A canonical matrix continued fraction solution of the minimal (partial) realization problem.

1. Introduction

The problem of minimal (partial) realization M(P)R is to find a (partial) model for the data

$M_1, M_2, M_3, \ldots$,

I.e. to find a rational function which has an expansion around infinity which is of the form

$$M_1 z^{-1} + M_2 z^{-2} + \cdots + M_n z^{-n} + \cdots .$$

The recursive solution of this problem has a long history and is in the case of scalar data M_k essentially given by the Euclidean algorithm. It constructs a continued fraction whose convergents are successive (minimal) partial realizations that approximate the data in a Padé -like sense. Some history on this problem and related topics is given in the introduction of [2]. Some key-papers are [13,12,10,7].

The extension of this theory for the case where the M_k are $p \times m$ matrices has a more recent history. The generalization of the Berlekamp-Massey algorithm was first given by Dickinson, Morf and Kailath [8] and Anderson, Brasch and Lopresti [1]. However the relation of their generalization to related topics of the Euclidean algorithm, continued fractions, Padé approximation etc. are not worked out. Fuhrmann [9] gives a matrix continued fraction and its relation to the Euclidean algorithm but some simple and nice properties are still lacking. Antoulas [2] then provides a theory which indeed seems to be the good approach to generalize all the nice properties of the scalar case. He uses a framework to derive his results that doesn't make clear how the original algorithms of [8] and [1] fit into this.

In a recent Ph.D. Thesis of the first author [14], it is shown how, starting from [8] (or [1]), you can obtain the results of Antoulas. We shall not give all the derivations here but only give a brief resume which is needed to formulate algorithm **MPR** of section 3. This is an explicit formulation of a slightly modified version of [8,1]. This modification makes the relation with the work of Antoulas [2] straightforward. It is not essential to start from [8,1] if you want to obtain the main results of this paper. They could as well have been obtained from [2] immediately. We shall give some suggestions in that direction below. However, we have chosen for the approach of this paper for two reasons. Firstly, it allows us to give an explicit and detailed algorithm which is closer to the presentation of [8,1], rather than a more theoretical characterization of the solution as in [2]. Secondly, because of the different approaches to the problem by [8,1] on one hand and [2] on the other, it is not trivial, except for a few specialists, to see that the relationship between the two is as intimate as given here.

The main results of this paper are in sections 4 and 5. In section 4 algorithm **MPR** is extended with some extra steps which are given in the definition of **MakeCanonical,** (see section 4) and which performs some normalization of the solution. The solution of the MPR problem is not unique but all possible solutions can be easily described once a solution is obtained by algorithm **MPR** (see section 3). The solution obtained by the normalization picks a specific solution and also fixes the other objects appearing in algorithm **MPR** uniquely. In this way everything is made canonical and we shall refer to this Canonical form of the algorithm **MPR** as **CMPR.** This canonical solution has a lot of very interesting properties which are given in section 5. Some of these (but not all) are also satisfied by the solution given by Antoulas. Section 6 gives an example.

The relations with the Euclidean algorithm and Padé approximapproximation is given elsewhere [6].

Although the approach used here is mainly like in [8,1], the results are close to those obtained in [2]. However, there are some essential differences which may be worth stressing. The main line of the recursive realization proposed in [2] is as follows. All (minimal) partial realizations for the data $M_1, M_2, \ldots, M_{n-1}$ are described in terms of some matrix rational function Z_n (which is basically the same as our Z_n of section 5). Then a closed form solution (formula (5.6) of [2]) is given for this Z_n to select among these a unique minimal partial realization of the previous data extended with M_n. Again all possible solutions can be described. It will then take some work to find a coprime polynomial numerator-denominator pair $C_n A_n^{-1}$ for Z_n. When these are obtained, one is ready to prepare for the next step.
The unique solution proposed by Antoulas in [2] also has a lot of interesting properties which are shared by the one proposed in this paper. However, our choice guarantees the physical realizability of the ladder network which can be associated with the algorithm (see the causality properties of theorem 5.10) while this is not true in general for the solution of [2]. Of course, since all the elements are present, it is possible to construct a physical realizable ladder network starting from the solution of [2]. However, we then have to go back to previous solutions and modify them. In this way, recursiveness is lost. (We mean here recursiveness in the sense of [2] and to a certain extend also recursiveness in the algorithmic sense.) As a matter of fact, the same problem appears in our solution. To define our canonical solution for $M_1, \ldots, M_n$, we need in general some information from $M_{n+1}, M_{n+2}, \ldots$. The problem is essentially the same : we anticipate the values to come, while in an adaptation of [2], one would wait until all this information is present and then go back as far as necessary to adapt previously computed solutions. The algorithm proposed in this paper is recursive in the algorithmic sense and it is possible to make it recursive in the sense of [2] if one postpones the computation of a new solution until all the information is available. This would make our algorithm considerably more complicated, in fact as complicated as the canonical adaptation of the solution of [2]. That is why we prefer not to do so.
Another difference between our approach and [2] is the following. For computational reasons, we think it more interesting to have a recursion for numerator and denominator, rather than having a "closed form" solution as in [2]. The formula (5.6) of [2] defines the Z_n mentioned before in terms of some Δ_n, which is the mismatching of the data M_n by a minimal partial realization of $M_1, \ldots, M_{n-1}$. To compute this, some inner product of block sequences involving its numerator and denominator are necessary. After that, if you want to go to the next stage of the realization, an algorithm has to be performed to find the polynomial numerator C_n and denominator A_n so that $Z_n = C_n A_n^{-1}$. In our approach, we first compute C_n and A_n and then we could set $Z_n = C_n A_n^{-1}$, but there is no need at all to do this.
Finally, because we propose an algorithm which works with the series S and R (see algorithm **MPR** in section 3), we get a so called "layer peeling" version of the algorithm, which does not need an inner product. The computational version of [2] requires an inner product in every step because it is basically a "layer adjoining" version. It is a known fact that for a parallel implementation, the latter is computationally less efficient. The terminology "layer peeling" and "layer adjoining" was introduced by T. Kailath and his coworkers when studying algorithms like this in a scattering framework. The Levinson algorithm is a typical example of a layer adjoining algorithm while the Schur version is a layer peeling analog.

2. The minimal (partial) realization problem

This problem has been given a lot of attention in linear system theory [11,12] and can be formulated as follows :
Consider the formal power series $M(z)$ around $z = \infty$:

$$M(z) = \sum_{k=1}^{\infty} M_k z^{-k} \quad \text{with} \quad M_k \in K^{p\times m}.$$

K is an arbitrary (possibly finite) field.
The coefficients M_k are called the Markov parameters of a linear constant strictly causal system. The power series $M(z)$ can be interpreted as the transfer function expanded in a power series around $z = \infty$. The theory is applicable not only to discrete time systems in which case the sequence $\{M_k\}$ is the impulse response sequence $\{h_k\}$ but also to continuous time systems in which case the M_k are equal to $\left.\frac{d^{k-1} h(t)}{dt^{k-1}}\right|_{t=0}$ where $h(t)$ is the impulse response.
We use the notation $K[z]$ to denote all polynomials with coefficients belonging to the field K; in a similar way $K^{p\times m}[z]$ and $K^m[z]$ are defined.
$K<z^{-1}>$ denotes the set of all Laurent series that can be written as $\sum_{k=n}^{\infty} c_k z^{-k}$, i.e. with a finite number of terms with a positive power of z. Similarly, $K^{p\times m} <z^{-1}>$ and $K^m <z^{-1}>$ are defined.
A partial realization of order n of $M(z)$ is a rational matrix function $C(z)A(z)^{-1}$ with $C(z) \in K^{p\times m}[z]$ and $A(z) \in K^{m\times m}[z]$ such that

$$C(z)A(z)^{-1} = M(z) + O(z^{-n-1}), \quad z \to \infty .$$

I.e. the first n Markov parameters of $M(z)$ and $C(z)A(z)^{-1}$ are the same. The partial realization is minimal if the degree of the determinant (deg det) of $A(z)$ is minimal. The qualification "partial" is skipped if $n = \infty$.

