Generation of Linear Models using Simulation Results

Georg Otte, Sven Reitz, Joachim Haase
Fraunhofer Institute for Integrated Circuits, Branch Lab Design Automation Dresden
Zeunerstraße 38, D-01069 Dresden, Germany
Phone: +49 (0)351 4640-705, Fax +49 (0)351 4640-703
e-mail: {otte, reitz, haase}@eas.iis.fhg.de

Abstract: A two step approach for the generation of linear models using impulse response values computed by simulation will be described. In the first step a model of a "middle" order is determined using a realization based method. In the second step the order of this model will be reduced using known model order reduction methods. The objective of this paper is to show, that this kind of modeling can be realized by a combination of well known methods and leads to accurate models of very low order.

1. Introduction

In many applications very fast models of as low order as possible are required. In the last time very powerful methods have been known which are used for the order reduction of high dimensional models [1]. Basing on the knowledge of the matrices \( \Sigma = (A, B, C, D) \) of the high dimensional model these methods determine a suitable subspace and project transient processes of the system into this subspace. On the powerful methods the subspace is determined using balanced truncation or as a Krylov-subspace, respectively. If the matrices \( \Sigma = (A, B, C, D) \) are not available but only the outputs of the system, e.g. computed via simulation, these methods cannot be used.

The aim of this paper is the generation of linear time discrete models from simulation results. Given the input and output values of a process the determination of a model is called an identification task. But the existing identification methods do not use all advantages of simulated data. For process identification there are established solutions, known as for example the "prediction error method PEM" or the "instrumental variable method IVM", respectively. These identification methods are developed for the determination of models from experimental, i.e. disturbed measurements. They are specialized for the handling of disturbances in the process output which occur in all real measurements. But if there is a high number of inputs and outputs or a high systems order their numerical properties are very poor [2]. In the case of nearly undisturbed data like simulation results these methods are without advantage.

Recently developed subspace-methods [2], [3], [4] have better numerical properties. Optimization steps are not needed. The multi-input multi-output case can be handled in the same way as the single-input single-output one. Initial values are not needed. An important advantage in practice is the simple determination of the model structure by only one parameter, the order of the model [6]. If it is possible to determine the step-response or impulse-response directly, simplified (basic-) versions of these methods can be used. These are much more robust and their application area can be extended to large systems.

Our approach to determine time discrete state-space-models directly from impulse response values follows this idea. Methods used here are closely related to the realization theory [2]. Partial realizations that are realizations of lower order approximate the given process data. The approximation is excellent at the beginning of a transient process, but it is not so good at the end. Often the partial realization has to be of relative high ("middle") order to obtain models with small static errors.

Therefore it will be suggested in such cases to reduce the order of the analytic "middle" order model in a second step with known standard methods like modal or balanced truncation, with or without singular perturbation, respectively [5].

This paper is organized like this: In section 2 the basic relations between system matrices, observability, controllability, Hankel matrices and the impulse response are presented. The construction of a suitable Hankel matrix from given impulse response values and the determination of the model from these Hankel matrix will be discussed. In section 3 methods for model reduction as second step of modeling are mentioned. Section 4 contains examples and section 5 concludes the paper.
2. First step: Realization based determination of a model

Suppose \( \Sigma = (A, B, C, D) \) is a linear time discrete system

\[
x_{t+1} = Ax_t + Bu_t \\
y_t = Cx_t + Du_t
\]

Here \( x_t \) is an \( N \)-dimensional state vector, \( u_t \) is a \( m \)-dimensional input vector and \( y_t \) is a \( q \)-dimensional output vector. The matrix \( A \) is \( N \times N \), \( B \) is \( N \times m \), \( C \) is \( q \times N \) and \( D \) is \( q \times m \). To this system belong an observability matrix \( O_p \), a controllability matrix \( R_p \) and a Hankel matrix \( H_p \).

\[
O_p = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{p-1} \end{bmatrix}, \quad R_p = \begin{bmatrix} B & AB & \cdots & A^{p-1}B \\ C & CA & \cdots & CA^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ C & CA^p & \cdots & CA^{p-2} \end{bmatrix}
\]

\[
H_p = O_p R_p = \begin{bmatrix} CB & CAB & \cdots & CA^{p-1}B \\ CAB & CA^2B & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ CA^{p-1} & CA^p & \cdots & CA^{2p-2}B \end{bmatrix}
\]

If \( \Sigma \) is full controllable and observable and if \( p \geq N \), all these matrices are of the same rank \( N \).

