ANNALS of the ORADEA UNIVERSITY.
Fascicle of Management and Technological Engineering, Volume VII (XVII), 2008

ON THE MODEL AND CONTROLLER REDUCTION OF LARGE-SCALE DYNAMICAL SYSTEMS

Dumitru NICOARA
Transilvania University of Brasov
tnicoara@unitbv.ro

Keywords: Model and controller reduction, Large-scale dynamic systems

Abstract: In this paper we make a comparative study of recently proposed methods for incorporating closed-loop system information into the plant and controller reduction process for large-scale systems. The aim is to obtain a systematic way to reduce the order of large-scale controllers using a reasonable amount of computational effort and storage that, involving efficient algorithms, would yield controllers feasible for real-time implementations.

1. INTRODUCTION

In the context of control synthesis the determination of the model of the plant the reduced order model and the design of the control law are not independent problems.

The model and controller reduction methods can be divided into two different classes [1]: Direct and Indirect. Figure 1 shows the direct method design from a high-order plant to a low-order controller. With direct methods, the parameters defining a low-order controller are computed by employing an optimization technique. With indirect methods, two design pathways are possible. A high-order controller can first be designed, and then a procedure can be used to reduce the controller complexity. In Figure 1, this pathway is illustrated in the upper right. Following a different procedure, a reduced-order plant can be found prior to the controller design, and then a reduced-order controller is designed for the reduced-order plant. In Figure 1, this pathway is illustrated in the lower left. There are many issues that arise with the indirect approach. In the overall design process, the plant model approximation is carried out at an early step of the design process, without the benefit of pertinent information about the low-order controller.

As discussed in [3], a good approximation of the plant requires knowledge of the controller. It is important to understand that the problem of controller reduction (closed-loop) is distinct from the problem of model reduction (open-loop), since it is after all, closed-loop performance that should be approximated.

Recently, Antoulas, et al. [2], [4] proposed a method for incorporating closed-loop system information into the plant and controller reduction process for large-scale systems, using rational interpolation through the poles of the large-scale closed-loop system and the large-scale controller.

Most computational methods currently employed for controller reduction [9] cannot effectively handle very large-scale problems that exhibit some sparsity. They frequently involve the solution of Riccati equations and linear matrix inequalities (LMI) in the controller reduction process. It is known that current methods exhibit computational cost associated with the algorithms on the order $O(n^3) - O(n^6)$ operations, where n is the number of states, thus becoming impractical for large-scale applications [2], [4]

A systematic way for reducing the order of large-scale controllers using a reasonable amount of computational effort and storage, that is, involving efficient algorithms, would yield controllers feasible for real-time implementations.