3. Solution of the minimal partial realization (MPR) problem

In this section we give an algorithm to solve the MPR problem very similar to the solution given by Dickinson, Morf and Kailath [8] and Anderson, Brasch and Lopresti [1]. More details about this algorithm, which we call algorithm **MPR**, can be found in these papers and in [14]. We shall briefly review the main ideas of this algorithm without proofs. Therefore we need some terminology and notation.
A nonzero vector $a(z) = \sum_{k=d}^{\infty} a_k z^{-k} \in K^m<z^{-1}>$ has degree d iff $a_d \neq 0$. The coefficient a_d is called the highest degree coefficient of $a(z)$. Note that this definition can also be applied to the special case of polynomial vectors. The degree of the zero vector is defined as $-\infty$.

A matrix

$$A(z) = \left[\sum_{k=d_1}^{\infty} a_{1,k} z^{-k} \quad \sum_{k=d_2}^{\infty} a_{2,k} z^{-k} \quad \cdots \quad \sum_{k=d_m}^{\infty} a_{m,k} z^{-k} \right] \in K^{p \times m} <z^{-1}> \quad a_{i,d_i} \neq 0$$

is called column reduced iff the matrix of the highest degree coefficients $a_{i,d_i} \neq 0$ is nonsingular. This column highest degree coefficient matrix is denoted by A_{hc}. Similarly row reducedness and the row highest degree coefficient matrix A_{hr} are defined.

A polynomial matrix $V(z) \in K^{m \times m}[z]$ is called unimodular iff $\det V(z) \in K \setminus \{0\}$. I.e. $V(z)$ has a polynomial inverse.

Every unimodular matrix $V(z)$ can be interpreted as a product of polynomial elementary transformation matrices :

- $E_{i,j}$ = the unit matrix in which column i and column j are interchanged.
- $E_{i,j}(q(z))$ = the unit matrix in which column j is replaced by column j plus $q(z) \in K[z]$ times row i.
- $E_i(k)$ = the unit matrix where column i is multiplied by $k \in K \setminus \{0\}$.

It is a well known fact [16] that every nonsingular polynomial matrix $A(z) \in K^{m \times m}[z]$ can be brought in column reduced form by postmultiplying with a unimodular matrix $V(z)$. I.e. there exists a unimodular matrix $V(z)$ such that $A(z)V(z)$ is column reduced. Because deg det $A(z)$ = deg det $A(z)V(z)$ we only have to look for minimal partial realizations $C(z)A(z)^{-1}$ with $A(z)$ column reduced. Algorithm **MPR** computes the columns $a_i(z)$ of such a column reduced denominator polynomial matrix $A(z)$. The numerator polynomial matrix $C(z)$ is easily computed when $A(z)$ is known : $C(z)$ = the polynomial part of $M(z)A(z)$.
The main point of [8,1,14] is the fact that the columns $a_i(z)$ of $A(z)$ have to be solutions of the set of homogeneous linear block Hankel equations determined by the first n Markov parameters $\{M_k\}_{k=1}^{n}$ if n is the order of the partial realization :

$$\begin{bmatrix} & & & M_1 \\ & M_1 & \cdots & M_\alpha \\ M_1 & M_2 & \cdots & M_{1+\alpha} \\ \vdots & & & \vdots \\ M_{n-\alpha} & & \cdots & M_n \\ M_{n-\alpha+1} & & \cdots & M_{n+1} \\ \vdots & & & \vdots \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_\alpha \end{bmatrix} = \begin{bmatrix} c_{\alpha-1} \\ c_{\alpha-2} \\ \vdots \\ c_0 \\ 0 \\ \vdots \\ 0 \\ r_1 \\ r_2 \\ \vdots \end{bmatrix} \tag{3.1}$$

where

$a(z) = \sum_{i=0}^{\alpha} a_i z^i$ represents a column of the denominator $A(z)$ of degree α

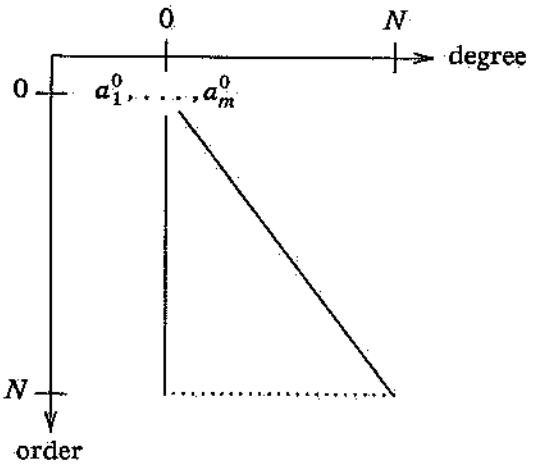
$c(z) = \sum_{i=0}^{\alpha-1} c_i z^i$ represents the corresponding column of the numerator polynomial matrix $C(z)$

$r(z) = \sum_{i=1}^{\infty} r_i z^{-n+\alpha-i}$ is called the corresponding residual series of $a(z)$.

Written in $K\langle z^{-1}\rangle$-notation (3.1) is equivalent with

$$M(z)\,a(z) = c(z) + r(z) \quad \text{where } c(z) \in K^p[z] \text{ and } r(z) = O(z^{-N+\alpha-1}),\ z \to \infty .$$

Let $a_i^k(z)$, $i = 1, \ldots, m$ be the columns of the column reduced denominator polynomial $A_k(z)$ of a k-th order MPR, $k = 0,1,2, \ldots$. We shall call these vectors solution vectors. They can be put in a triangular table called the solution table :



Note that $[a_1^0\ a_2^0\ \cdots\ a_m^0]$ can be chosen as the $m \times m$ unit matrix I_m. Note also that the degree α_i^k of $a_i^k(z)$ is at most k. The residual $Ra_i^k(z)$ is defined as the coefficient of degree $-k-1+\alpha_i^k$ of the residual series $r_i^k(z)$. It is called discrepancy in [1]. This is the first coefficient in $r_i^k(z)$ which is not explicitly annihilated by (3.1). The subspace of K^p generated by the residuals

