

PROPERTIES OF MODEL REDUCTION TECHNIQUES BASED ON THE RETENTION OF FIRST- AND SECOND-ORDER INFORMATION

WIESLAW KRAJEWSKI*, ANTONIO LEPSCHY**, UMBERTO VIARO**

In the recent literature on model simplification considerable interest has been focused on the techniques leading to reduced models that match a suitable number of both first-order and second-order information indices. By limiting attention to the information supplied by the Markov parameters and the entries of the impulse-response Gramian, respectively, the paper considers three main approaches. The related algorithms are briefly presented and discussed. Some examples concerning both SISO and MIMO systems illustrate the procedures and compare their performance with that of alternative reduction techniques.

1. Introduction

The problem of model simplification is an important topic in linear system theory and has attracted considerable attention over the past decades. Many methods have been developed to deal with this problem. Among the methods considered earlier on, the most popular and important ones were the aggregation and the Padé techniques (Aoki, 1968; Bultheel and Van Barel, 1986; Davison, 1966). Later on, methods based on the minimization of the L_2 norm of the approximation error were suggested (Wilson, 1970; 1974). In the 1980s, techniques related to the minimization of the Hankel norm (Glover, 1984) and to the truncation of balanced realizations (Moore, 1981) were extensively studied.

In the recent literature, remarkable interest has been devoted to the construction of reduced models that retain a selected set of both first- and second-order information indices of the given original system (Agathoklis and Sreeram, 1990; Anderson and Skelton, 1988; Krajewski *et al.*, 1994a; Yousuff *et al.*, 1985). In the following, we are concerned with the properties of these methods and with the performance of the corresponding reduced-order models.

The first-order information indices are usually provided by the so-called Markov parameters. In the case of continuous-time systems, they correspond to the initial values ($t = 0$) of the impulse response and its derivatives, which in turn coincide with the coefficients of the asymptotic series expansion of the transfer function ($s = \infty$).

* Systems Research Institute, Polish Academy of Sciences, ul. Newelska 6, 01-447 Warsaw, Poland

** Department of Electronics and Informatics, University of Padova, via Gradenigo 6/A, 35131 Padova, Italy

Alternatively, one may consider the coefficients of its MacLaurin expansion ($s = 0$), which are related to the time moments of the impulse response. Expansions at different points can also be taken into account.

The second-order information is usually provided by the entries of the impulse-response Gramian. Alternatively, the so-called system Gram matrix can be used. In the SISO case, the essential second-order information is supplied by the diagonal entries of such matrices because any other entry can be obtained from a diagonal entry and from a suitable number of first-order information indices (Markov parameters or time moments, respectively) (Nagaoka, 1987). In the MIMO case, the situation is not substantially different (Krajewski *et al.*, 1994b).

Reduced models that retain first-order indices only, can easily be obtained via the Padé technique. A serious drawback of this computationally simple procedure is that the stability of the reduced model of a stable system is not ensured. On the contrary, matching second-order information indices ensures stability. The techniques considered in this paper, which force the coincidence of an equal number of both first- and second-order information indices, achieve this objective without appreciably increasing the computational complexity. In addition to that, the McMillan degree of the corresponding transfer function matrices is related to the number of retained indices.

In the sequel, we briefly analyse and compare such techniques. Some illustrative examples are also discussed.

2. Preliminaries

Consider an asymptotically stable time-invariant linear system with p inputs and m outputs described by the k -th order matrix differential equation relating the output vector $y(t)$ to the input vector $u(t)$:

$$y^{(k)}(t) + \sum_{i=0}^{k-1} F_i y^{(i)}(t) = \sum_{i=0}^{k-1} G_i u^{(i)}(t) \quad (1)$$

where the superscript (i) denotes the i -th derivative, $F_i \in \mathbb{R}^{m \times m}$ and $G_i \in \mathbb{R}^{m \times p}$. This equation corresponds to the left matrix fraction description (MFD):

$$\hat{y}(s) = F^{-1}(s)G(s)\hat{u}(s) = \widehat{W}(s)\hat{u}(s) \quad (2)$$

where the hat denotes the Laplace transform, $F(s) = Is^k + F_{k-1}s^{k-1} + \dots + F_0$, $G(s) = G_{k-1}s^{k-1} + \dots + G_0$ and $\widehat{W}(s)$ is the transfer matrix. Assume additionally that a minimal state-space realization of the above system is

