorithm B to derive the solution vectors for shift value $k+1$ and level n once algorithm A has computed the solution vectors for shift k and level $n+1$. To find the solutions for the partial $k+1$ realization for the series around $z=0$, we have to adapt algorithm B exactly as we adapted algorithm A above. Without giving further details you can derive easily the following tables for shift value $k = 0,1,2,3$ as derived from table 14. Only B^* and B are given. The results are given in tables 15-18.

degree → / level ↓	0	1	2	3
1	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$			
2	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$			
3	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$		$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}z^2+1\\0\end{bmatrix}$	
4			$\begin{bmatrix}3\\-2\end{bmatrix}$ $\begin{bmatrix}3-z^2\\-2+z^2\end{bmatrix}$	$\begin{bmatrix}z^3-z-1\\1\end{bmatrix}$ $\begin{bmatrix}z^3-z-1\\1\end{bmatrix}$

Table 15: Shift value $k = 0$

degree → / level ↓	0	1	2
1	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$		
2	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}z+1\\0\end{bmatrix}$	
3		$\begin{bmatrix}3\\-2\end{bmatrix}$ $\begin{bmatrix}3-z\\-2+z\end{bmatrix}$	$\begin{bmatrix}-z-1\\1\end{bmatrix}$ $\begin{bmatrix}z^2-z-1\\1\end{bmatrix}$

Table 16: Shift value $k = 1$

degree → / level ↓	0	1
1	$\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$ $\begin{bmatrix}-1\\1\end{bmatrix}$ $\begin{bmatrix}1\\0\end{bmatrix}$	
2	$\begin{bmatrix}3\\-2\end{bmatrix}$ $\begin{bmatrix}3\\-2\end{bmatrix}$	$\begin{bmatrix}-z-1\\1\end{bmatrix}$ $\begin{bmatrix}-z-1\\1\end{bmatrix}$

Table 17: Shift value $k = 2$

degree → / level ↓	0
1	$\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Table 18: Shift value $k = 3$

The previous computations can be put in the diagram of table 19. In this table k represents the shift value and n is the number of time moments used.

−2	−1	0	1	2	3	4	5	k
	o →	1						
	o →	2	o →	o				
	o →	3	o →	o →	o			
	o →	4	o →	o →	o →	o		
		5						
		6						
		n						

Table 19

We can first compute the solution vectors for the lowest value of k and then determine the solution vectors for increasing values of k.

21. Connection with Padé approximation

To make the connection with Padé approximation, we consider the scalar case, i.e. $p = m = 1$. Given the scalar formal power series

$$T(z) = T_0 + T_1 z + T_2 z^2 + \cdots \quad , T_k \in K,$$

we shall find by the previous algorithm two polynomials $D(z)$ and $B(z) \in K[z]$ such that $B(0) \neq 0$ and degree $D(z) \leqslant$ degree of $B(z)$ + shift parameter k. They are such that

(a) $D(z)/B(z) = T(z) + O_+(z^N)$, i.e. of order $\geqslant N$

(b) the degree of $B(z)$ is minimal.

We call the degree of $B(z)$, the degree of the approximant.

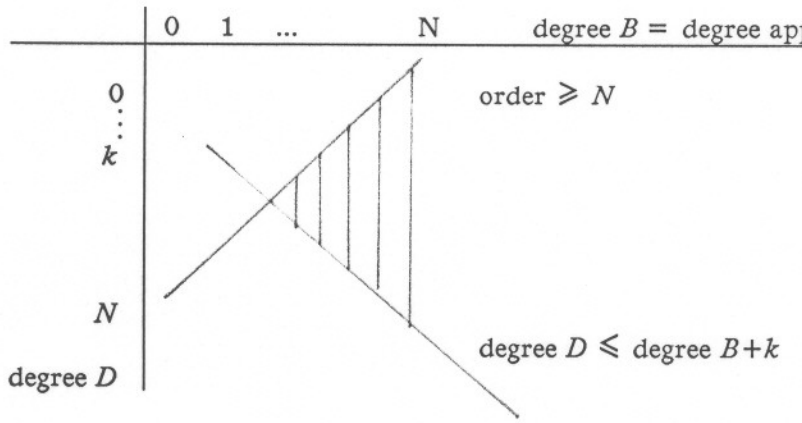


Table 20

In a Padé table we have the situation as scetched in Table 20.

