

MODEL REDUCTION BY KAUTZ FILTERS

A.C. den Brinker

Eindhoven University of Technology
P.O. Box 513, NL-5600 MB Eindhoven, Netherlands
Tel: +31 40 2473628; fax +31 40 2448375
e-mail: A.C.d.Brinker@ele.tue.nl

ABSTRACT

A method is presented for model reduction. It is based on the representation of the original model in an (exact) Kautz series. The Kautz series is an orthonormal model and is non-unique: it depends on the ordering of the poles. The ordering of the poles can be chosen such that the last sections contribute least or the first sections contribute most to the overall impulse response of the original system (in a quadratic sense). Having a specific ordering, the reduced model order, say n , can be chosen by considering the energy contained in a truncated representation. The resulting reduced order model is obtained simply by truncation of the Kautz series at the n th term.

1 INTRODUCTION

In many applications it is convenient to have the disposal of an adequate description of a linear system by an IIR filter of lowest possible order (according to some error criterion). Often, a quadratic norm is used to compare the original system with its lower order approximation. Taking an ARMA model description [1], the aim is to determine the poles and zeros (or numerator and denominator polynomials) of the lower order model. It is well known that using a quadratic norm the error surface usually contains many local minima and that especially the poles turn out to be numerically poorly conditioned. Furthermore, the model order has to be chosen beforehand and thus if the best model within this class is not adequate, the whole estimation procedure has to be repeated.

For these reasons, among others, several suboptimal optimization procedures were developed. One of these is Prony's method [2], but it turns out that this method generally overestimates the damping terms of the poles [3]. Another well-known technique is balanced model order reduction [4], which is also suboptimal in the sense of a quadratic norm. Although this method is appealing from several points of view, it is also associated with numerical problems as a consequence of the required matrix decompositions and inversions.

The numerical problems associated with model reduction can be easily illustrated. Suppose we have a function which is the sum of two exponential sequences $A_1 p_1^k$ and $A_2 p_2^k$. In order that model reduction is possible at all, the poles should be in each others proximity, i.e. a large part of the

sequence p_1^k can be modelled by p_2^k (and vice versa). (We exclude the trivial case that one of the exponential sequences contains hardly any energy at all.) A reduced model involves finding a sequence Bq^k such that the original function is 'best' represented by this simpler representation. However, finding q is a numerically difficult task: small variations in q around its optimal value typically have little effect on the error. Instead of trying to solve this problem, we avoid it by exploiting the property that p_1^k, p_2^k and its optimal lower order model q^k have so much in common. In essence, the technique proposed is the following. Since p_1^k offers already a fair description of p_2^k (and vice versa) we simply select q as that p_i which models most of the energy of both sequences and calculate B accordingly. (Note that the excluded trivial case where one of the exponential sequences contains relatively little energy also fits in nicely with such a procedure and in that case the poles do not need to be located in each others proximity).

Clearly, such technique is suboptimal in a quadratic sense: we do not optimize over the poles, but make a selection on an already given set of poles. The essential part of the procedure is an ordering algorithm, which has been simplified by using the orthogonalization of exponential sequences as given by a Kautz series. Furthermore it is noted that the order selection and model calculation are performed simultaneously and that if the original model is stable, so is the reduced model.

In the next section the Kautz functions (orthogonalized exponential sequences) are introduced. In the two subsequent sections the ordering of the poles and the actual model reduction are discussed. Section 5 contains some preliminary results of the proposed algorithm; large scale tests have as yet not been carried out. The paper concludes with a discussion.

2 KAUTZ FILTERS

Suppose we have a discrete-time model of large order N given by its transfer function

$$H(z) = \sum_{n=1}^N \frac{R_n z}{z - p_n} = \frac{\sum_{n=1}^N B_n z^n}{\prod_{n=1}^N (z - p_n)}, \quad (1)$$

with p_n the poles and R_n the residues ($R_n \neq 0$). We assume that $H(z)$ is a stable filter $|p_n| < 1$. We introduce A as the set

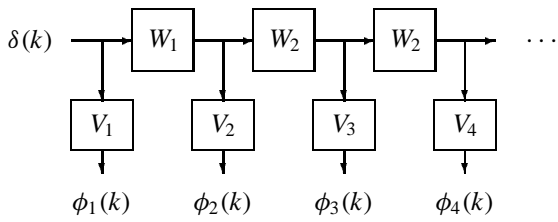


Figure 1: A Kautz filter

of poles, i.e., $A = \{p_n; n = 1, \dots, N\}$, and we assume that $p_n \neq p_m$ for $n \neq m$. Thus we have for the impulse response $h(k) = \sum_{n=1}^N R_n p_n^k$.