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3)$$

$$y(t) = Cx(t) \quad (4)$$

where $x(t) \in \mathbb{R}^n$ and A, B and C have the appropriate dimensions. Note that, if $F^{-1}(s)G(s)$ is irreducible, then n is equal to mk ; otherwise $n < mk$ (Kailath, 1980).

As already observed, the first-order information can be given by the Markov parameters, which are equal to the coefficients $W_i \in \mathbb{R}^{m \times p}$ of the asymptotic expansion of $\widehat{W}(s)$, i.e.

$$\widehat{W}(s) = \sum_{i=1}^{\infty} W_i s^{-i} \tag{5}$$

or the coefficients of the MacLaurin expansion of the system impulse-response matrix $W(t)$, i.e.

$$W(t) = \sum_{i=1}^{\infty} W_i \frac{t^{i-1}}{(i-1)!} \tag{6}$$

with $W_i = W^{(i-1)}(0)$. The Markov parameters can also be expressed in terms of the matrices A, B and C in (3) and (4) as

$$W_i = CA^{i-1}B \tag{7}$$

As regards the second-order information indices, one may refer to the MacLaurin expansion coefficients $R_i \in \mathbb{R}^{m \times m}$ of the output correlation function corresponding to a white noise input:

$$E\{y(t + \tau)y^*(t)\} = \sum_{i=0}^{\infty} R_i \frac{t^i}{i!} \tag{8}$$

where E denotes expected value and the superscript $*$ denotes (conjugate) transpose. The covariance matrices R_i are given by

$$R_i = \lim_{\tau \rightarrow 0} \lim_{t \rightarrow \infty} \frac{d^i E\{y(t + \tau)y^*(t)\}}{dt^i} = CA^iXC \tag{9}$$

where X (steady-state covariance) satisfies the Lyapunov equation:

$$AX + XA^* + BB^* = 0 \tag{10}$$

With reference to the k -th order differential equation (1), the second-order information is also conveyed by the entries of the impulse-response Gramian:

$$P = [P_{i,j}]_{i,j=1,\dots,k} \tag{11}$$

with

$$P_{i,j} = \int_0^{\infty} W^{(i-1)}(t)(W^*)^{(j-1)}(t) dt = P_{j,i}^* \tag{12}$$

Observe that, in the SISO case, for $n = k$ we get

$$P = OW_cO^T \tag{13}$$

where W_c is the controllability Gramian and O is the observability matrix.

It has been shown (Krajewski *et al.*, 1994a) that in the SISO case all the elements of P can be obtained from its k diagonal entries (energies) and from the first k Markov parameters, whereas in the MIMO case the "essential" second-order information is provided by the so-called pseudoenergies, whose number is m^2k , since all the other entries of P can be formed from them and from the entries of the first k matrix Markov parameters.

3. Model Reduction Techniques

By limiting attention to the reduction methods that use the Markov parameters as first-order data and the elements of P as second-order data, three distinct approaches have been followed in the literature.

Approach 1

The q -Markov COVER (covariance equivalent realization) method suggested in (Yousuff *et al.*, 1985) aims at the retention of q Markov parameters and q output covariances of the original system. Precisely, the reduced-order model

$$\dot{x}_r(t) = A_r x_r(t) + B_r u(t) \quad (14)$$

$$y_r(t) = C_r x_r(t) \quad (15)$$

is a q -Markov COVER equivalent realization of system (3), (4) if and only if the matrices A_r, B_r, C_r satisfy the constraints

$$C_r A_r^{i-1} B_r = C A^{i-1} B, \quad i = 1, 2, \dots, q \quad (16)$$

$$C_r A_r^{i-1} X_r C_r = C A^{i-1} X C, \quad i = 1, 2, \dots, q \quad (17)$$

where X is the solution of (10) and X_r the solution of the equation

$$A_r X_r + X_r A_r^* + B_r B_r^* = 0_r \quad (18)$$

According to (Yousuff *et al.*, 1985) model (14), (15) can be obtained using the following algorithm.