If the degree of B has to be as low as possible we have to select the left top element in the shaded region whenever there is one. If there is not a unique top element, there will be two elements of the same degree : one of order N and one of order $N+1$. In that case we can select one of both e.g. the one with the lowest possible degree for D or the other one. Of course, all these elements should exist which is the case in a normal Padé table. Take the example of a normal Padé table with $k < 0$ as given in table 21.

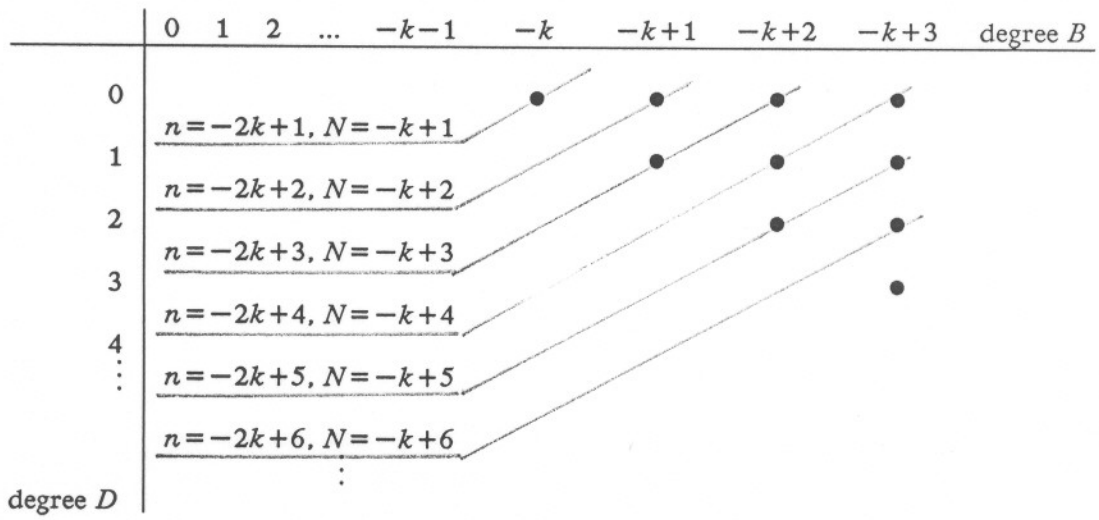


Table 21

At level $2k+1$ we get $[0/-k]$ of degree $-k$ and order $N = -k+1$.
At level $-2k+2n$ we get $[n-1/-k+n]$ and $[n/-k+n]$ of degree $-k+n$ and order $N = -k+2n$.
At level $-2k+2n+1$ we get $[n/-k+n]$ of degree $-k+n$ and order $N = -k+2n+1$.
To make the solutions unique we can impose the condition that the degree of the numerator is minimal. Then we get the staircase $[0/-k]$, $[0/-k+1]$, $[1/-k+1]$, $[1,-k+2]$, $[2/-k+2]$.

As an example for a Padé table with a singular block consider the following one for $k=2$. In

the table we indicate the degrees of the Padé approximants that exist. With the technique illustrated in table 21 we select the following solutions displayed in table 22.

	0	1	2	3	4	5	6	... ∂ B	level	order	degree	solution
0	0	1	2	3	4	5			1	3	0	[2/0]
1	0	1	2	3	4	5			2	4	1	[2/1]
2	0	1	1	1	4	5			3	5	1	[2/1]
3		1	1		4	5			4	6	1	[2/1]
4		1			4	5			5	7	3	[5/3]
5				3	4	5			6	8	3	[5/3]
⋮					4	5			7	9	3	[5/3]
									8	10	4	[5/4] or [6/4]
									9	11	4	[6/4]
∂ D									⋮	⋮	⋮	⋮

singular block

Table 22

If there are different solutions for a denominator polynomial, then every linear combination is again a solution on condition that the highest and lowest degree coefficient is not zero.