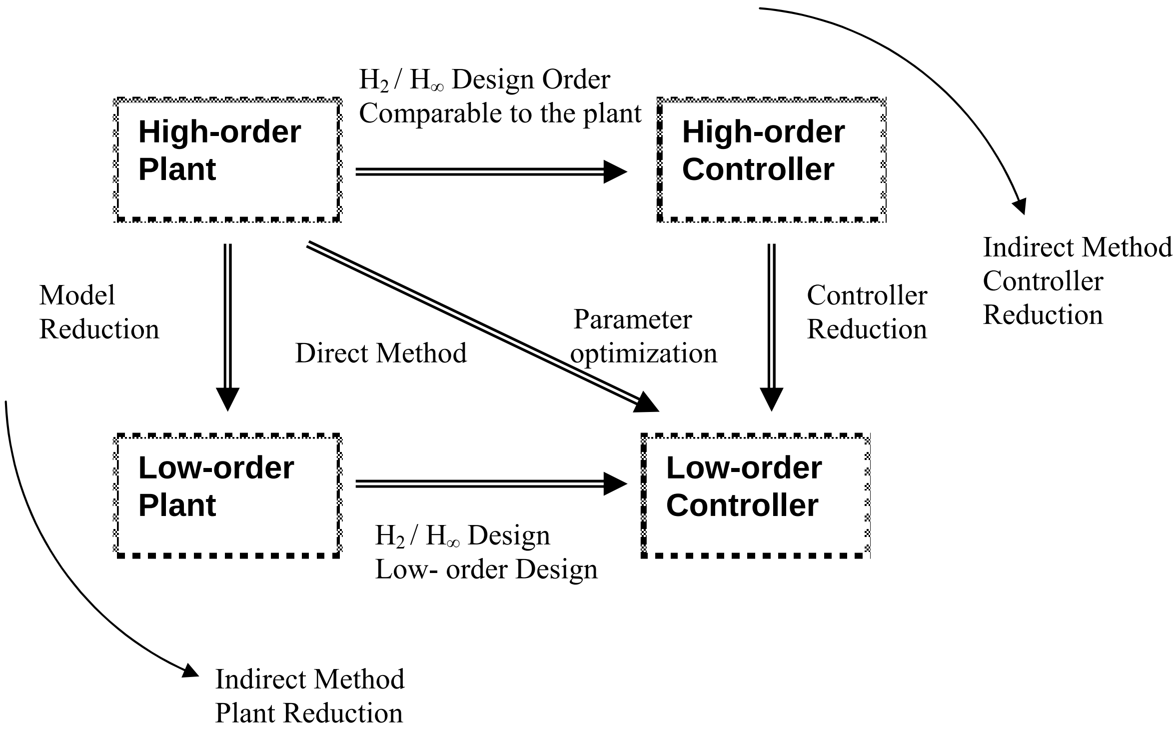

Figure 1 Direct and indirect Approaches for model and controller reduction

2. MODEL REDUCTION PROBLEM

Given a dynamic system model *G* of usually high order n, a model reduction method is a procedure that yields some approximate model G_r of order $n_r << n$. We note that by abuse of notation, both the underling dynamical system and its transfer function (TMF) are denoted by $G(s)$. The quality of the approximation is usually evaluated by looking at the model reduction error, that is, the signal obtained as the difference between the outputs of the original system and the outputs of the reduced order model driven by the same input signal. The goal is to produce a low dimensional systems that has similar response characteristic as the original system with far lower storage requirement and evaluation time. The resulting reduced-order model might be used to replace the original system as a component in a large simulation or it might be used to develop a low dimensional controller suitable for real time applications.

The unifying feature of all model and controller reduction techniques is that they are obtained by means of a projection. This process applies to controller and the closed-loop system as well.

Also, for an efficient reduction algorithm, one has to guarantee that:

- the dimension of reduced-order model is $r << n$;
- the behavior of reduced-order model approximates the original with certain accuracy, i.e., there is a small error bound on $\|\mathbf{y}(t) - \mathbf{y}_r(t)\|_{2;\infty}$;
- the procedure is computationally stable and efficient.

Consider a linear, continuous time- invariant (LTI) large-scale dynamical system or plant, single-input single-output (SISO)

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,u(t) \\ y(t) &= \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\,u(t),\ \mathbf{x}(0) = \mathbf{x}_0 \end{aligned} \quad \Leftrightarrow \quad \sum = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \in \mathrm{R}^{(n+p)\times(n+m)} \qquad (1)$$

where the state vector $\mathbf{x}(t) \in \mathrm{R}^n$, and $u(t), y(t)$ are the scalar input and output, respectively. The transfer function is

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}, \tag{2}$$

In model reduction we are faced with the problem of finding a reduced-order LTI system

$$\begin{aligned} \dot{\mathbf{x}}_r(t) &= \mathbf{A}_r\mathbf{x}_r(t) + \mathbf{B}_r\mathbf{u}(t) \\ \mathbf{y}_r(t) &= \mathbf{C}_r\mathbf{x}_r(t) + \mathbf{D}_r\mathbf{u}(t) \end{aligned} \tag{3}$$

of order $r, r << n$, and associated TMF

$$G_r(s) = \mathbf{C}_r(s\mathbf{I} - \mathbf{A}_r)^{-1}\mathbf{B}_r + \mathbf{D}_r \tag{4}$$

which approximate $G(s)$ such that the following properties are satisfied:

- the predicted input-output behavior is close, e.g., $\|G - G_r\|_\infty$ or $\|G - G_r\|_2$ are small.
- system properties, like stability, passivity, are preserved.
- the procedure is computationally efficient.