Algorithm 1

1. Determine the solution X of (10).
2. Form the q -th observability matrix:

$$\mathcal{O}_q^* = [C^*, A^* C^*, \dots, (A^*)^{q-1} C^*]^* \in \mathbb{R}^{n \times m q} \quad (19)$$

3. Construct a matrix U_1 whose columns form an orthonormal basis for the range of \mathcal{O}_q .
4. Compute the matrices

$$L_r = U_1^* \mathcal{O}_q \quad (20)$$

$$T_r = X \mathcal{O}_q^* U_1 (U_1^* \mathcal{O}_q X \mathcal{O}_q^* U_1)^{-1} \quad (21)$$

5. Compute

$$A_r = L_r A T_r \tag{22}$$

$$B_r = L_r A \tag{23}$$

$$C_r = C T_r \tag{24}$$

The resulting model is not the only q -Markov COVER, as shown in (Anderson and Skelton, 1988) where a parametrization of all the possible q -Markov COVER's in terms of (multivariable) cost-decoupled Hessenberg forms is given. Note, however, that the transformation to such a form becomes a formidable task for large n .

Approach 2

In (Agathoklis and Sreeram, 1990), a method for constructing a reduced- order model of an original SISO system has been suggested which leads, in general, to matching the first q Markov parameters and the $q \times q$ leading principal submatrix of the original impulse-response Gramian. Consequently, the energies (squared L_2 norms) of the impulse response and its first $q - 1$ successive derivatives are retained. This result is achieved using the following algorithm.

Algorithm 2

1. Transform the given n -th order SISO system (A, b, c) into a controllability form $(\bar{A}, \bar{b}, \bar{c})$

$$\bar{A} = C^{-1} A C, \tag{25}$$

$$\bar{b} = C^{-1} b, \tag{26}$$

$$\bar{c} = c C \tag{27}$$

where

$$C = [b, Ab, \dots, A^{n-1}b] \tag{28}$$

2. Compute the impulse-response Gramian P by solving

$$\bar{A}^T P + P \bar{A} + \bar{c}^T \bar{c} = 0 \tag{29}$$

3. Take the $q \times q$ leading principal submatrices P_q and Q_q of P and $Q = \bar{c}^T \bar{c}$, respectively.

4. Compute the controllability form \bar{A}_r of the reduced-order system from the equation

$$\bar{A}_r^T P_q + P_q \bar{A}_r + Q_q = 0 \tag{30}$$

5. Form the corresponding reduced vectors \bar{b}_r and \bar{c}_r as

$$\bar{b}_r = [1, 0, \dots, 0] \in \mathbb{R}^q \tag{31}$$

$$\bar{c}_r = [w_1, w_2, \dots, w_q] \tag{32}$$

where w_i is the (scalar) i -th Markov parameter.

Approach 3

A reduction method, valid for SISO and MIMO systems and based on MFD's, has been presented in (Krajewski *et al.*, 1994a; 1994b). It leads, in general, to matching the first q scalar (SISO case) or matrix (MIMO case) Markov parameters and the first q impulse energies (SISO case) or m^2q pseudoenergies (MIMO case). Such approximant has therefore been named q -Markov ENER (energy equivalent realization) (Krajewski *et al.*, 1994a).

The impulse-response Gramian P of the original system can directly be evaluated from the coefficients of the corresponding MFD as shown in the quoted papers. The reduced model can be found according to the following algorithm.