The elements of the Hankel matrix are \( p \times m \) matrices \( g_k = CA^{k-1}B, k \geq 1 \),

\[
H = \begin{bmatrix} g_1 & g_2 & g_3 & \cdots \\ g_2 & g_3 & g_4 & \cdots \\ g_3 & g_4 & g_5 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}
\]

which are equal to the impulse response values of \( \Sigma \) for \( k \geq 1 \).

If the \( g_k \) are known, e.g. as a simulation result, matrices \( H \) can be constructed without explicit knowledge of the matrices \( A, B, C, D \) and the Order \( N \).

Given the \( g_k \), the first task is to assemble a Hankel matrix \( H \) having a sufficient high numeric rank \( N_n \). The second task is to determine the matrices \( \tilde{\Sigma} = (\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}), n \leq N_n \) of an approximate system.

2.1. Construction of a Hankel matrix and choosing the model order

If the system matrices and the system order are known, the Hankel matrix of correct order \( N \) can be constructed easily. Like in all identification processes, these information is not available here. In our case the matrix \( H \) is to assemble by a sufficient number of block-columns and block-rows of the \( g_k \) by numeric experiments. It will be assumed that there exists a large number \( K \) of instants of the impulse response \( \{g_k, k = 1, \cdots K\} \). Thus the numeric rank of \( H \) is not limited by the columns number but it can be limited by the rows number. To find a sufficient number of rows, these can be increased step by step until the rank does not more grow or it grows only slowly,

Fig 1 numeric rank of Hankel matrices of different rows (example)  
Fig 2 singular values of Hankel matrices of a system with a high number of outputs (example)

caused by small numeric disturbances. In this way a Hankel matrix of sufficient numeric rank can be established. Fig. 1 shows the rank values versus row block numbers of an example. In such cases, the order of the model looked for can be chosen as high as the rank.
In other applications, if there exists a very high number of outputs, may be, many thousands, one block row can be sufficient. The numeric rank of the Hankel matrix can be as high as the output number. Here the model order cannot be chosen as high as the rank. Fig. 2 shows the singular values of such an Hankel matrix of an example with approximately 100 outputs. As can be seen, the use of more block rows does not really change the relations. The large number of singular values which are very small and do not decrease further on is caused by numeric disturbances. In such cases the model order should be chosen near the bottom of the left part of the graph, before the blip.

In other cases, if the rank of the Hankel matrix in the range of interest increases permanently but slowly and the singular values decrease regularly slow, too, the sufficient rank and the system order are to determine by numerical experiments.

The choice of a relative high model order is not critical. Superfluous states will be removed in the next step.

2.2. Determination of the model from Hankel matrix

From \( H \) it is possible to compute two matrices \( \hat{O} \) and \( \hat{R} \) with any full rank factorization, so that \( H = \hat{O}\hat{R} \).

Different factorizations lead to different realizations of the same system. A well suitable factorization method is the singular value decomposition (svd) which calculates matrices \( U, S, V \) with 
\[
USV = H.
\]

Furthermore, the matrices \( \hat{O} \) and \( \hat{R} \) calculated by svd are well suitable to carry out an approximation of the Hankel matrix if the numeric rank is too large. In
\[
H = \begin{pmatrix} \tilde{U} & U_{\text{neg}} \end{pmatrix} \begin{pmatrix} \tilde{S} & 0 \\ 0 & S_{\text{neg}} \end{pmatrix} \begin{pmatrix} V' \\ V_{\text{neg}} \end{pmatrix} = \tilde{U}S\hat{V} + U_{\text{neg}}S_{\text{neg}}V_{\text{neg}}.
\]

the components indexed by "neg" are connected to small singular values \( s_i \), \( n < i \leq N_y \). These components are negligible. This can be done, by setting \( s_i = 0 \), \( n < i \leq N_y \).