3. METHODS FOR MODEL AND CONTROLLER ORDER REDUCTION

3.1 Lyapunov balancing methods

The most commonly used model reduction scheme is the so-called balanced model reduction, which was first introduced by Mullis and Roberts [6] and then in a systems and control framework by Moore [5] The main idea of this technique is a change of the state coordinate basis, called a balancing transformation, such that the controllability and observability grammians are both equal to some diagonal matrix, Σ_d, where the magnitudes of the diagonal entries of the grammians reflect the contributions of different entries of the state vector of the system. This is achieved by simultaneously diagonalizing the reachability, $\mathcal{G}_C$, and the observability, $\mathcal{G}_O$, grammians, which are solutions to the reachability and the observability Lyapunov equations:

$$\mathbf{A}\mathcal{G}_C + \mathcal{G}_C\mathbf{A}^T + \mathbf{B}\mathbf{B}^T = 0 \ , \ \ \mathbf{A}^T\mathcal{G}_O + \mathcal{G}_O\mathbf{A} + \mathbf{C}^T\mathbf{C} = 0 \tag{5}$$

Although balanced model reduction and its variants have nice system theoretic properties, such as preservation of stability and computation of an error bound, they become computationally prohibitive for large-scale systems. This drawback stems from the fact that they require dense matrix actorizations, such as solving two Lyapunov equations, and therefore the computational cost on the order O(n^3) and storage of order O(n^2) becomes impractical for systems of order n > 1000.
Lyapunov balancing methods are now explored with applications to model reduction.
This discussion will support the application of the concepts to the problem of closed-loop controller reduction by balanced truncation.

Consider a stable LTI system model, G(s), given by its state-space realization and transfer function as described in Eq. (1).

The reachable, observable and stable system **G**(s) is called Lyapunov-balanced if

$$\mathcal{G}_C = \mathcal{G}_O = \Sigma_d = diag\left(\sigma_1 \mathbf{I}_{m_1} + \sigma_2 \mathbf{I}_{m_2} + ... + \sigma_p \mathbf{I}_{m_p}\right) \quad (7)$$

where Σ_d is a diagonal matrix with $\sigma_1 > \sigma_2 > \cdots \sigma_q$, and m_i are the multiplicities of σ_i ,so that $m_1 + \cdots + m_q = n$.

The following results formally describe the balancing method [7]:

Proposition 1. Consider a stable system $\mathbf{G}(s)$ and suppose $\mathbf{G}(s) = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ is a balanced realization, i.e., its controllability and observability grammians are equal and diagonal given by $\mathcal{G}_C = \mathcal{G}_O = \Sigma_d$, which satisfy the following Lyapunov equations

$$\mathbf{A}\mathbf{\Sigma}_d + \mathbf{\Sigma}_d \mathbf{A}^* + \mathbf{B}\mathbf{B}^T = 0 \quad (8)$$

$$\mathbf{A}^*\mathbf{\Sigma}_d + \mathbf{\Sigma}_d \mathbf{A} + \mathbf{C}^*\mathbf{C} = 0 \quad (9)$$

where $(\cdot)^*$ denotes the complex conjugate transpose of a matrix.