Algorithm 3

1. Form the submatrices $P_{i,j}$, $i = 1, \dots, q+1$, $j = 1, \dots, q$, of the original impulse-response Gramian P (see (11) and (12)).
2. Determine the denominator coefficient matrices $F_{r,0}, F_{r,1}, \dots, F_{r,q-1}$ of the reduced MFD by solving the set of matrix equations

$$\sum_{i=1}^q F_{r,i-1} P_{i,j} = -P_{q+1,j}, \quad j = 1, \dots, q \quad (33)$$

3. Compute the numerator coefficient matrices $G_{r,0}, G_{r,1}, \dots, G_{r,q-1}$ of the reduced MFD:

$$G_{r,i} = \sum_{j=1}^{q-i-1} F_{r,i+j} W_j + W_{q-i}, \quad i = 0, \dots, q-1 \quad (34)$$

As noted earlier, it follows from (34) that the (matrix) Markov parameters $W_{r,i}$, $i = 1, \dots, q$, of the reduced model always coincide with the corresponding parameters W_i of the original system, whereas the relevant entries of the impulse-response Gramian are matched in the generic case in which the resulting reduced MFD is irreducible.

4. Main Features

- i) The covariance coefficients R_i , defined in (8) and (9) and used in the construction of q -Markov COVER's (cf. Approach 1 of Section 3), are equal to the first (block) column entries of the impulse-response Gramian P defined in (11) and (12).
- ii) In the SISO case, the off-diagonal entries of P can be uniquely determined from its diagonal entries (energies of the impulse response and its derivatives) and from an equal number of Markov parameters. Similarly, in the MIMO case, all the entries of P can be uniquely determined from a suitable number of pseudoenergies and from the entries of the relevant Markov parameter matrices.
- iii) From the previous points, it follows that the q -Markov ENER's (cf. Approach 3 of Section 3) are particular q -Markov COVER's, but q -Markov COVER's exist that are not q -Markov ENER's.

- iv) The q -Markov ENER is generically unique, whereas there is an infinite number of reduced models of order q that match q Markov parameters and q covariances. A parametrization of all q -Markov COVER's is given in (Anderson and Skelton, 1988). The q -Markov COVER obtained according to Approach 1 of Section 3 is one of these, precisely, that corresponding to the choice of the matrix U_1 .
- v) The external stability of the reduced models obtained according to all three algorithms of Section 3 is always ensured even if the resulting system (A_r, B_r, C_r) is not minimal or the resulting MFD is not irreducible. In these cases, all the relevant Markov parameters are matched, whereas not all the covariances or energies are matched.
- vi) Approach 2 of Section 3 could be extended to MIMO systems according to the procedure used in (Krajewski *et al.*, 1994b).
- vii) Approach 3 of Section 3 leads, in general, to an irreducible MFD of order m_q . This, however, can easily be converted into a minimal state-space representation. In fact, it can be realized in the block observability form, where A_r is the $m_q \times m_q$ block companion form with the first block row equal to $[-F_{q-1}, -F_{q-2}, \dots, -F_0]$, B_r is equal to $[W_q^T, \dots, W_1^T]^T$ and $C_r = [0, \dots, 0, I]$.

5. Examples

In this section, we consider three rather simple examples and compare the results obtained using the algorithms of Section 3 with those obtained using different techniques (L_2 -optimal, Moore's balancing). For this purpose, a reference is made to the relative squared L_2 error norm:

$$\delta = \frac{\|W - W_r\|^2}{\|W\|^2} \tag{35}$$

Example 1. Let us first refer to the elementary SISO example already considered in (Kabamba, 1985) to show that Moore's balancing approach does not always lead to models that are good in the L^2 sense. The original system is described by the triplet

$$A = \begin{bmatrix} 0.005 & -0.99 \\ -0.99 & -5000.0 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 100 \end{bmatrix}, \quad c = [1 \quad 100] \tag{36}$$

or, alternatively, by the transfer function

$$\widehat{W}(s) = \frac{10001s + 4852}{s^2 + 5000.005s + 24.0199} \tag{37}$$

All the algorithms of Section 3 can be used and lead to the first-order model

$$A_r = -4951.5, \quad b_r = 10001, \quad c_r = 1 \tag{38}$$

or, equivalently,

$$\widehat{W}_r(s) = \frac{10001}{s + 4951.5} \quad (39)$$

which is very close to the optimal model in the L_2 sense (Kabamba, 1985). In fact, the value of δ for (38) and (39) is 0.00956, whereas $\delta = 0.00851$ for the L_2 -optimal model. Note that, in this case, Moore's balancing method leads to very bad value of $\delta = 0.99$.