22. Fitting time moments : another approach

As we have seen in earlier sections, the fitting of time moments was a bit more complicated than the fitting of Markov parameters. The reason for this is that the column-reducedness condition for the denominator is in fact a condition at $z=\infty$, while time moments are conditions at $z=0$.

Algorithm A generated solution vectors a which could be used to construct a basis for systems of the form (17). What we actually needed were reciprocals of these a vectors, i.e. the $b^*(z) = z^{\alpha}a(1/z)$. It was because of the latter operation that we had to add some extra step to algorithm A. Mathematically speaking, it is possible to transform the problem for time moments into the same problem as for Markov parameters but the extra condition of linear independency of the constant terms of the solution vectors should be added. Thus the algorithm should take care of this. This transformation goes as follows : Invert the block row and block column order of the system (17) and you get :

$$\begin{bmatrix} T_{k+n} & \cdots & T_{k+n-\beta} \\ \vdots & & \vdots \\ T_{k+\beta+1} & \cdots & T_{k+1} \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_\beta \end{bmatrix} = \begin{bmatrix} r \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and by shifting this over one block row you get

$$\begin{bmatrix} M_{k+2} & \cdots & M_{k+2+\beta} \\ \vdots & & \vdots \\ M_{n+k+1-\beta} & \cdots & M_{n+k+1} \end{bmatrix} \begin{bmatrix} b_0 \\ \vdots \\ b_\beta \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ s \end{bmatrix}$$

where we have renamed T_l as $M_{2k+n-l+1}$. Now the problem is formulated in exactly the same way as for the problem solved by algorithm A. However the interpretation of the method of solution is different. Indeed, in algorithm A, the solution vectors for level n are generated from the solution vectors of previous levels. The systems that are solved at level $n-l$ however, when written in terms of time moments, are the systems one has to solve to fit time

moments for level $n-l$ but with a shift parameter value which is $k+l$ if k is the shift parameter for the system of level n. Thus the value of the shift parameter decreases as the levels increase. Thus one solves the triangular family of systems of table 23.

$(1,0)^+_{k+n-1}$				
$(2,0)^+_{k+n-2}$	$(2,1)^+_{k+n-2}$			
$\vdots$				
$(n-1,0)^+_{k+1}$	$(n-1,1)^+_{k+1}$	$\cdots$	$(n-1,n-2)^+_{k+1}$	
$(n,0)^+_k$	$(n,1)^+_k$	$\cdots$	$(n,n-2)^+_k$	$(n,n-1)^+_k$

Table 23

This is clearly not very interesting from a system theoretic point of view, but when regarded as an approximation problem for matrix sequences, it is a valuable alternative.

23. Fitting both time moments and Markov parameters

Now that we know how to approximate Markov parameters and how to approximate time moments, it will be an easy task to combine these two techniques to approximate say $N-1$ Markov parameters and M time moments. To obtain this we shall have to find denominator columns that should have k-order $n = N-k-1$, for the series

$$T(z) = M_1 z^{-1} + M_2 z^{-2} + \cdots \quad \in K^{p\times m}[[z^{-1}]].$$

Hence, its coefficients satisfy the following system

$$\begin{bmatrix} & & M_1 \\ & & \vdots \\ M_1 & \cdots & M_{1+\alpha} \\ M_2 & \cdots & M_{2+\alpha} \\ \vdots & & \vdots \\ M_{k+1} & \cdots & M_{k+1+\alpha} \\ M_{k+2} & \cdots & M_{k+2+\alpha} \\ \vdots & & \vdots \\ M_{k+n-\alpha} & \cdots & M_{k+n} \\ M_{k+n+1-\alpha} & \cdots & M_{k+n+1} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_\alpha \end{bmatrix} = \begin{bmatrix} c_{\alpha+k} \\ \vdots \\ \\ \\ \\ c_0 \\ 0 \\ \vdots \\ 0 \\ r \end{bmatrix} . \tag{21}$$

The coefficients c_i that appear in the right hand side are the coefficients of the corresponding numerator.