$Ra_i^k(z)$, $i = 1,2,\ldots,m$; $k < n$ is called the residual space for order n. The potential degree [8] or accumulation index [1] with respect to order n of a solution vector $a_i^k(z)$ of degree α_i^k is defined as $\pi_j^n = n-k+\alpha_i^k$. If the residual space has dimension s, then there exist s solution vectors with minimal potential degrees with respect to order n whose residuals span the residual space. These will be called auxiliary vectors, and we denote them by $x_j^n(z)$, $j = 1,\ldots,s$. The following property is basic for the algorithm :

Theorem 3.1 :

A basis for the kernel of the block Hankel matrix

$$H = \begin{bmatrix} M_1 & \cdots & M_{1+\alpha} \\ \vdots & & \vdots \\ M_{n-\alpha} & \cdots & M_n \end{bmatrix}$$

or equivalently, a basis for the space of all $a(z) = \sum_{i=0}^{\alpha} a_i z^i$ satisfying

$$H\,[a_0^T \cdots a_\alpha^T]^T = 0$$

is given by

$$B(n,\alpha) = A_\alpha \bigcup X_\alpha$$

with

$$A_\alpha = \bigcup_{j=1}^{m} \{a_j^n(z)\}_{\alpha_j^n}^{\alpha} \quad \text{and} \quad X_\alpha = \bigcup_{j=1}^{s} \{x_j^n(z)\}_{\pi_j^n}^{\min(\alpha,n)}$$

where

$$\begin{aligned} \{a(z)\}_\alpha^\beta &= \{a(z), za(z), \ldots, z^{\beta-\alpha}\} \quad \text{if} \quad \alpha \leqslant \beta \\ &= \varnothing \qquad\qquad\qquad\qquad\quad \text{if} \quad \alpha > \beta \end{aligned}$$

and

$a_j^n(z)$ are the solution vectors of degree α_j^n
$x_j^n(z)$ are the auxiliary vectors of potential degree π_j^n

for the MPR problem of order n for the sequence $M_1, M_2, \ldots$

Proof : See [1], [14, Theorem 4.1].

□

Using this basis, algorithm **MPR** constructs the solution vectors $a_i^{n+1}(z)$ of order $n+1$ based on the solution vectors $a_j^n(z)$ and auxiliary vectors $x_j^n(z)$. This is done by updating the G matrix for the successive orders. This G matrix contains all the quantities involved : (we drop the explicit indication of the order for notational simplicity)

$$G = \begin{bmatrix} Y & C \\ X & A \\ S & R \end{bmatrix} \quad \text{with}$$

$X(z) = [X_{ij}]$ with $X_{ij} = \sum_{d=0}^{\xi_{ij}} X_{i,j,d}\, z^d$, $\xi_{ij} = \deg X_{ij}$ (auxiliary vectors)

$A(z) = [A_{ij}]$ with $A_{ij} = \sum_{d=0}^{\alpha_{ij}} A_{i,j,d}\, z^d$, $\alpha_{ij} = \deg A_{ij}$ (solution vectors)

$S(z) = [S_{ij}]$ with $S_{ij} = \sum_{d=0}^{\sigma_{ij}} S_{i,j,d}\, z^d$, $\sigma_{ij} = \deg S_{ij}$ (residual series of auxiliary vectors)

$R(z) = [R_{ij}]$ with $R_{ij} = \sum_{d=0}^{\rho_{ij}} R_{i,j,d}\, z^d$, $\rho_{ij} = \deg R_{ij}$ (residual series of solution vectors)

Note that if one of these degrees is $-\infty$ we define $\sum_{d=0}^{-\infty} c_d\, z^d \equiv 0$.

The column degrees are defined by the following functions :

$\alpha_j = \max \{\alpha_{ij} \mid i = 1, \ldots, m\}$

$\xi_j = \max \{\xi_{ij} \mid i = 1, \ldots, m\}$

$\sigma_j = \max \{\sigma_{ij} \mid i = 1, \ldots, p\}$

$\rho_j = \max \{\rho_{ij} \mid i = 1, \ldots, p\}$

and we define

$\alpha_{\min} = \min \{\alpha_j \mid j = 1, \ldots, m\}$

$\sigma_{\min} = \min \{\sigma_j \mid j = 1, \ldots, p\}$.

a_i denotes the i-th column of $A(z)$ (similarly for the other blocks).

Note that algorithm **MPR** described below constructs G for the next order such that A_{hc} is a unit upper triangular matrix. This implies also that the degrees α_j for order n are equal to the controllability indices $\kappa_j(n)$ [14,4]. The G matrix for order $n+1$ is equal to the G matrix of order n multiplied to the right by the unimodular matrix $V_{n+1}(z)$. These unimodular matrices $V_k(z)$ contain all the information to construct a matrix continued fraction whose convergents are the MPRs $C(z)A(z)^{-1}$.

MPR :=

InitializeG

For $n = 0,1,2, \ldots$

GenerateSomeSolution

GenerateV

MakeCanonical

end MPR

InitializeG :=

{A possible choice for G of order 0}

Set $G = \begin{bmatrix} Y & C \\ X & A \\ S & R \end{bmatrix} = \begin{bmatrix} I_p & 0 \\ 0 & I_m \\ -I_p & M(z) \end{bmatrix}$

end Initialization

GenerateSomeSolution :=

{Compute G of order $n+1$ based on G of order n}

For $i = 1, \ldots, m$

Set Ra_i = the coefficient of $z^{-n-1+\alpha_i}$ in r_i.

If $Ra_i = 0$ **then** $V^{(i)} = I_{p+m}$.

If $Ra_i \neq 0$ **then**

Write Ra_i as a linear combination of the columns of S_{hc}, i.e. $Ra_i + S_{hc}\,[c_1, \ldots, c_p]^T = 0$

Choose some l such that $\sigma_l = \max \{\sigma_j \mid c_j \neq 0\,,\ j = 1, \ldots, p\}$

If $\sigma_l > \rho_i$ **then**

$$V^{(i)} = E_{l,p+i}\, E_{p+i}\,(c_l) \prod_{\substack{j=1 \\ j \neq l}}^{p} E_{j,p+i}\,(c_j z^{\sigma_l - \sigma_j})\, E_{l,p+i}\,(z^{\sigma_l - \rho_i})$$

If $\sigma_l \leqslant \rho_i$ **then**

$$V^{(i)} = \prod_{j=1}^{p} E_{j,p+i}\,(c_j\, z^{\rho_i - \sigma_j})$$

Replace G by $G\ V^{(i)}$

end GenerateSomeSolution

GenerateV :=

$$\text{Set } V = \prod_{i=1}^{m} V^{(i)}$$

end GenerateV

For the moment the definition of **MakeCanonical** is empty. Thus for the moment it can be ignored. In the next section when a canonical solution will be generated out of the present one, a definition of this step will be given wich requires some non trivial operations.