The transfer function $H(z)$ is written as a Kautz series

$$H(z) = \sum_{n=1}^N w_n \Phi_n(z) \quad (2)$$

with

$$\Phi_n(z) = \sqrt{1 - q_n q_n^*} \frac{z}{z - q_n} \prod_{l=1}^{n-1} \frac{1 - z q_l^*}{z - q_l} \quad (3)$$

the Kautz functions [5],[6] where $*$ denotes conjugation and $\{q_n; n = 1, \dots, N\} = A$. This Kautz series is of finite order and exact. The Kautz filter is shown in Fig. 1 where $W_n(z) = (1 - z q_n^*) / (z - q_n)$ and $V_n(z) = \sqrt{1 - q_n q_n^*} z / (z - q_n)$.

The Kautz functions are orthonormal

$$\sum_{k=0}^{\infty} \phi_m(k) \phi_n^*(k) = \delta_{mn} \quad (4)$$

with δ_{mn} the Kronecker delta and $\phi_n(k) = \mathcal{Z}^{-1}\{\Phi_n(z)\}$. As a consequence of the orthonormality, the weights w_n can be determined by

$$w_n = \sum_k h(k) \phi_n^*(k) = \frac{1}{2\pi j} \oint H(z) \Phi_n^*(z) \frac{dz}{z}, \quad (5)$$

where $j = \sqrt{-1}$ and the contour integral is taken counterclockwise around the unit circle. The Kautz series is non-unique: there are $N!$ orderings of p_i resulting in allowable q_i 's. In fact, (2) can be regarded as $N!$ possible Gram-Schmidt orthogonalization procedures.

We note that the energy E contained in the impulse response can be calculated in the time-domain, in the frequency domain using Parseval, or from the expansion coefficients in the Kautz series:

$$E = \sum_{k=0}^{\infty} h(k) h^*(k) = \sum_{n=1}^N |w_n|^2. \quad (6)$$

3 ORDERING OF THE POLES

3.1 Backward Ordering

In order to be able to do the model reduction we can choose the ordering of the poles q_n such that the least energy is

contained in the last weights. This is done in a recursive way. First, we select the last pole q_N as that pole p_n which yields minimum weight $|w_N|^2$. Next, we select q_{N-1} as that pole from the remaining set of poles $A \setminus q_N$ such that $|w_{N-1}|^2$ is minimum, etc. In this way we push as much as energy as possible into the first weights of the Kautz series.

In the m th iteration step we select q_{N+1-m} from the $N + 1 - m$ yet unselected poles p_n . The weights that have to be calculated are called w_{mn} where n only takes on those values associated with not yet selected poles p_n . The weights w_{mn} can be calculated by

$$\begin{aligned} w_{mn} &= \frac{1}{2\pi j} \oint H(z) \Phi_{N+1-m}^*(z) \frac{dz}{z} \\ &= \sqrt{1 - p_n p_n^*} \sum_{l=1}^N \frac{R_l}{z - p_l p_n^*} \prod_{i=1}^{N-m} \frac{p_l - q_i}{1 - p_l q_i^*}. \quad (7) \end{aligned}$$

where the q_i 's ($i = 1, \dots, N - m$) have to be interpreted as the unselected poles with the exception of p_n . We choose w_{N+1-m} as that value of w_{mn} with minimum absolute value and q_{N+1-m} as the associated p_n . In order to obtain a unique Kautz series it is required that this minimum is unique. We assume that in practical situations this is always the case.

3.2 Forward Ordering

In the forward ordering scheme, the poles of the Kautz series q_n are selected in ascending order. This is done by selecting in the m th iteration the pole q_m from the remaining $(N - m + 1)$ poles as that one which yields a maximum absolute value for the weight w_m . Analogous to the previous method it is now assumed that in each iteration a unique maximum occurs.