From the approximation \( \hat{H} \) of rank \( n \) of the Hankel matrix
\[
\hat{H} = \hat{U}\hat{S}\hat{V}
\]
appropriate observability and controllability matrices can be found, e.g.: \( \hat{O} = \hat{U}\sqrt{S} \) and \( \hat{R} = \sqrt{S}\hat{V} \).

The approximations \( \hat{O} \) and \( \hat{R} \) contain on the top the matrices \( \tilde{B} \) and \( \tilde{C} \). In MATLAB notation
\[
\tilde{B} = \hat{R}(1:n,1:m) \quad \text{(m: number of inputs)}
\]
\[
\tilde{C} = \hat{O}(1:q,1:n) \quad \text{(q: number of outputs)}
\]

Because of the shift property of \( \hat{O} \) one gets
\[
\hat{O}(1:end-q+1:1:n)A = \hat{O}(q+1:end,1:n)
\]
\[
\hat{A} = \hat{O}(1:end-q+1:1:n)^*\hat{O}(q+1:end,1:n)
\]

where \( (\cdot)^* \) denotes the Moore-Penrose pseudo inverse. These inverse is to build for a matrix of the low order \( n \).

3. Second step: Order reduction

Using the shortly described method, the demand of a small static error leads to models of relatively high ("middle") order \( n \). In most applications these order can be reduced further on without significant reducing the accuracy of the model. The model matrices are known now. This enables the application of well established order reduction methods like truncated modal or balanced realization, with or without singular perturbation, respectively. The main properties of these methods are summarized in [5].

For a truncation of the model it is to transform in a suitable form. The state space vector has to be divided in a retained \( \nu \)-vector \( x_1 \) and a \((n-\nu)\)-vector \( x_2 \) that can be discarded. \( x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \). The matrices \( A, B, C \) have to be partitioned according these reordering
\[
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}
\]

By omitting the states \( x_2 \) the following lower order system is obtained.
\[ \xi_{k+1} = A_1 \xi_k + B_1 u_k \]
\[ \eta_k = C_1 \xi_k + D u_k \]

These elimination of \( x_2 \) can cause a static error. Therefore in the second way, the singular perturbation approximation will be used, to correct the static behavior. Only the "dynamic" of the omitted states will be neglected. Setting \( x_{2,k+1} = x_{2,k} \) and eliminating \( x_2 \) from the remaining equations yields (if \( E - A_2 \) is nonsingular)
\[ \xi_{k+1} = (A_1 + A_2 (E - A_2)^{-1} A_2) \xi_k + (B_1 + (E - A_2)^{-1} B_2) u_k \]
\[ \eta_k = (C_1 + C_2 (E - A_2)^{-1} A_2) \xi_k + (D + C_2 (E - A_2)^{-1} B_2) u_k \]

The truncation of modal realizations is common in engineering practice. In this case the eigenvalues of \( A \) have to be computed and ordered by aspects of practice (e.g. ordered so that \( |\lambda_i| \), \( \text{Re}(\lambda_i) \) or \( \text{Im}(\lambda_i) \), respectively, is nondecreasing).

For the balancing truncation a state transformation is needed, which arranges the (transformed) states according to their amount of energy transformation from input to output, i.e. according to their transfer significance. So, the states are omitted, which are most nonsignificant.

It is very difficult to predict, which method in a special application gives the better result. This question is to answer by numerical experiments.

Thus, the order of the model which is determined in the first step can be reduced in a second step. Finally one gets a model of the order \( \nu \) with \( \nu < n < N \).

4. Examples

By three examples it will be shown that the two-step method yields very compact models and works reliable.

4.1 An example from MATLAB

The first example "gasf" was taken from the MATLAB Control System Toolbox, described in get_start.pdf, 2.26. The original system contains 4 outputs, 6 inputs, and 25 states. It is characterized by values with great differences (static factors from \( 10^{-5} \) to \( 10^{+5} \)) between amplifications of different input-output branches.

The impulse response instants were obtained using the step und diff functions.

The start of the first step is the construction of the Hankel matrix, the svd and the choice of order \( n=17 \). This setting is motivated by the graph of the singular values of the Hankel matrix, see the vertical black bar in fig. 3.