Notes :

1. Algorithm **MPR** essentially eliminates the residual Ra_i of each solution vector a_i. If this residual is zero, a_i is also a solution vector of order $n+1$. If not, then we have to take a linear combination c of a_i and the auxiliary vectors x_j such that the residual of c with respect to order n is zero. If the highest potential degree π_l of the auxiliary vectors active in the linear combination c is greater than the degree α_i of the solution vector a_i, the a_i can substitute x_l. If $\pi_l \leqslant \alpha_i$ then the auxiliary vectors can be left unchanged. Because the residual degree σ_i of s_i and the potential degree π_i of x_i are related by $\sigma_i = -n-1-\pi_i$, it is easy to see that the tests $\pi_l > (\leqslant) \alpha_i$ turn out to be equivalent with the tests $\sigma_l > (\leqslant) \rho_i$ used in the algorithm.
2. The V matrix obtained in the step **GenerateV** is the matrix V_{n+1}.
3. Algorithm **MPR** is a so called "layer peeling" variant of a Berlekamp-Massey type algorithm since it updates the residual series of solution and auxiliary vectors together with the vectors themselves. This is extremely simple since they obey the same recursion. Therefore, the residuals are directly taken from the residual series and are not computed by multiplying a block row of the Hankel matrix with the solution vectors. The latter strategy is called "layer adjoining" and this is the one that is used in the scalar Berlekamp-Massey algorithm. The algorithm can be easily adapted to get this variant. One can drop the residuals from the recursion and use the equivalent tests described in note 1. For a parallel implementation however, the layer peeling variant is more efficient because it is free of inner products.
4. We also included the numerators in the algorithm. They are not necessary to keep the algorithm going (only the residual series, the degrees of the solution vectors and the order are needed), but the Y and C blocks will be used to construct a matrix continued fraction later on. Moreover, they are a part of the V matrix which plays a central role in the theory.

Once the solution vectors and the auxiliary vectors are known, it is easy to write down all the different solutions. We have the following theorem :

Theorem 3.2 :

All the different solutions of the n-th order minimal partial realization problem are given by

$C^{\lambda}(z)A^{\lambda}(z)^{-1}$ with the columns $a_j^{\lambda}(z)$ of $A^{\lambda}(z)$ given by :

$$a_j^{\lambda}(z) = a_j(z) + \sum_{k=1}^{p} \sum_{i=\pi_k}^{\alpha_j} \lambda_{k,i}^{j} z^{i-\pi_k} x_k(z) \ , \ j = 1,2, \ldots ,m \ . \tag{3.2}$$

π_k is the potential degree with respect to order n of auxiliary vector $x_k(z)$
α_j is the degree of solution vector $a_j(z)$ of order n
$\lambda_{k,i}^{j} \in K$ form a minimal set of free parameters.

Proof : See [1] [14, Theorem 4.2].

□

4. A canonical solution of the problem

When we use algorithm **MPR** to compute certain MPRs $C(z)A(z)^{-1}$ of successive orders, we obtain the unimodular matrices $V_1, V_2, \ldots, V_n, V_{n+1}, \ldots$. Later on we indicate how these V_n matrices define a matrix continued fraction whose convergents are the MPRs $C(z)A(z)^{-1}$. As can be seen e.g. from (3.2) these V_n matrices are not unique. We shall put some extra conditions on the G matrices so that they become unique and and consequently also the V_n matrices will be unique. In this section we shall state the extra conditions and an algorithm to transform G by elementary polynomial transformations to its unique form defined by these extra conditions.
Suppose G_n is the G matrix of order n. Algorithm **MPR** constructs a unimodular matrix V_{n+1} such that $G'_{n+1} = G_n V_{n+1}$ is a certain choice for the G matrix of order $n+1$. We shall construct a unimodular matrix W_{n+1} such that $G_{n+1} = G'_{n+1} W_{n+1} = G_n V_{n+1} W_{n+1}$ will also satisfy the extra conditions. From now on, when we use the notation $A = O(z^{f(i,j)})$ for a matrix A, this means that $A_{ij} = O(z^{f(i,j)})$, $z \to \infty$. The four extra conditions are :

(a) $A(z)$ has to be canonical (i.e. column and row reduced) with

A_{hc} a unit upper triangular matrix and
A_{hr} the $m \times m$ unit matrix I_m .

(b) $A(z)^{-1}X(z)$ has to be strictly proper i.e.

$$A(z)^{-1}X(z) = O(z^{-1})$$

Together with condition (a) this means that

$$X(z) = O(z^{\kappa_i - 1})$$

where $\kappa_i = \alpha_i$ are the controllability indices for the sequence $M_1, M_2, \ldots$

(c) $S(z)$ has to be canonical with

$-S_{hc}$ unit lower triangular and
$-S_{hr}$ the $p \times p$ unit matrix I_p.

(d) $S(z)^{-1}R(z)$ has to be strictly proper i.e. assuming $S(z)$ is row reduced with row (and column) degrees σ_i (see condition (c)) we make $R(z) = O(z^{\sigma_i - 1})$.

We describe now the steps of algorithm **MakeCanonical** that constructs the unimodular matrix $W = W_{n+1}$. This definition should replace the empty definition of the previous section. For further reference, we shall call the algorithm **MPR** with this non empty definition the **CMPR** algorithm (C for Canonical).

MakeCanonical :=

InitializeW

MakeACanonical

MakeShcLowerTriangular

MakeAinverseXStriclyProper

MakeSCanonical

MakeSinverseRStrictlyProper

MakeV

end MakeCanonical

InitializeW :=

Set $W(z) = I_{p+m}$

end InitializeW

MakeACanonical :=

{Make $A(z)$ canonical (condition (a)) ; note that A_{hc} is already unit upper triangular}

For $j = 1,2,\ldots,m$

For $d = \alpha_j, \alpha_{j-1}, \ldots, \alpha_{\min}$

For $i = m, m-1, \ldots, 1$

If $i \neq j$ and $d \geqslant \alpha_i$ and $A_{i,j,d} \neq 0$ **then**

$\gamma = d - \alpha_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{p+i,p+j}(-M_{i,j,d}\, z^{\gamma})$

end MakeACanonical

MakeShcLowerTriangular :=

{Before we make $A(z)^{-1}X(z)$ strictly proper (condition (b)) we can make $-S_{hc}$ unit lower triangular (condition (c)). Note that $S(z)$ is already column reduced}

For $i = 1,2,\ldots,p$

Choose some l such that $e = \sigma_{il} = \min\{\sigma_{ij} \mid \sigma_{ij} = \sigma_j\ ,\ j = i,\ldots,p\}$

Set $q = -S_{i,l,e}^{-1}$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{i,l}E_i(q)$

For $j = i+1,\ldots,p$

If $\sigma_i \leqslant \sigma_j$ and $S_{i,j,\sigma_j} \neq 0$ **then**

Set $\gamma = \sigma_j - \sigma_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{i,j}(S_{i,j,\sigma_j}z^{\gamma})$

end MakeShcLowerTriangular

MakeAinverseXStrictlyProper :=

{Make $A(z)^{-1}X(z)$ strictly proper (condition (b))}

For $j = 1,2,\ldots,p$

For $d = \xi_j,\ldots,\alpha_{\min}$

For $i = m,m-1,\ldots,1$

If $d \geqslant \alpha_i$ and $X_{i,j,d} \neq 0$ **then**

$\gamma = d - \alpha_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{p+i,j}(-X_{i,j,d}z^{\gamma})$

end MakeAinverseXStrictlyProper

MakeSCanonical :=

{Make $S(z)$ canonical (condition (c)) while condition (b) is maintained}

For $j = 1, \ldots, p$

For $e = \sigma_j, \sigma_j - 1, \ldots, \sigma_{\min}$

For $i = 1, \ldots, p$

If $i \neq j$ and $e \geqslant \sigma_i$ and $S_{i,j,e} \neq 0$ **then**

$\gamma = e - \sigma_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{i,j}(S_{i,j,e} z^{\gamma})$