This denominator column should also have k-order $q = M-k$ for the series

$$T_0 + T_1 z + T_2 z^2 + \cdots \quad \in K^{p\times m}[[z]].$$

Therefore, it must also satisfy

$$\begin{bmatrix} T_{k+q} & \cdots & T_{k+q-\alpha} \\ T_{k+q-1} & \cdots & T_{k+q-\alpha-1} \\ \vdots & & \vdots \\ T_{\alpha+k+1} & \cdots & T_{k+1} \\ T_{\alpha+k} & \cdots & T_k \\ \vdots & & \vdots \\ T_\alpha & \cdots & T_0 \\ \vdots & & \\ T_0 & & \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_\alpha \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ \vdots \\ 0 \\ c_{\alpha+k} \\ \vdots \\ \\ \\ c_0 \end{bmatrix} . \tag{22}$$

Thus, if we subtract (22) from (21) we get :

$$\begin{bmatrix} M_{-k-q+1} & \cdots & M_{-k-q+\alpha+1} \\ \vdots & & \vdots \\ M_{k+n-\alpha} & \cdots & M_{k+n} \\ M_{k+n+1-\alpha} & \cdots & M_{k+n+1} \end{bmatrix} \begin{bmatrix} a_0 \\ \vdots \\ a_\alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r \end{bmatrix}$$

where we have set $M_{-i} = -T_i$, $i \leqslant 0$. The matrix denominator which consists of all these columns should be column reduced and of minimal degree. If some time moments are involved, the lowest degree coefficient of this matrix polynomial should be nonsingular. Thus it is expected that algorithms similar to the one which solved the time moment problem can be used. Indeed, if the number of Markov parameters one wants to use is fixed, then $k+n+1$ is fixed and in that case we can use algorithm A as adapted for time moments to fit $M_{k+n+1}, M_{k+n}, \ldots, M_1, M_0, \ldots, M_{-k-q+1}, \ldots$. If the number of time moments which one wants to use is fixed, a similar algorithm can be designed to match $M_{-k-q+1}, M_{-k-q+2}, \ldots, M_0, M_1, \ldots, M_{k+n+1}, \ldots$ since now $-k-q+1$ is fixed. The development of such an algorithm is along the same lines as for the time moment adaptation of section 22.
We have to solve

$$\begin{bmatrix} M_{-k-q+1} & \cdots & M_{-k-q+\alpha+1} \\ \vdots & & \vdots \\ M_{k+n+1-\alpha} & \cdots & M_{k+n+1} \end{bmatrix} \begin{bmatrix} b^n_{i,0} \\ \vdots \\ b^n_{i,\beta_i} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r_i \end{bmatrix} , i = 1,2, \ldots, m \tag{23}$$

and β_i as low as possible. $B = [b^n_1, b^n_2, \ldots, b^n_m]$.
The highest degree coefficients $HB = [b^n_{1,\beta_1}, \ldots, b^n_{m,\beta_m}]$ should be linearly independent.
Algorithm A can be used to construct solutions of the systems

$$\begin{bmatrix} M_{-k-q+1} & \cdots & M_{-k-q+\alpha+1} \\ \vdots & & \vdots \\ M_{k+n+1-\alpha} & \cdots & M_{k+n+1} \end{bmatrix} \begin{bmatrix} a^n_{i,0} \\ \vdots \\ a^n_{i,\alpha_i} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r_i \end{bmatrix} , i = 1,2, \ldots, m \tag{24}$$

with α_i as low as possible. $A = [a^n_1, \ldots, a^n_m]$. The highest degree coefficients $HA = [a^n_{1,\alpha_1}, \ldots, a^n_{m,\alpha_m}]$ are linearly independent. However the lowest degree coefficients may not be linearly independent. Therefore we add some *generating vectors of the second kind* which we denote again as g^n_i since they play a role which is similar to the generating vectors g^n_i we used when matching time moments. They are solutions of the systems (23) that have linear independent lowest degree coefficients. So if La^n_i is not linearly independent of the previous coefficients Lb^n_j, $j < i$, we replace it by

$b_i^n = z^{\gamma_i} a_i^n + g_r$.