![Fig.3 Relative singular values of the Hankel matrices, constructed with a different number of block rows](image)

In the following only the step responses with the largest and with the smallest magnitude are considered. From the figures 4 and 5 you can see the good correspondence between the original and the both approximations. Only at time step \( k=0 \) the second approximation of order 12 doesn't match the green circle (the original). For the order reduction in the second step modal truncation (MATLAB function modreal) was used.

By choosing the order \( \nu=12 \) already for the model 1 the correspondence of the step responses in the finish part will be very poor, compare fig. 6.
Fig. 4 Step response from input #1 to output #1, step response with maximal magnitude, the left diagram shows original and both approximations at the start of step response.

Fig. 5 Step response from input #6 to output #4, step response with minimal magnitude, the left diagram shows original and both approximations at the start of step response.

Fig. 6 Step response from input #6 to output #4, original and approximation of order 12 already in the first step.

This example demonstrates:
- the robustness of the first modeling step. It is capable to work on systems with extreme different static branches.
- the advantage of the second step for reduction the model order further on.
4.2 Temperature distribution of an analog Amplifier-IC

Fig. 7 shows the locations of critical points, where heat is produced (black arrows). The temperature distribution is of interest in 102 points, marked in fig. 8 on the surface of the IC. The arrows show the locations of the temperature sensible input stages (in our model the outputs) numbered by 36 and 41.

The step responses between the (normalized) two heat power inputs of the power stages and the 102 marked nodes were calculated by an FEM model, consisting of 27,956 nodes.

Fig. 7 Energy entry points (power transistors)  
Fig. 8 102 output-points on the surface (magenta color)

The step responses between the (normalized) two heat power inputs of the power stages and the 102 marked nodes were calculated by an FEM model, consisting of 27,956 nodes.

Fig. 9 Relative singular values of the Hankel matrices, constructed with a different number of block rows. The block row consists of 102 rows.

<table>
<thead>
<tr>
<th>Output No. #36</th>
<th>Output No. #41</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Input 1]</td>
<td>![Input 1]</td>
</tr>
<tr>
<td>![Input 2]</td>
<td>![Input 2]</td>
</tr>
<tr>
<td>![Output #36]</td>
<td>![Output #41]</td>
</tr>
<tr>
<td>![Output #41]</td>
<td>![Output #36]</td>
</tr>
</tbody>
</table>

Fig. 10 Temperature step responses of outputs #36 and #41 caused by temperature steps of the power stages.
The order of the first model was chosen as $n=21$, compare fig. 9. For the order reduction in the second step balanced truncation (MATLAB functions `balreal` and `modred`) was used. The order was reduced to $\nu=15$. The step responses of the second model agree with the original and the first model apart from the first three / four time instants.

In this example already the first model is of comparative small order. Nevertheless these order can be reduced further on. The correspondence of the model and the original remains apart from the first time points.

### 4.3 A micro system acceleration sensor

Fig. 11 shows the scheme of a micro system acceleration sensor. The FEM model describing the output signal as a function of the acceleration consists of 5365 nodes with a degree of freedom of 3. Therefore the system is of order 16095.

![Fig. 11 Scheme of a micro system acceleration sensor](image1)

The singular values of the Hankel matrix in fig. 12 suggest an order $n=8$. In the second step the order can be reduced to $\nu=2$ without loss of accuracy. The step response outputs of the original and of the two models are shown in fig. 13. Fig. 14 shows zoomed the starting sector of the step response. Both of the models match the points given. The order reduction in the second step was carried out by modal truncation (MATLAB function `modreal`).

![Fig. 12 Relative singular values of the Hankel matrices, constructed with a different number of block rows](image2)

![Fig. 13 Step responses of original and models](image3)

![Fig. 14 Origin sector of the Stepp responses](image4)

A difference between the models determined is visible in the Bode-plot, see fig. 15. As can be seen, they differ on very high frequencies at the end of the plot.
Fig. 15 Bode plot of the two accelerator models of order $n=8$ and order $\nu=2$

5. Conclusions

The combination of well known approaches to

- calculate a partial realization via factorization of the Hankel matrix using svd and
- model order reduction by modal or balanced truncation

yields a powerful and fast tool to generate time discrete state space models from given impulse response values. The two-step approach yields in most applications significantly better results than the modeling of order $\nu$ already in the first step. The method works well on large systems too. The models are accurate and of very low order.

6. References