{This part of the program makes $A(z)^{-1} X(z)$ strictly proper again}

For $d = \xi_j, \xi_j - 1, \ldots, \alpha_{\min}$

For $i = m, m-1, \ldots, 1$

If $d \geqslant \alpha_i$ and $X_{i,j,d} \neq 0$ **then**

$\gamma = d - \alpha_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{p+i,j}(-X_{i,j,d} z^{\gamma})$

end MakeSCanonical

MakeSinverseRStrictlyProper :=

{Make $S(z)^{-1}\ R(z)$ strictly proper (condition (d)) while $A(z)$ remains canonical}

For $j = 1, \ldots, m$

For $e = \rho_j, \rho_j - 1, \ldots, \sigma_{\min}$

For $i = 1, \ldots, p$

If $e \geqslant \sigma_i$ and $R_{i,j,e} \neq 0$ **then**

$\gamma = e - \sigma_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{i,p+j}(R_{i,j,e} z^{\gamma})$

{Make $A(z)$ canonical again (condition (a))}

For $d = \alpha_j - 1, \alpha_j - 2, \ldots, \alpha_{\min}$

For $i = m, m-1, \ldots, 1$

If $d \geqslant \alpha_i$ and $A_{i,j,d} \neq 0$ **then**

$\gamma = d - \alpha_i$

Replace $\begin{bmatrix} G \\ W \end{bmatrix}$ by $\begin{bmatrix} G \\ W \end{bmatrix} E_{p+i,p+j}(-A_{i,j,d} z^{\gamma})$

end MakeSinverseRStrictlyProper

MakeV :=

Set $V(z) = V(z)W(z)$

end MakeV

5. Properties of the solution matrices G_n and the unimodular matrices V_n

In this section we shall state and proof some very interesting properties of the matrices G_n and V_n obtained by applying algorithm **CMPR**. We shall use the following notation

$$G_n = \begin{bmatrix} Y_{1,n} & C_{1,n} \\ X_{1,n} & A_{1,n} \\ S_{1,n} & R_{1,n} \end{bmatrix} = G_0 \, V_1 \, V_2 \, \cdots \, V_n = \begin{bmatrix} V_{1,n} \\ S_{1,n} \quad R_{1,n} \end{bmatrix} \tag{5.1a}$$

$$V_n = \begin{bmatrix} Y_n & C_n \\ X_n & A_n \end{bmatrix} \quad \text{and} \tag{5.1b}$$

$$\Lambda_n = V_n^{-1} = \begin{bmatrix} D_n & -B_n \\ -U_n & T_n \end{bmatrix} \quad \text{and} \tag{5.1c}$$

$$\Lambda_{1,n} = V_{1,n}^{-1} = \begin{bmatrix} D_{1,n} & -B_{1,n} \\ -U_{1,n} & T_{1,n} \end{bmatrix}. \tag{5.1d}$$

Theorem 5.1 :

The inverse of a square column reduced matrix $A(z)$ with column degrees α_j is a row reduced matrix $B(z)$ with $[B(z)]_{hr} = [A(z)]_{hc}^{-1}$ and row degrees $-\alpha_i$. Similarly for a row reduced matrix $A(z)$.

Proof : (trivial).

□

Theorem 5.2 :

The blocks of the matrix G_n, associated with $M(z)$ by the **CMPR** algorithm of the previous section satisfy

$$M(z)\, A_{1,n}(z) - C_{1,n}(z) = R_{1,n}(z)$$

and

$$M(z)\, X_{1,n}(z) - Y_{1,n}(z) = S_{1,n}(z)\,.$$

Proof : (trivial).

□

Theorem 5.3 :

The MPR of order n for $M(z)$ as generated by algorithm **CMPR** satisfies

(e) $\quad M(z) = C_{1,n}(z)\, A_{1,n}(z)^{-1} + O(z^{-\kappa_j(n)+\sigma_i(n)-1})$

where $\kappa_j(n)$ are the controllability indices of $M_1, \ldots, M_n$ and $\sigma_i(n)$ are the column degrees of $S_{1,n}(z)$.

Proof :

From theorem 5.2 it follows that

$$M(z) = C_{1,n}A_{1,n}^{-1} + R_{1,n}\,A_{1,n}^{-1}.$$

From condition (d) together with (c) we know that

$$R_{1,n}(z) = O(z^{\sigma_i(n)-1}).$$

Condition (a) and theorem 5.1 imply that $A_{1,n}^{-1}$ is column reduced with column degrees $-\kappa_j(n)$.

So

$$R_{1,n}A_{1,n}^{-1} = O(z^{\sigma_i(n)-\kappa_j(n)-1})$$

and this proves the theorem.

□

Theorem 5.4 :

Conditions (a), (b), (c) and (d) imply that

(f) $D_{1,n}(z)$ is canonical with row degrees $-\sigma_i(n)$ and

D_{hr} is a unit lower triangular matrix and
D_{hc} is the $p\times p$ unit matrix I_p

The $\sigma_i(n)$ are the column degrees of $S_{1,n}(z)$. $D_{1,n}(z)$ is defined by (5.1).

Proof : (see [14, theorem 6.5])

Using the definition of $S_{1,n}$ and (e) we get :

$$\begin{aligned} S_{1,n} &= M\,X_{1,n} - Y_{1,n} \\ &= (C_{1,n}\,A_{1,n}^{-1} + O(z^{-\kappa_j(n)+\sigma_i(n)-1}))\,X_{1,n} - Y_{1,n} \\ &= (C_{1,n}\,A_{1,n}^{-1} - Y_{1,n}) + O(z^{-\kappa_j(n)+\sigma_i(n)-1})\,X_{1,n}. \end{aligned} \tag{5.2}$$

From (a) and (b) it follows that

$$X_{1,n} = O(z^{\kappa_i(n)-1})$$

thus that the last term in (5.2) gives

$$O(z^{-\kappa_j(n)+\sigma_i(n)-1})\,X_{1,n} = O(z^{\sigma_i(n)-2}).$$

From $V_{1,n}\Lambda_{1,n} = I_{p+m}$ we get

$$-D_{1,n}^{-1} = C_{1,n}\,A_{1,n}^{-1}\,X_{1,n} - Y_{1,n}$$

so that

$$S_{1,n} = -D_{1,n}^{-1} + O(z^{\sigma_i(n)-2}) .$$

If (c) is satisfied

$$S_{1,n} - O(z^{\sigma_i(n)-2}) = \text{diag}\,(z^{\sigma_1(n)}, \ldots, z^{\sigma_p(n)})\,[S_{1,n}]_{hr} + \text{lower order terms}$$

so $D_{1,n}$ will be column reduced with

$$[D_{1,n}]_{hc} = I_p = -[S_{1,n}]_{hr} .$$

To prove the row reducedness of $D_{1,n}$ we use the fact that the auxiliary vectors x_j of order n are solution vectors of order $\omega_j < n$ that are not solution vectors of order ω_j+1. I.e.

$$S_{1,n} = M\,X_{1,n} - Y_{1,n} = O(z^{-\omega_j-1+\xi_j}) \tag{5.3}$$

or $-\omega_j-1+\xi_j = \sigma_j(n)$ where ξ_j is the degree of x_j.