If the degree of b_i^n, which is $\beta_i^n \geqslant \alpha_i^n$ has to be kept as low as possible, then γ_i should be as low as possible, preferably zero. The value of γ_i is influenced by the potential degree of g_r. Thus if the generating vectors of the second kind have potential degrees $\sigma_1 \leqslant \sigma_2 \leqslant \cdots \leqslant \sigma_m$, then we should keep r as low as possible and use the relations indicated by $\rightarrow$ in table 1, otherwise the lowest degree coefficients of a_i^n would not be changed by adding g_r. Furthermore, we should not affect the linear independency of the highest degree coefficients of the a_i^n. Thus if π_r is the exact degree of g_r, we should have $\gamma_i + \alpha_i^n > \pi_r$.
Suppose we want to construct $B_{n+1} = [b_1^{n+1}, \ldots, b_m^{n+1}]$ of degrees $\beta_1^{n+1}, \ldots, \beta_m^{n+1}$, given the sets.
The A set for level $n+1$:
$A_{n+1} = [a_1^{n+1}, \ldots, a_m^{n+1}]$ of degrees $\alpha_1^{n+1}, \ldots, \alpha_m^{n+1}$, the G set of generating vectors of the second kind for level n. $G_n = [g_1^n, \ldots, g_m^n]$ of potential degree with respect to level n: $\sigma_1^n, \ldots, \sigma_m^n$ and of strict degrees $\pi_1, \ldots, \pi_m$.
Initially $G = [e_1, e_2, \ldots, e_m]$, the unit vectors of K^m, all of degree zero.
We now construct for $i = 1,2, \ldots, m$ the vectors b_i^{n+1} as follows : At the start B_{n+1} is empty. If La_i^{n+1} is linearly independent of the previous Lb_j^{n+1}, $j < i$, then $b_i^{n+1} = a_i^{n+1}$ is added to the set B_{n+1} of degree $\beta_i^{n+1} = \alpha_i^{n+1}$.
If La_i^{n+1} is linearly dependent of Lb_j^{n+1}, $j < i$, then look for the smallest r, such that Lg_r is linearly independent of Lb_j^{n+1}, $j < i$. There are the following possibilities :

1. $\sigma_r + 1 \leqslant \alpha_i^{n+1}$. Then set

   $b_i^{n+1} = a_i^{n+1} + g_r^n$ degree of $\beta_i^{n+1} = \alpha_i^{n+1}$.

   Since $\pi_r \leqslant \sigma_r < \alpha_i^{n+1}$, the highest degree coefficient of a_i^{n+1} is not affected.

2. $\sigma_r + 1 > \alpha_i^{n+1}$. Then set

   $b_i^{n+1} = z^{\gamma_i} a_i^{n+1} + g_r^n$ with $\gamma_i = \sigma_r + 1 - \alpha_i^{n+1}$ of degree $\beta_i^{n+1} = \sigma_r + 1$.

   Indeed, in this case it is impossible to add to a_i^{n+1} a vector which is in the solution space $S(n+1, \alpha_i^{n+1})_k$ with linear independent lowest degree coefficient. So we have to shift a_i^{n+1} to make it of a higher degree. The least possible shift is given by γ_i such that $z^{\gamma_i} a_i^{n+1}$ is of degree $\sigma_r + 1$. Again, since $\pi_r \leqslant \sigma_r < \beta_i^{n+1} = \sigma_r + 1$, the highest degree coefficient of a_i^{n+1} is maintained.

Thus algorithm A can be adapted as follows :

Change A.1.3 as follows :

A.1.3 The set of solution vectors $A = [1,0]_k = \{e_1, \ldots, e_m\}$ all of (potential) degree $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$.
The corresponding set of residuals $RA = M_{-M+k+2}$ (the m columns of M_{-M+k+2}).

Add the initialization step

A.1.4 Set $B := [1,0]_k = A$. The B-vectors of degree $\beta_i = \alpha_i = 0$, $i = 1,2, \ldots, m$.

A.1.5 The set G can be initialized as $G := [1,0]_k$ with potential degrees $\sigma_i = 0$ and strict degrees $\pi_i = 0$, $i = 1,2, \ldots, m$.