Partitioning the balanced gramians as $\mathbf{\Sigma}_d = \begin{bmatrix} \mathbf{\Sigma}_1 & 0 \\ 0 & \mathbf{\Sigma}_2 \end{bmatrix}$ with

$\mathbf{\Sigma}_1 = diag\left(\sigma_1 \mathbf{I}_{m_1} + \sigma_2 \mathbf{I}_{m_2} + ... + \sigma_k \mathbf{I}_{m_k}\right)$ and $\mathbf{\Sigma}_2 = diag\left(\sigma_{k+1} \mathbf{I}_{m_{k+1}} + ... + \sigma_q \mathbf{I}_{m_q}\right)$

and partitoning the balanced system accordingly

$$\mathbf{G}(s) = \begin{bmatrix} \mathbf{A}_b & | & \mathbf{B}_b \\ - & | & - \\ \mathbf{C}_b & | & \mathbf{D}_b \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & | & \mathbf{A}_{12} & \mathbf{B}_1 \\ - & - & - & - \\ \mathbf{A}_{11} & | & \mathbf{A}_{22} & \mathbf{B}_1 \\ \mathbf{C}_1 & | & \mathbf{C}_2 & \mathbf{D} \end{bmatrix}$$

where the dimensions are $\mathbf{A}_{11} \in \mathrm{R}^{k \times k}$, $\mathbf{B}_1 \in \mathrm{R}^{k \times m}$,$\mathbf{C}_1 \in \mathrm{R}^{p \times r}$,$\mathbf{D}_1 \in \mathrm{R}^{p \times m}$.

Then, the reduced order model $\mathbf{G}_r(s) = \begin{bmatrix} \mathbf{A}_{11} & | & \mathbf{B}_1 \\ - & | & - \\ \mathbf{C}_1 & | & \mathbf{D} \end{bmatrix}$ obtained by truncation is asymptotically stable, balanced, minimal (controllable and observable) and satisfies

$$\left\| \mathbf{G}(s) - \mathbf{G}_r(s) \right\|_{H_\infty} \le 2\left(\sigma_{k+1} + \cdots \sigma_q\right) \quad (10)$$

Equality holds if $\mathbf{\Sigma}_2 = \sigma_q \mathbf{I}_{m_q}$.

In order to compute the simultaneous diagonalization of $\mathcal{G}_C$ and $\mathcal{G}_O$, several algorithms have been proposed in the literature [2].

3.2 Krylov-Based Model Reduction

SVD-based methods are not suitable for large-scale systems due to the use of dense matrix factorizations of O(n^3) and storage of O(n^2). As an alternative, model reduction techniques that rely on matrix-vector multiplication and that can be implemented iteratively in a numerically efficient manner become good choices for large-scale systems. Krylov subspace techniques provide this alternative [2].

The key ingredient of Krylov-based methods is moment matching. The idea is to match moments of the original higher-order model by the moments of a lower-order model. This is achieved by iteratively constructing matrices that span certain (generalized) Krylov subspaces of $\mathbf{A}$ and $\mathbf{B}$ (controllability subspace) and/or $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{C}^{\mathrm{T}}$ (observability subspace).

Model reduction by Krylov techniques is not based on minimization, as with the SVD-based reduction methods. Instead Krylov techniques are based on moment matching, where one attempts to match the coefficients of a power series expansion of the transfer function for the original and reduced-order models

$$\mathbf{G}(s) = \mathbf{C}\,(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}; \quad \mathbf{G}_r(s) = \mathbf{C}_r\,(s\mathbf{I}_r - \mathbf{A}_r)^{-1}\mathbf{B}_r \qquad (11)$$

If the transfer functions in Eq. 11 are expanded in a Laurent series around a given point in the complex plane, $\sigma \in C$, then

$$G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \sum_{i=0}^{\infty} m_i(\sigma - s)^i \qquad (12)$$

where the coefficients denoted by m_i are called moments of the system at a point σ. Similarly, one obtains the moments of the reduced-order model as m_{ri} . The main idea of model-order reduction by moment matching is to match a given number of moments as $m_i = m_{ri}$, $i = 1,2,...,l$, $l << n$ of the original and reduced-order transfer functions. Several special cases of the moments can be determined depending on the location of the expansion points in the complex plane. In the case of $\sigma = \infty$, the moments are the well-known Markov parameters of the system. The Markov parameters represent the values of the zero-state impulse response, or transfer function G(s), and subsequent derivatives of the impulse response at t = 0. Since matching the Markov parameters emphasizes the behavior at t = 0, the reduced-order model may be dominated by rapid decaying dynamics, not representing accurately the behavior at later time. In the frequency domain, once can show good matching of the frequency response of the system at high frequencies. A power series expansion can also be performed about σ = 0. In this case the reduced order model will be a good approximation to the steady-state response of the original system. As a summary, Table 1 is constructed based on Eq. (12).