Example 2. Let us consider now the 6-th order SIMO system (Gawronski and Juang, 1990) described by the triplet:

$$A = \begin{bmatrix} -0.21053 & -0.10526 & -0.0007378 & 0 & 0.0706 & 0 \\ 1 & -0.03537 & -0.000118 & 0 & 0.0004 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -605.16 & -4.92 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -3906.25 & -12.5 \end{bmatrix}$$

$$B^T = \begin{bmatrix} -7.211 & -0.05232 & 0 & 794.7 & 0 & -448.5 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0.000334 & 0 & -0.007728 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (40)$$

The 4-th order state-space model obtained using Algorithm 1 is as follows:

$$A_r = \begin{bmatrix} -14.15 & -1.794 & -1.515 & 5480.99 \\ 0.709 & -0.123 & -0.076 & -0.15 \\ 0.017 & 0.71 & -0.016 & 0.71 \\ -0.705 & -0.105 & -0.074 & -0.132 \end{bmatrix}, \quad B_r = \begin{bmatrix} -43.107 \\ -7.558 \\ 0.014 \\ -2.637 \end{bmatrix}$$

$$C_r = \begin{bmatrix} 0 & 0.707 & 0.018 & -0.707 \\ 0 & 0 & 0.999 & 0.025 \end{bmatrix} \quad (41)$$

A model with the McMillan degree equal to 4 can also be obtained using Algorithm 3 with $q = 2$. The MFD of the reduced model turns out to be given by

$$F_r(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} s^2 + \begin{bmatrix} 14.206 & -3873.19 \\ -0.0025 & 0.219 \end{bmatrix} s + \begin{bmatrix} 3875.738 & -135.52 \\ 0.0257 & 0.111 \end{bmatrix}$$

$$G_r(s) = \begin{bmatrix} -3.479 \\ -0.052 \end{bmatrix} s + \begin{bmatrix} 110.1 \\ -7.212 \end{bmatrix} \quad (42)$$

which can be directly realized in the (minimal) block observability form (cf. item (vii) of Section 4)

$$\begin{aligned}
 A_{ro} &= \begin{bmatrix} -14.206 & 3873.19 & -3875.738 & 135.52 \\ 0.0025 & -0.219 & -0.0257 & -0.111 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & B_{ro} &= \begin{bmatrix} -41.883 \\ -7.209 \\ -3.479 \\ -0.052 \end{bmatrix} \\
 C_{ro} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. & & (43)
 \end{aligned}$$

Models (41) and (43) are equivalent. The corresponding value of δ is $8.9473 \cdot 10^{-6}$ which is of the same order of magnitude as the value for Moore's model ($\delta = 7.4618 \cdot 10^{-6}$), whereas the L_2 -optimal model is characterized by $\delta = 5.8262 \cdot 10^{-6}$. Note that the δ -values corresponding to different reduction techniques considered in (Gawronski and Juang, 1990) are appreciably larger.

Example 3. Let us consider now a MIMO system with 2 inputs and 2 outputs described by the triplet

$$\begin{aligned}
 A &= \begin{bmatrix} -15 & 4000 & -4000 & 100 \\ 0.002 & -0.3 & -0.03 & -0.1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & B &= \begin{bmatrix} -40 & -3838 \\ -9.993 & -0.72 \\ -4 & -10 \\ 0.05 & -1 \end{bmatrix} \\
 C &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & & (44)
 \end{aligned}$$

Both algorithms, i.e. Algorithm 1 and Algorithm 3 lead to the same reduced model:

$$A_r = \begin{bmatrix} -0.1854 & -0.1027 \\ 0.5281 & -0.0139 \end{bmatrix}, \quad B_r = \begin{bmatrix} -4 & -10 \\ -0.05 & -1 \end{bmatrix}, \quad C_r = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (45)$$

The corresponding value of δ is 1.21378, whereas Moore's model and L_2 -optimal model are characterized by $\delta = 0.0753$ and $\delta = 0.075$, respectively. However, the time of computations is much larger for the latter models.