After step A.2.1 and before step A.2.2 add the following operations :

A'.2.1 for $i = 1,2, \ldots, m$
{B is the set of B-vectors which is empty at the beginning of this block and will contain the m B-vectors for level $n+1$ at the end}

A'.2.1.1 If $La_i^{n+1} \in \text{span}\{Lb_j^{n+1}, j=1,2,\ldots,i-1\}$ then

A'.2.1.1.1 Look for the first r such that

$$Lg_r^n \notin \text{span}\{Lb_j^{n+1}, j=1,2,\ldots,i-1\}.$$

It will always exist and r will be $\leqslant i$.

A'.2.1.1.2 If $\sigma_r+1 \leqslant \alpha_i$ then

$$b_i^{n+1} = a_i^{n+1} + g_r^n \text{ of degree } \beta' = \alpha_i.$$

A'.2.1.1.3 If $\sigma_r+1 > \alpha_i$ then

$$b_i^{n+1} = z^{\sigma_r+1-\alpha_i} a_i^{n+1} + g_r^n \text{ of degree } \beta' = \sigma_r+1$$

A'.2.1.2 If $La_i^{n+1} \notin \text{span}\{Lb_j^{n+1}, j=1,2,\ldots,i-1\}$ then

$$b_i^{n+1} = a_i^{n+1} \text{ of degree } \beta' = \alpha_i.$$

A'.2.1.3 Add b_i^{n+1} to the set B' of degree β'.

Step A.2.2 becomes :

A.2.2.1 Replace A by A'

A.2.2.2 Increase the potential degrees of the vectors in X and G with 1.

A.2.2.3 The new set G is formed by m vectors of the old G and the set A whose lowest degree coefficients are linearly independent and who are of lowest potential degree. These vectors are ordered according to increasing potential degree.

Example 10

$$T(z) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^0 + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z^1 + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} z^2 + \cdots \quad z \to 0$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^{-2} + \cdots \quad z \to \infty$$

If we fix $k = -1$ and $m = 4$, then we should apply the algorithm to the data

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

which gives :

$$A_1 = B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad G_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ (0) & (0) \end{bmatrix}$$

$$M_1 = \begin{bmatrix} | & 1 & 0 \\ | & 0 & 1 \\ & (0) & (0) \end{bmatrix}.$$

$$RM_1 = \left[\begin{array}{c|cc} & -1 & -1 \\ \hline & -1 & -1 \end{array}\right] \rightarrow A_2 = \begin{bmatrix} -1 & z \\ 1 & 0 \end{bmatrix} \rightarrow B_2 = \begin{bmatrix} -1 & z+1 \\ 1 & 0 \end{bmatrix}, \; G_2 = \begin{array}{c} \left[\begin{array}{cc} -1 & 1 \\ 1 & 0 \end{array}\right] \\ (0)\ (1) \end{array}$$

$$M_2 = \begin{array}{c} \left[\begin{array}{c|cc} 1 & -1 & z \\ 0 & 1 & 0 \end{array}\right] \\ (1)\quad (0)\ (1) \end{array},$$

$$RM_2 = \left[\begin{array}{c|cc} -1 & 0 & 0 \\ -1 & 0 & 0 \end{array}\right] \rightarrow A_3 = \begin{bmatrix} -1 & z \\ 1 & 0 \end{bmatrix} \rightarrow B_3 = \begin{bmatrix} -1 & z^2+1 \\ 1 & 0 \end{bmatrix}, \; G_3 = \begin{array}{c} \left[\begin{array}{cc} -1 & 1 \\ 1 & 0 \end{array}\right] \\ (0)\ (2) \end{array}$$

$$M_3 = \begin{array}{c} \left[\begin{array}{c|cc} 1 & -1 & z \\ 0 & 1 & 0 \end{array}\right] \\ (2)\quad (0)\ (1) \end{array},$$