(5.2) can also be written as :

$$\begin{aligned} S_{1,n} &= -D_{1,n}^{-1} + O(z^{-n-1+\xi_j}) \\ &= O(z^{-\omega_j-1+\xi_j}) \qquad \text{(from (5.3))} \end{aligned}$$

with $\omega_j < n$.

Because $S_{1,n}$ is column reduced with $[-S_{1,n}]_{hc}$ equal to a unit lower triangular matrix, it is clear that $D_{1,n}$ is row reduced with $[D_{1,n}]_{hr}$ unit lower triangular.

□

Theorem 5.5 :

Under the same conditons as in theorem 5.4 we have :
The row (or column) degrees $-\sigma_i(n)$ of $D_{1,n}$ are equal to the observability indices $\nu_i(n)$ of the sequence $M_1, M_2, \ldots, M_n$.

Proof :

Because $\Lambda_{n+1}V_{n+1} = I_{p+m}$ we get that $D_{1,n}^{-1}B_{1,n}$ is a left coprime matrix fraction description for $C_{1,n}\,A_{1,n}^{-1}$. Because $D_{1,n}$ is row reduced and $[D_{1,n}]_{hr}$ is unit lower triangular (condition (f)) it follows that $-\sigma_i(n) = \nu_i(n)$.

□

Corollary 5.6 :

From theorem 5.3 and 5.5 we get that the MPR of $M(z)$ generated by algorithm CMPR satisfies

$$M(z) - C_{1,n}\,A_{1,n}^{-1} = O(z^{-\kappa_j(n)-\nu_i(n)-1})$$

where $\kappa_j(n)$ and $\nu_i(n)$ are the Kronecker indices, i.e., controllability resp. observability indices for the sequence $M_1, \ldots, M_n$.

□

Theorem 5.7 :

For each order $n = 0,1,2,3,\ldots$ there is one and only one matrix G_n such that conditions (a), (b), (c) and (d) are satisfied.

Proof :

Suppose that there are two different matrices G and G' of order n satisfying conditions (a), (b), (c) and (d). (We drop the n from the notation for simplicity). From corollary 5.6 we get :

$$M(z) - C\,A^{-1} = O(z^{-\kappa_j-\nu_i-1})$$

$$M(z) - D'^{-1}\,B' = O(z^{-\kappa_j-\nu_i-1}).$$

Subtracting these two equations leads to :

$$C\,A^{-1} - D'^{-1}\,B' = O(z^{-\kappa_j-\nu_i-1})\,.$$

Multiplying to the right by A and to the left by D' gives :

$$D'\,C - B'\,A = O(z^{-1}).$$

The left hand side of this equation is a polynomial matrix, so the right hand side has to be zero. It follows then that

$$C\,A^{-1} = D'^{-1}\,B' = C'\,A'^{-1}.$$

Because A and A' are canonical by condition (a) we get that [3]

$$C = C' \quad \text{and} \quad A = A' \quad \text{or similarly} \quad D = D' \quad \text{and} \quad B = B'\,. \tag{5.4}$$

From (5.4) it follows that (see (5.1) for notation)

$$\begin{bmatrix} D' & -B' \\ -U' & T' \end{bmatrix} \begin{bmatrix} Y & C \\ X & A \end{bmatrix} = \begin{bmatrix} I & 0 \\ P & I \end{bmatrix} \quad \text{with} \quad P = T'X - U'Y$$

or

$$\begin{bmatrix} Y & C \\ X & A \end{bmatrix} = \begin{bmatrix} Y' & C' \\ X' & A' \end{bmatrix} \begin{bmatrix} I & 0 \\ P & I \end{bmatrix} = \begin{bmatrix} Y'+C'P & C' \\ X'+A'P & M' \end{bmatrix}.$$

This implies that $X = X' + A'P$. Using condition (b) we get $A^{-1}X = A'^{-1}X + P$. Thus we find $P = 0$.
This gives us : $U' = U$, $T' = T$ and similarly $Y' = Y$, $X' = X$.
Because the upper parts V and V' of the matrices G and G' are identical also the residual sequences will be the same.

□

Note :

A very simple proof can be given using a result of [4] and [5] where it is proved that there is a unique MPR $M_{1,n}(z)$ of order n such that :

$$M(z) - M_{1,n}(z) = O(z^{-\kappa_j(n)-\nu_i(n)-1}) . \tag{5.5}$$

From corollary 5.6, it follows that algorithm **CMPR** choses this unique MPR.

The role of the conditions (a), (b), (c) and (f) can be explained as follows :

If a specific MPR is chosen together with its G matrix G_n, then all G matrices coupled with this MPR are such that the upper part $V_{1,n}$ is unimodular while the denominator $A_{1,n}$ is unique up to a unimodular transformation. Hence, all G_n matrices are then given by

$$G_n \begin{bmatrix} U_2 & O \\ P & U_1 \end{bmatrix}$$

where

$U_1 \in K^{m\times m}[z]$ and $U_2 \in K^{p\times p}[z]$ are unimodular and $P \in K^{m\times p}[z]$ polynomial.

Condition (a) determines U_1 ; condition (c) or (f) determine U_2 and condition (b) fixes P. Condition (d) is implicitly used to pick the unique MPR satisfying (5.5).

Corollary 5.8 :

Because the G_n matrices are unique under the previous conditions, also their upper parts $V_{1,n}$ and the building blocks V_n will be uniquely defined.

□

The unique MPR generated by algorithm **CMPR** and all the objects related to it which will also be unique will be given the adjective canonical.

The following theorem gives the connection with the continued fractions as given by Antoulas [2].

Theorem 5.9 :

Let the canonical unimodular matrices

$$V_n = \begin{bmatrix} Y_n & C_n \\ X_n & A_n \end{bmatrix}$$

be as defined in step **MakeV** of algorithm **CMPR**, then the convergents of the following canonical matrix continued fraction

$$\cfrac{}{C_1 + Y_1 \cfrac{}{A_2 + X_2 \cfrac{}{\cdots}}} \Bigg/ \cfrac{}{A_1 + X_1 \cfrac{}{C_2 + Y_2 \cfrac{}{\cdots}}} \quad , \quad \frac{A}{B} = AB^{-1}$$

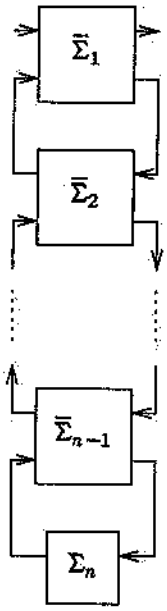
are the canonical minimal partial realizations $C_{1,n}(z)\ A_{1,n}(z)^{-1}$.

Proof : [2].

□

The last property which we shall show here is about the physical realizability of the matrix continued fraction of theorem 5.9 by a ladder network.