Frequency to be Approximated	Power Series Expansion of the TF	i^{th} Moment
About $\sigma = \infty$	$\sum_{i=1}^{\infty} m_{-i} s^{-i}$	$\mathbf{C}\mathbf{A}^{i-1}\mathbf{B}$
About $\sigma = 0$ (Padé)	$\sum_{i=1}^{\infty} m_{i-1} s^{i-1}$	$-\mathbf{C}\{\mathbf{A}^{-1}\}^i\mathbf{B}$
About $s = \sigma$ (shifted Padé)	$\sum_{i=1}^{\infty} m_{i-1}(s-\sigma)^{i-1}$	$\mathbf{C}\{(\sigma\mathbf{I} - \mathbf{A})^{-1}\}^i\mathbf{B}$

Table 1 Expansions and moments to be matched.

A straightforward approach to constructing reduced-order models can be obtained by explicitly computing $2r$ moments of the original system as in Table 1, where r is the size of the reduced-order model. Then, the frequency response of the reduced-order system is forced to correspond to the selected moments. This can be viewed as a selection of the coefficients for the numerator and denominator of the reduced-order transfer function through the solution of a linear system involving Hankel matrices. Unfortunately, numerical drawbacks of the explicit moment-matching can occur, such as ill-conditioned
Hankel matrices, sensitivity of the partial realization, moment scaling, and the stability of the approximation.

5. CONCLUDING REMARKS

In state-space, truncation of the state vector is the “natural” choice for obtaining a reduced order model. The fundamental question is what states are “important” and should be kept in the reduced order model? In the context of model reduction, and given a state-space realization (*A,B,C,D*) of order n, one might wonder whether the given realization is minimal in the sense that there exists no other transfer equivalent realization (A_r,B_r,C_r,D_r) of order nr with n_r smaller than n. The answer to this question relies on the concepts of controllability and observability.

As discussed, a good approximation of the plant requires knowledge of the controller. It is important to understand that the problem of controller reduction (closed-loop) is distinct from the problem of model reduction (open-loop), since it is after all, closed-loop performance that should be approximated.

REFERENCES

[1] Anderson, B.D.O., Y. Liu. Controller reduction: Concepts and approaches, IEEE Transactions of Automatic Control, 34:802–812, 1989.
[2] Antoulas, A.C., A new result on positive real interpolation and model reduction Systems and Control Letters, 54:361–374, 2005.
[3] Enns, D.F., Model reduction for control systems design. PhD thesis, Dep. Aeronaut. Astronaut., Stanford University, Stanford, C.A., 1984.
[4] Gugercin S., A.C. Antoulas. On model reduction by the Lanczosand the rational Krylov methods, ECE Dep., Rice University. Technicalreport, 2002.
[5] Moore, B.C., Principal component analysis in linear systems: Controllability, observability, and model reduction. IEEE Trans. Automat. Control, AC-26, 17–32 (1981).
[6] Mullis, C., Roberts, R.A.: Synthesis of minimum roundoff noise fixed point digital filters. IEEE Trans. Circuits and Systems, CAS-23:9, 551–562 (1976).
[7] D. D. Nicoara: ON THE COMPUTATIONAL ASPECTS IN MODEL REDUCTION OF DYNAMICAL SYSTEMS, International Workshop ″Advanced Researches in Computational Mechanics and Virtual Engineering″ pag.149-155, CD Edition, ISBN 973-635-823-2, 2006, Brasov, Romania.
[8] Nicoara D.: Large-Scale Dynamical Systems and Pseudospectra , Annals of the Oradea University, Volume V (XV), pag. 127-134, 2005
[9] Zhou,K., J.C. Doyle, and K. Glover. Robust and Optimal Control.Prentice Hall, 1996.