$$RM_3 = \left[\begin{array}{c|cc} -1 & 0 & 0 \\ -1 & 0 & -1 \end{array}\right] \rightarrow A_4 = \begin{bmatrix} -1 & z^3 \\ 1 & 0 \end{bmatrix} \rightarrow B_4 = \begin{bmatrix} -1 & z^3+1 \\ 1 & 0 \end{bmatrix}, \; G_4 = \begin{array}{c} \left[\begin{array}{cc} -1 & 1 \\ 1 & 0 \end{array}\right] \\ (0)\ (3) \end{array}$$

$$M_4 = \begin{array}{c} \left[\begin{array}{cc|cc} z & 1 & -1 & z^3 \\ 0 & 0 & 1 & 0 \end{array}\right] \\ (2)\ (3)\quad (0)\ (3) \end{array},$$

$$RM_4 = \left[\begin{array}{cc|cc} 0 & -1 & -1 & 1 \\ -1 & -1 & -1 & 1 \end{array}\right] \rightarrow A_5 = \begin{bmatrix} -z^3-1 & 0 \\ z^3 & z^3 \end{bmatrix} \rightarrow B_5 = \begin{bmatrix} -z^3-1 & -1 \\ z^3 & z^3+1 \end{bmatrix},$$

$$G_5 = \begin{array}{c} \left[\begin{array}{cc} -1 & -z^3-1 \\ 1 & z^3 \end{array}\right] \\ (1)\quad (3) \end{array}$$

$$M_5 = \begin{array}{c} \left[\begin{array}{cc|cc} -1 & z & -z^3-1 & 0 \\ 1 & 0 & z^3 & z^3 \end{array}\right] \\ (1)\ (3)\quad\quad (3)\quad (3) \end{array}$$

$$RM_5 = \left[\begin{array}{cc|cc} -1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 \end{array}\right] \rightarrow A_6 = \begin{bmatrix} -z^3-1 & z \\ z^3 & z^3 \end{bmatrix} \rightarrow B_6 = \begin{bmatrix} -z^3-1 & z-1 \\ z^3 & z^3+1 \end{bmatrix},$$

$$G_6 = \begin{bmatrix} -1 & -z^3-1 \\ 1 & z^3 \\ (2) & (3) \end{bmatrix}$$

The numerator coefficients can be found from system (21) or (22) which gives :

$$D_6(z) = \begin{bmatrix} -z^2 & 0 \\ -z^2-1 & z \end{bmatrix} .$$

$D_6(z)B_6(z)^{-1}$ will now fit 3 time moments and 2 Markov parameters. Indeed,

$$D_6(z)B_6(z)^{-1} = \begin{bmatrix} -z^2 & 0 \\ -z^2-1 & z \end{bmatrix} \begin{bmatrix} -z^3-1 & z-1 \\ z^3 & z^3+1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} -z^2 & 0 \\ -z^2-1 & z \end{bmatrix} \begin{bmatrix} z^3+1 & -z+1 \\ -z^3 & -z^3-1 \end{bmatrix} / [-(z^3+1)^2 - z^4 + z^3]$$

$$= \begin{bmatrix} z^5+z^2 & -z^3+z^2 \\ z^5+z^4+z^3+z^2+1 & z^4-z^3+z^2+1 \end{bmatrix} / (z^6+z^4+z^3+1)$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} z^{-1} + \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^{-2} + \begin{bmatrix} 0 & -1 \\ 0 & -1 \end{bmatrix} z^{-3} + \cdots \qquad z \to \infty$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^0 + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} z^2 + \begin{bmatrix} 0 & -1 \\ 0 & -2 \end{bmatrix} z^3 + \cdots \qquad z \to 0.$$

Note

The idea of putting time moments and Markov parameters together into one block Hankel matrix has been used by Bistritz [4,5,6] who gives a state space algorithm to fit both time moments and Markov parameters. His algorithm may not be applicable if the approximant has a pole in $z = 0$, which means that the time moments for the model do not exist, and therefore they can't match the given time moments. In our formulation this corresponds to the condition that the lowest degree coefficient of the matrix polynomial denominator should be nonsingular. For example Bistritz's algorithm applied to the data

$$T(z) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} z^0 - \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} z^1 + \cdots$$

would generate the solution :

$$C(zI-A)^{-1}B$$

with $C = [0 \;\; 1]^T$, $B = [1 \;\; 1]$ and $A = 0$ which is the only possible solution and this gives the approximation

$$\begin{bmatrix} 0 & 0 \\ z^{-1} & z^{-1} \end{bmatrix}$$

which doesn't match the data. It is however expected that his algorithm can be adapted to take

care of this anomaly.