It is described in Antoulas [2] how the n-th convergent of the continued fraction can be realized by a ladder network



where the transfer function of Σ_n is $Z_n = C_n A_n^{-1}$ and the transfer function of the augmented system $\bar{\Sigma}_j$ is

$$\bar{Z}_j = \begin{bmatrix} Z_j & D_j^{-1} \\ A_j^{-1} & -A_j^{-1}X_j \end{bmatrix} = \begin{bmatrix} I & -C_j \\ 0 & A_j \end{bmatrix}^{-1} \begin{bmatrix} 0 & Y_j \\ I & -X_j \end{bmatrix}$$

$$= \begin{bmatrix} I & 0 \\ -U_j & T_j \end{bmatrix} \begin{bmatrix} 0 & I \\ D_j & -B_j \end{bmatrix}^{-1} .$$

The overall transfer function being $Z_{1,n} = C_{1,n}A_{1,n}^{-1}$.
If each system $\bar{\Sigma}_j$ has to be physically realizable, then $\bar{Z}_j(z)$ should be proper rational. For a canonical continued fraction this will be true i.e. :

Theorem 5.10 :

For a canonical continued fraction, each $\bar{\Sigma}_j$ will be causal, i.e. $\bar{Z}_j$ is proper. We shall prove the more general relations for $l \geq k+1$

1. $C_{k+1,l}A_{k+1,l}^{-1} = Z_{k+1,l} = O(z^{-1})$ (strictly proper)
2. $A_{k+1,l}^{-1} = O(z^0)$ (proper)
3. $D_{k+1,l}^{-1} = O(z^0)$ (proper)
4. $A_{k+1,l}^{-1}X_{k+1,l} = O(z^{-1})$ (strictly proper)

which give for $l = k+1$ the causality of $\bar{\Sigma}_{k+1}$.

Proof :

1. Because of corollary 5.6

$$(\Delta Z)_{k+1,l} = (M-Z_{1,k}) - (M-Z_{1,l})$$
$$= O(z^{-\kappa_j(k)-\nu_i(k)-1}) - O(z^{-\kappa_j(l)-\nu_i(l)-1})$$

and because the $\kappa_j(n)$ and $\nu_i(n)$ are not decreasing with n, we have

$$(\Delta Z)_{k+1,l} = O(z^{-\kappa_j(k)-\nu_i(k)-1}).$$

It was shown by Antoulas [2,p.1125] that $Z_{1,l} = Z_{1,k} + (\Delta Z)_{k+1,l}$ with

$$(\Delta Z)_{k+1,l} = D_{1,k}^{-1}\, Z_{k+1,l}\, A_{1,k}^{-1}\, (I + X_{1,k}\, Z_{k+1,l}\, A_{1,k}^{-1})^{-1}$$

or

$$D_{1,k}^{-1}O(z^{-\kappa_j(k)-\nu_i(k)-1}) = Z_{k+1,l}A_{1,k}^{-1}(I + X_{1,k}Z_{k+1,l}A_{1,k}^{-1})^{-1}.$$

Because $D_{1,k}$ is column reduced with column degrees $\nu_i(k)$, the left hand side will be $O(z^{-\kappa_j(k)-1})$.

Condition (b) states that $A_{1,k}^{-1}X_{1,k}$ is strictly proper, while condition (a) includes that $A_{1,k}$ is row reduced with row degrees $\kappa_j(k)$. Therefore row j of $X_{1,k}$ will be of degree $\leqslant \kappa_j(k)-1$. Thus we have

$$O(z^{-\kappa_j(k)-1})(I+X_{1,k}Z_{k+1,l}A_{1,k}^{-1}) = Z_{k+1,l}A_{1,k}^{-1}$$

or

$$O(z^{-\kappa_j(k)-1}) + O(z^{-2})Z_{k+1,l}A_{1,k}^{-1} = Z_{k+1,l}A_{1,k}^{-1}$$

or

$$O(z^{-\kappa_j(k)-1})A_{1,k} + O(z^{-2})Z_{k+1,l} = Z_{k+1,l}.$$

Again because $A_{1,k}$ is row reduced with row degrees $\kappa_j(k)$, we have

$$O(z^{-1}) + O(z^{-2})Z_{k+1,l} = Z_{k+1,l}$$

or

$$(I+O(z^{-2}))Z_{k+1,l} = O(z^{-1})$$

or

$$Z_{k+1,l} = O(z^{-1}).$$

2. Because

$$X_{1,k}C_{k+1,l} + A_{1,k}A_{k+1,l} = A_{1,l}$$

we have

$$(A_{1,k}^{-1}X_{1,k})(C_{k+1,l}A_{k+1,l}^{-1}) + I = A_{1,k}^{-1}A_{1,l}A_{k+1,l}^{-1}.$$

$A_{1,k}^{-1}X_{1,k} = O(z^{-1})$ by condition (b) and $Z_{k+1,l} = O_-(z^{-1})$ as we have proved in 1. Hence

$$A_{1,l}^{-1}A_{1,k}(I+O(z^{-2})) = A_{k+1,l}^{-1}$$

Because of the canonicity of the $A_{1,n}$ and because the $\kappa_j(n)$ are nondecreasing with n we find that $A_{1,l}^{-1}A_{1,k} = O(z^0)$, so that also $A_{k+1,l}^{-1} = O(z^0)$.

3. This point can be shown as in 2 by duality.

4. By condition (b) we have that $A_{1,k}^{-1}X_{1,k}$ is strictly proper, because $A_{1,k}^{-1}X_{1,k} = U_{1,k}D_{1,k}^{-1}$, the latter will also be strictly proper. Now consider

$$U_{1,l}D_{1,l}^{-1} = [U_{k+1,l}D_{1,k}+T_{k+1,l}U_{1,k}]\,[D_{k+1,l}D_{1,k}+B_{k+1,l}U_{1,k}]^{-1}.$$

Use in the numerator of the right hand side

$$T_{k+1,l} = A_{k+1,l}^{-1}X_{k+1,l}B_{k+1,l} + A_{k+1,l}^{-1}$$

and

$$U_{k+1,l} = A_{k+1,l}^{-1}X_{k+1,l}D_{k+1,l}$$

to get

$$U_{1,l}D_{1,l}^{-1} = [A_{k+1,l}^{-1}X_{k+1,l}] + A_{k+1,l}^{-1}U_{1,k}\,[D_{k+1,l}D_{1,k}+B_{k+1,l}U_{1,k}]^{-1}.$$

The second term can be rewritten as

$$A_{k+1,l}^{-1}U_{1,k}D_{1,k}^{-1}\,[I+D_{k+1,l}^{-1}B_{k+1,l}U_{1,k}D_{1,k}^{-1}]^{-1}D_{k+1,l}^{-1}.$$

Now is as shown in 1

$$D_{k+1,l}^{-1}B_{k+1,l} = C_{k+1,l}A_{k+1,l}^{-1} = O(z^{-1})$$

while

$$U_{1,k}D_{1,k}^{-1} = O(z^{-1})$$

as said in the beginning of this proof. Also

$$A_{k+1,l}^{-1} = O(z^0) \quad \text{and} \quad D_{k+1,l}^{-1} = O(z^0)$$

as shown under 2 and 3. Hence we may conclude that the whole term is $O(z^0)$ and therefore

$$\begin{aligned} A_{k+1,l}^{-1}X_{k+1,l} &= U_{1,l}D_{1,l}^{-1} - O(z^0) \\ &= O(z^{-1}) - O(z^0) = O(z^0) \end{aligned}$$

which is 4.
Note that by duality also $U_{k+1,l}D_{k+1,l}^{-1}$ will be $O(z^0)$.