6. Conclusions

The reduction methods aiming at the retention of both first-order and second-order information indices ensure the stability of the model of a stable original system as well as a good fit of its response. This result is achieved in a fairly simple way, i.e. by solving sets of linear equations.

The three approaches considered in this paper refer to the first-order information supplied by the Markov parameters and to the second-order information given by appropriate entries of the impulse-response Gramian. The first and the third approach of Section 3 can also be applied to MIMO systems, whereas the second holds for SISO systems only.

The guidelines of the algorithms corresponding to the three approaches have been given; the related MATLAB computer programs are available on request. As shown by the examples of Section 5, they lead to satisfactory results.

The methods can be extended in a rather easy manner to the cases in which the first-order information is provided by different expansion coefficients, e.g. the Taylor expansion of $\widehat{W}(s)$ about suitable points instead of the asymptotic expansion, and the second-order information by the energies of suitable system responses instead of the impulse-response energies.

References

- Agathoklis P. and Sreeram V. (1990): *Identification and model reduction from impulse response data*. — Int. J. Systems Sci., v.21, No.8, pp.1541–1552.
- Anderson B.D.O. and Skelton R.E. (1988): *The generation of all q -Markov covers*. — IEEE Trans. Circ. Systems, v.CAS-35, No.4, pp.375–384.
- Aoki M. (1968): *Control of large-scale systems by aggregation*. — IEEE Trans. Aut. Contr., v.AC-13, No.3, pp.246–253.
- Bultheel A. and Van Barel M. (1986): *Pade techniques for model reduction in linear system theory: A survey*. — J. Comp. Appl. Math., v.14, No.3, pp.401–438.
- Davison E.J. (1968): *A method for simplifying linear dynamic systems*. — IEEE Trans. Aut. Contr., v.AC-11, No.1, pp.93–101.
- Gawronski W. and Juang J. (1990): *Model reduction for flexible structures*. In: Control and Dynamic Systems, (C.T. Leondes, Ed.). — San Diego: Academic Press, pp.143–222
- Glover K. (1984): *All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds*. — Int. J. Contr., v.39, No.6, pp.1115–1193.
- Kabamba P.T. (1985): *Balanced gains and their significance for L_2 model reduction*. — IEEE Trans. Aut. Contr., v.AC-31, No.7, pp.796–797.
- Kailath T. (1980): *Linear Systems*. — Englewood Cliffs, NJ: Prentice-Hall.
- Krajewski W., Lepschy A. and Viaro U. (1994a): *Approximation of continuous-time linear systems using Markov parameters and energy indices*. — Archives of Control Sciences, v.3 (XXXIX), No.1-2, pp.21–30.
- Krajewski W., Lepschy A. and Viaro U. (1994b): *Reduction of linear continuous-time multivariable systems by matching first- and second-order information*. — IEEE Trans. Aut. Contr., v.39, No.10, pp.2126–2129.
- Moore B.C. (1981): *Principal components analysis in linear systems: controllability, observability and model reduction*. — IEEE Trans. Aut. Contr., v.AC-26, No.1, pp.17–32.
- Nagaoka H. (1987): *Mullis-Roberts-type approximation for continuous-time linear systems*. — Electronics and Communications in Japan, part 1, v.70, No.10, pp.41–52.

- Wilson D. A. (1970): *Optimum solution of model reduction problem.* — Proc. IEE Part D - Control Theory Appl., v.117, No.6, pp.1161-1165.
- Wilson D. A. (1974): *Model reduction for multivariable systems.* — Int. J. Contr., v.20, No.1, pp.57-64.
- Yousuff A., Wagie D.A., and Skelton R.E. (1985): *Linear system approximation via covariance equivalent realizations.* — J. Math. Anal. Appl., v.106, No.1, pp.91-115.

Received: December 6, 1944