24. Conclusion

We have given a recursive description of a basis for the kernels of certain block Hankel matrices. This, together with the notion of minimal potential degree for these basis vectors has resulted in a recursive algorithm along the lines of Dickinson, Morf and Kailath for the solution of the minimal partial realization problem for matrix sequences. It is an extension of the Berlekamp-Massey algorithm to the MIMO case.
Its relation to this algorithm and to the Padé approximation problem is explained.
A modification is given for the algorithm so that it can be used to match time moments instead of Markov parameters and it can also be used or modified to match both of them.
Many examples are given to illustrate the algorithms.
We also gave a relation with the work of Graves-Morris on vector valued rational interpolation and show that a restructuring of the problem is not always allowed because this could result in a different problem.
Our approach doesn't seem to be useful in giving a continued fraction interpretation of the algorithm, while such an interpretation appears naturally in the work of Antoulas [1]. This is a problem which is under further investigation. Also the question if our approach gives all the elements to describe the fine block structure of a table of matrix Padé approximants and how they should be defined is still partially open.

25. References

[1] A.C. ANTOULAS : On recursiveness and related topics in linear systems, Techn. Report TR8507, Rice University, Houston, Texas (1985)

[2] E.R. BERLEKAMP : Nonbinary BCH decoding, International symposium on Information Theory, San Remo, Italy (1962)

[3] E.R. BERLEKAMP : *Algebraic coding theory*, McGraw-Hill, New York, chapts 7 and 10 (1968)

[4] Y. BISTRITZ : Nested bases of invariants for minimal realizations of finite matrix sequences, SIAM J. Control Opt., **21** , 804-821 (1983)

[5] Y. BISTRITZ, U. SHAKED : Discrete multivariable system approximations by minimal Padé type state models, IEEE Trans. Circuits and Systems, **CAS-31** (4) 382-390 (1984)

[6] Y. BISTRITZ, U. SHAKED : Minimal Padé model reduction for multivariable systems, Journal of Dynamic Systems, Measurement, and Control, **106** , 293-299 (1984)

[7] A. BULTHEEL : Recursive algorithms for nonnormal Padé tables, SIAM J. Appl. Math., **39** , 106-118 (1980).

[8] J. CONAN : A recursive procedure for the solution of minimal partial realization for scalar rational sequences, Revue Roum. Math. Pures Appl., **30** , 625-645 (1985)

[9] B.W. DICKINSON, M. MORF, T. KAILATH : A minimal realization algorithm for matrix sequences, IEEE Trans. Autom. Control, **AC-19** , 31-38 (1974)

[10] W.B. GRAGG, A. LINDQUIST : On the partial realization problem, Lin. Alg. Appl., **50** , 277-319 (1983)

[11] P.R. GRAVES-MORRIS : Vector valued rational interpolants II, IMA Journ. Numer. Anal., **4** , 209-224 (1984)

[12] T. KAILATH : *Linear Systems*. Prentice Hall, Englewood Cliffs, N.J. (1980)

[13] R.E. KALMAN : On minimal partial realizations of a linear input/output map, in : Aspects of Nerwork and Systems Theory, R.E. Kalman, N. Declaris eds., Holt, Reinhart

and Winston, New York, 385-407 (1971)

[14] R.E. KALMAN : On partial realization, transfer functions and canonical forms, Acta Polytechnica Scandinavica, **31** , 9-32 (1979)

[15] J.L. MASSEY : Shift register synthesis and BCH decoding, IEEE Trans. Information Theory, **IT-15** , 122-127 (1969)

[16] T. SHAMIR : The matrix partial realization problem : an algebraic approach, Int. J. Control, **43** (4), 1177-1199 (1986)

[17] W.A. WOLOVICH : The determination of state-space representations for linear multivariable systems, Automatica, **9** , 97-106 (1973)