□

6. Example

Consider the Markov parameters :

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

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

For order 0 we have that

$$[Y_0 \mid C_0] = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

$$[X_0 \mid A_0] = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

with residuals

$$[S_0 \mid R_0] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{array}\right] z^0 +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] z^{-1} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{array}\right] z^{-2} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] z^{-3} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^{-4} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] (z^{-5} + z^{-6} + z^{-7} + \cdots).$$

With algorithm **MPR** we find for order 1

$$[Y_{1,1} \mid C_{1,1}] = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$[X_{1,1} \mid A_{1,1}] = \left[\begin{array}{cc|ccc} 1 & 0 & z & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & z \end{array}\right]$$

$$[S_{1,1} \mid R_{1,1}] = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right] z^{-1} +$$

$$\left[\begin{array}{cc|ccc} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right] z^{-2} +$$

$$\left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] z^{-3} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^{-4} + \cdots$$

To get the canonical solution we continue as follows

1. $A_{1,1}$ is canonical
2. $A_{1,1}^{-1} X_{1,1} = O(z^{-1})$
3. To make $S_{1,1}$ canonical multiply the first 2 columns by -1.
4. To make $S_{1,1}^{-1} R_{1,1} = O(z^{-1})$ we have to add the first column (with the sign change) to the last column and add the second column (with sign change) to the third column.

We then get

$$V_1 = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ - & - & - & - & - \\ -1 & 0 & z & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & z \end{array}\right] \quad \text{and} \quad \Lambda_1 = \left[\begin{array}{cc|ccc} z & -1 & -1 & 0 & 0 \\ -1 & z & 0 & 0 & -1 \\ - & - & - & - & - \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

while

$$[S_{1,1} \mid R_{1,1}] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{array}\right] z^{-1} +$$

$$\left[\begin{array}{cc|ccc} 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 \end{array}\right] z^{-2} +$$

$$\left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] z^{-3} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^{-4} + \cdots$$

For order 2, algorithm **MPR** will have nothing to do, so that

$$V_2 = \Lambda_2 = I_{p+m}.$$

For order 3, it finds

$$[Y_{1,3} \mid C_{1,3}] = \left[\begin{array}{cc|ccc} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$[X_{1,3} \mid A_{1,3}] = \left[\begin{array}{cc|ccc} -1 & 0 & z & 0 & -1 \\ 0 & 1 & 0 & z^2 & z \\ 0 & 0 & -1 & -1 & z \end{array}\right]$$

and

$$[S_{1,3} \mid R_{1,3}] = \left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^{-1} +$$

$$\left[\begin{array}{cc|ccc} 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{array}\right] z^{-2} +$$

$$\left[\begin{array}{cc|ccc} -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right] z^{-3} +$$

$$\left[\begin{array}{cc|ccc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] z^{-4} + \cdots$$

To make this canonical, only $S_{1,3}$ has to be made canonical. This is obtained by multiplying the second column by -1 and by substracting z times this sign changed column from the first one. This results in

$$V_3 = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ - & - & - & - & - \\ 0 & 0 & 1 & 0 & 0 \\ z & -1 & 0 & z^2 & z \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] \text{ and } \Lambda_3 = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ z & z^2 & 0 & -1 & z \\ - & - & - & - & - \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

and this is also a solution for order 4, which gives

$$V_4 = \Lambda_4 = I_{p+m}.$$

It can easily be checked that each subsystem that is connected with a V_i or Λ_i, $i = 1,2,3,4$ is physically realizable.

7. Conclusions

In this paper algorithm **CMPR** is formulated to give a canonical solution of the minimal (partial) realization problem. Algorithm **CMPR** is found by taking the simpler algorithm **MPR** which computes some solution and extend it with some extra steps to make the solution generated by **MPR** canonical. Algorithm **MPR** is a variant of the algorithms of Dickinson, Morf and Kailath [8] or Anderson, Brasch and Lopresti [1] to generate some solution of the MPR problem. It is given in a form that makes the intimate relation with the work of Antoulas [2] abvious. The extra steps transform the solution into a canonical solution i.e. a solution of the MPR problem that is made unique by some additional constraints. The unimodular matrices V_n generated by algorithm **CMPR** are the building blocks of a matrix continued fraction which is physically realizable. This is an advantage over the unique solution proposed by Antoulas [2] which doesn't necessarily have this property. Some other properties of this canonical solution are stated and proved. The note of theorem 5.7 shows the connection with the work of Bistritz [4] and Bosgra and Van der Weiden [5]. The canonical solution method can be used to define matrix Padé approximations [6]. An explicit relation with the Euclidean algorithm is also given in [6]. The description of an Ada program which implements algorithm **CMPR** can be found in [15].

8. References

[1] B.M.ANDERSON, F.M.BRASCH Jr, P.V.LOPRESTI : The sequential construction of minimal partial realizations from finite input-output data. SIAM J. Control, 13(3), 552-571, 1975.

[2] A.C.ANTOULAS : On recursiveness and related topics in linear systems. IEEE Trans. Autom. Control, AC-31(12), 1121-1135, 1986.

[3] S.BEGHELLI, R.GUIDORZI : A new input-output canonical form for multivariable systems. IEEE Trans. Autom. Control, AC-21, 692-696, 1976 and AC-23, 93, 1978 (comments).

[4] Y.BISTRITZ : Nested bases of invariants for minimal realizations of finite matrix sequences. SIAM J. Control and Optimization, 21(5), 804-821, 1983.

[5] O.H.BOSGRA, A.J.J.VAN DER WEIDEN : Input-Output Invariants for Linear Multivariable Systems. IEEE Trans. Autom. Control, AC-25(1), 20-36, 1980.

[6] A.BULTHEEL, M. VAN BAREL : A matrix Euclidean algorithm and matrix Padé approximation. Report TW104, January 1988, Department of Computer Science, K.U.Leuven.

[7] J.CONAN : A recursive procedure for the solution of the minimal partial realization problem for scalar rational sequences. Rev. Roum. Math. Pures Appl., 30(8), 625-645, 1980.

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

[9] P.A.FUHRMANN : A matrix Euclidean algorithm and matrix continued fraction expansion. Syst. Control Letters, 3, 263-271, 1983.

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

[11] T.KAILATH : Linear systems. Prentice Hall, Englewood Cliffs, NJ, 1980.

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

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

[14] M.VAN BAREL : Nested minimal partial realizations and related matrix rational approximants. Ph.D. Thesis, Department of Computer Science, K.U.Leuven, January 1989.

[15] M.VAN BAREL, A.BULTHEEL : A minimal partial realization algorithm for MIMO systems, VII:An Ada program. Report TW110, March 1988, Department of Computer Science, K.U.Leuven.

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