3.3 Remarks

- If the model is exactly reducible (i.e., a model with pole-zero cancellations, or in terms of (1), one or more R_n equal zero) then the backward ordering will give the exact lower order model. The forward ordering method will in general not produce that result.
- The forward ordering method requires less computations than the backward ordering scheme.
- In the forward ordering method one can set a level of accuracy E_n/E beforehand. Then the ordering process can be terminated whenever this accuracy is reached. The backward ordering always requires to complete the entire procedure.
- The algorithm for the ordering of the poles requires a unique minimum or maximum for the possible weights that are calculated in each iteration step. In view of symmetry, this condition will not be met in the case of a real model having one or more complex-conjugated pole pairs. Furthermore, if one starts from a real system, one usually wants a reduced model that is real as well. In order to obtain this, the pole selection procedure has to be adapted in order to guarantee that complex-conjugated pole pairs occur sequentially.

3.4 Reduction of Real Models

The adaptation of the pole selection scheme for a real model always keeps complex-conjugated pole pairs together. This

reduces the number of Gram-Schmidt orthogonalization procedures from $N!$ to $(N_r + N_c)!$ where N_r and N_c are the number of real poles and complex-conjugated pole pairs, respectively, and $N = N_r + 2N_c$.

For the forward ordering process we propose the following algorithm.

1. Add an extra first-order section with a real-valued pole if
 - a. the absolute value of the weight associated with this pole is largest of all (as yet unselected) real poles
 - b. the squared absolute value of this weight is larger than the sum of squared absolute values of the weights of two additional first-order sections for any (as yet unselected) complex-conjugated pole pair.
2. If there is no real pole satisfying the previous requirement, we select two additional first-order sections either with a complex-conjugated pole pair or two real-valued poles. The selection is as before based on a maximum additional energy.

For the backward ordering procedure similar changes can be introduced.

4 MODEL REDUCTION

We now have the Kautz series with the desired ordering of the poles. We can plot E_n defined as

$$E_n = \sum_{m=1}^n |w_m|^2 \quad (8)$$

or E_n/E as a function of n and select the appropriate order as the minimum of n for which $E_n/E > 1 - \epsilon$ where ϵ is the admissible relative deviation in squared norm between the original model and its lower order approximation. The reduced order model is then defined as the n th order model $H_n(z) = \sum_{i=1}^n w_i \Phi_i(z)$ where Φ_i are governed by the ordered poles. We then have deleted the $N - n$ exponential sequences (poles) which, after projection on the remaining set of exponential sequences (poles), contribute least to the overall impulse response according to a quadratic norm.

If the pole ordering outlined in Section 3.4 is used, not every order n of the reduced system yields a real system. In order for the reduced model to be real, truncation in between a complex-conjugated pole pair ($q_{n+1} = q_n^*$) must be prohibited.

5 EXAMPLE

As an example we generated a 30th order random model in MatLab. The generated system is real. Both the forward and backward ordering procedures were used keeping complex-conjugated pole pairs together.

The results in terms of E_n/E are shown in Fig. 2. The results show that a reduction in order by typically a factor 2 or 3 is possible for such a 30th order model with very little loss. A second typical result is that for small orders the forward ordering provides mostly but not always better results (see bottom plot in Fig. 2). Since we are usually interested in a substantial order reduction and in view of the already mentioned advantages of the forward ordering

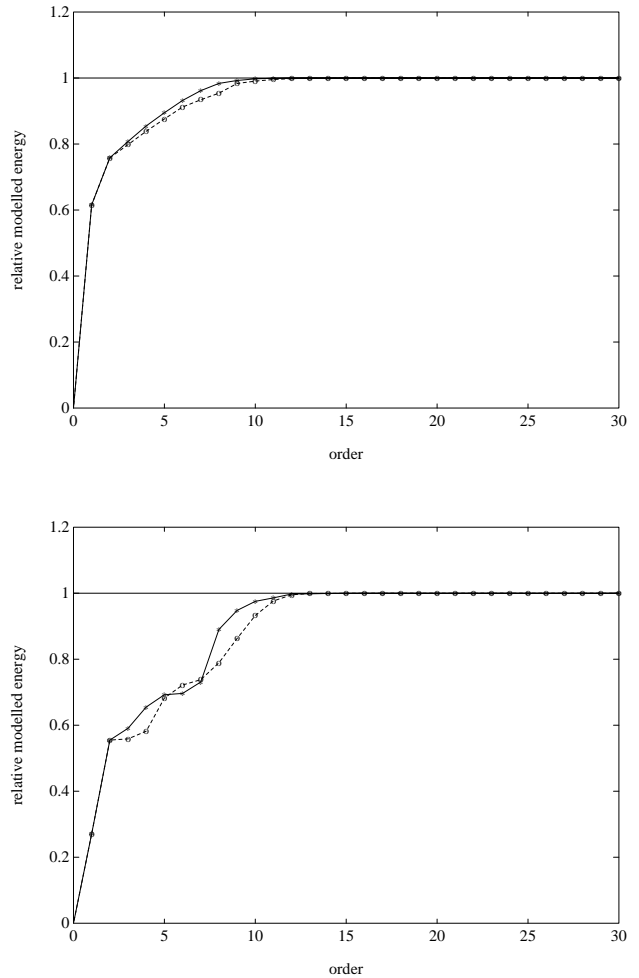


Figure 2: The relative modelled energy E_n/E as a function of the reduced model order. The solid and dashed line indicate the results obtained by forward and backward ordering, respectively.

method, this is considered to be the most promising ordering procedure. A third remark is that the MatLab random model generator typically produces an impulse response where the first values are equal to zero. This can only be accounted for with difficulty in our Kautz model since such initial zero response has to be achieved by destructive interference. Thus, the results will probably improve if the Kautz filter is preceded by a delay line of appropriate length.

6 DISCUSSION

The procedure can in principle be adapted to the case of multiple occurring poles. However, we lose degrees of freedom in the ordering process of the poles. If all the poles are equal, as for instance starting with an FIR model, there is no flexibility at all. Therefore, we restrict ourselves to the case of simple poles. The other argument is of course that in practice

this will be the usual case.

The transfer function $H(z)$ is a restricted N th order rational function. The general case can also be treated in the previous way using the following procedure:

$$H(z) = \frac{\sum_{n=0}^N B_n z^n}{\prod_{n=1}^N (z - p_n)} = C + \frac{\sum_{n=0}^{N-1} B'_n z^n}{\prod_{n=1}^N (z - p_n)}. \quad (9)$$

The constant C represents the direct feedthrough. The second part can be used in the model reduction scheme presented before using one-time delayed versions of the Kautz functions (3).

The presented ordering procedure is numerically well-conditioned; it requires no matrix decomposition, matrix inversions or other numerically cumbersome methods. However, should one start with a transfer function given by

$$H(z) = \frac{\sum_{n=0}^N B_n z^n}{\sum_{n=0}^N A_n z^n} \quad (10)$$

with $A_N = 1$, one must find the roots of the denominator polynomial. This may give numerical problems. Two remarks can be made. First of all the procedure can also be applied by starting with selecting poles q_i from a set A' not identical to the exact poles given by A . Even the dimensions of the sets A and A' need not be equal. The only restriction would be that by the set A' a very good Kautz model of the original system is available. This leads us to the second remark that we can start from the roots of the denominator polynomial even if in itself this procedure might not be very stable. In order to obtain a stable model however, it is required that the roots obtained from the polynomial are within the unit circle. As a side remark we mention that in this case we would not select the poles by calculating w_{mn} according to (7) but rather by

$$w_{mn}^* = \frac{1}{2\pi j} \oint H^*(z) \Phi_{N+1-m}(z) \frac{dz}{z} \quad (11)$$

and splitting Φ in a sum of terms $z/(z - q_i)$.

Our preliminary results suggest that the presented model order reduction technique can be valuable to obtain a first and numerically well-conditioned step in a more profound model order reduction scheme. In the second step, one can use existing methods for model reduction which allow adapting not only the zeros but also the poles of the model.

REFERENCES

- [1] R.N. McDonough and W.H. Huggins, "Best least-squares representation of signals by exponentials," *IEEE Trans. Autom. Control* **AC-13**, 408-412, 1968.
- [2] S.L. Marple, Jr., *Digital Spectral Analysis with Applications*. Englewood Cliffs (NJ): Prentice, 1987.
- [3] M.L. Van Blaricum and R. Mitra, "Problems and solutions associated with Prony's method for processing transient data," *IEEE Trans. Antennas Propag.* **AP-26**, 174-182, 1978.
- [4] B.C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Trans. Autom. Control* **AC-26**, 17-32, 1981.
- [5] P.W. Broome, "Discrete orthonormal sequences," *J. ACM* **12**, 151-168, 1965.
- [6] W.H. Kautz, "Transient synthesis in the time domain," *IRE Trans. Circuit Theory* **CT-1**, 29-39, 1954.