Time-Varying Systems and Computations
Lecture 4

Klaus Diepold

26. November 2012

Realization Theory for Finite Matrices

From Toeplitz Operator to Hankel Operator

We recall the external input-output description of our system

$$\left[\begin{array}{c} \vdots \\ y_{-2} \\ y_{-1} \\ \hline y_0 \\ y_1 \\ \vdots \end{array}\right] = \left[\begin{array}{ccc|ccc} \ddots & \vdots & \vdots & \vdots & \vdots & \\ & T_{-2,-2} & T_{-2,-1} & T_{-2,0} & T_{-2,1} & \cdots \\ \cdots & T_{-1,-2} & T_{-1,-1} & T_{-1,0} & T_{-1,1} & \cdots \\ \hline \cdots & T_{0,-2} & T_{0,-1} & \boxed{T_{0,0}} & T_{0,1} & \cdots \\ \cdots & T_{1,-2} & T_{1,-1} & T_{1,0} & T_{1,1} & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] \left[\begin{array}{c} \vdots \\ u_{-2} \\ u_{-1} \\ \hline u_0 \\ u_1 \\ \vdots \end{array}\right],$$

which is described by use of the *Toeplitz* Operator T. We can subdivide the Toeplitz operator in four parts, each desribing the mapping of past inputs to the past outputs ($u_p \mapsto y_p$), of past inputs to future outputs ($u_p \mapsto y_f$), of future inputs to past outputs ($u_f \mapsto y_f$) and of future inputs to future outputs ($u_f \mapsto y_f$), as depicted in Figure 1.

The lower left part of T, notably $T_{p,f}$ represents the map from past input to future output. It is this map $T_{p,f}$ that contains all the necessary information about the internal workings of our system, in particular it contains the information about the states in the system. A state can be considered to comprise all information originating from the past that still has influence on the future. In the literature this notion is often called the *Markov* property.

Cutting out this part of the Toeplitz operator produces the equation

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \cdots & T_{0,-3} & T_{0,-2} & T_{0,-1} \\ \cdots & T_{1,-3} & T_{1,-2} & T_{1,-1} \\ \cdots & T_{2,-3} & T_{2,-2} & T_{2,-1} \\ & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ u_{-3} \\ u_{-2} \\ u_{-1} \end{bmatrix}.$$

For notational convenience (having all dots going downwards or to the right) we flip the entries of the input vector u upside down, which causes a corresponding permutation in the matrix, id est, the columns

$$\begin{array}{c} \text{past} \uparrow \\ \text{future} \downarrow \end{array} \left[\begin{array}{c} y_p \\ \hline y_f \end{array}\right] = \left[\begin{array}{c|c} T_{p,p} & T_{f,p} \\ \hline \underbrace{T_{p,f}}_{\mathcal{H}} & T_{f,f} \end{array}\right] \left[\begin{array}{c} u_p \\ \hline u_f \end{array}\right] \begin{array}{c} \uparrow \text{past} \\ \downarrow \text{future} \end{array}$$

Figure 1: Hankel operator is a part of the Toeplitz operator.

are flipped left-right. The result is the equation

$$\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} T_{0,-1} & T_{0,-2} & T_{0,-3} & \cdots \\ T_{1,-1} & T_{1,-2} & T_{1,-3} & \cdots \\ T_{2,-1} & T_{2,-2} & T_{2,-3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u_{-1} \\ u_{-2} \\ u_{-3} \\ \vdots \end{bmatrix}.$$

The part $T_{p,f}$ of the *Toeplitz* operator with the flipped rows is so important in systems theory that it has its own name, it is called the *Hankel* operator, which we will denote with the letter $\mathcal{H}$

$$\mathcal{H}_0 = \begin{bmatrix} T_{0,-1} & T_{0,-2} & T_{0,-3} & \cdots \\ T_{1,-1} & T_{1,-2} & T_{1,-3} & \cdots \\ T_{2,-1} & T_{2,-2} & T_{2,-3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where the subscript $\mathcal{H}_0$ means that this Hankel matrix is valid if we consider the time point $k = 0$ as the present.

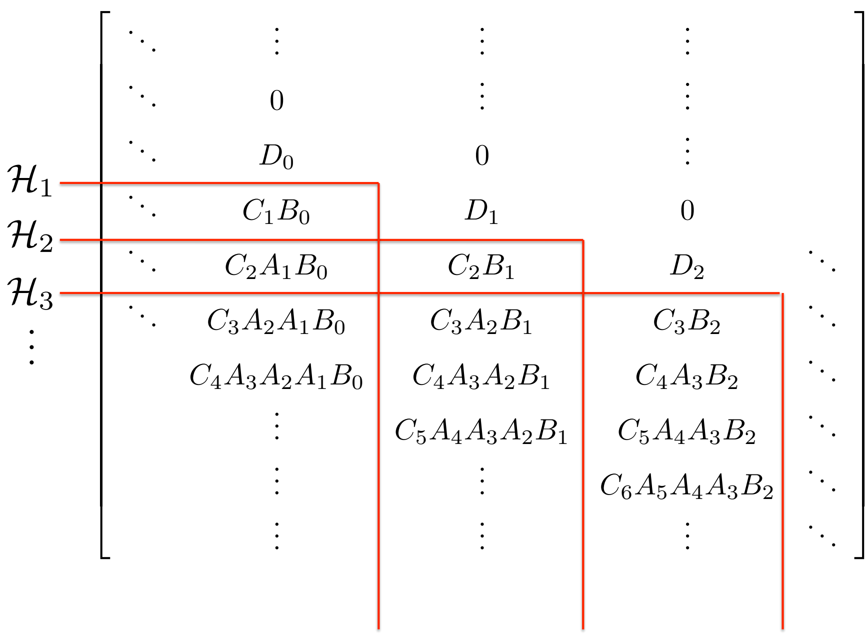

Figure 2: Schematic representation of time-varying Hankel operators for a causal system.

Looking at Figure 2 we can identify the first three Hankel operators as

$$\mathcal{H}_1 = \begin{bmatrix} C_1B_0 & C_1A_0B_{-1} & C_1A_0A_{-1}B_{-2} & \cdots \\ C_2A_1B_0 & C_2A_1A_0B_{-1} & C_2A_1A_0A_{-1}B_{-2} & \cdots \\ C_3A_2A_1B_0 & C_3A_2A_1A_0B_{-1} & C_3A_2A_1A_0A_{-1}B_{-2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

$$\mathcal{H}_2 = \begin{bmatrix} C_2B_1 & C_2A_1B_0 & C_2A_1A_0B_{-1} & \cdots \\ C_3A_2B_1 & C_3A_2A_1B_0 & C_3A_2A_1A_0B_{-1} & \cdots \\ C_4A_3A_2B_1 & C_4A_3A_2A_1B_0 & C_4A_3A_2A_1A_0B_{-1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

and

$$\mathcal{H}_3 = \begin{bmatrix} C_3B_2 & C_3A_2B_1 & C_3A_2A_1B_0 & \cdots \\ C_4A_3B_2 & C_4A_3A_2B_1 & C_4A_3A_2A_1B_0 & \cdots \\ C_5A_4A_3B_2 & C_5A_4A_3A_2B_1 & C_5A_4A_3A_2A_1B_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

The generic case of a Hankel Operator for arbitrary k looks like

$$\mathcal{H}_k = \begin{bmatrix} T_{k,k-1} & T_{k,k-2} & T_{k,k-3} & \cdots \\ T_{k+1,k-1} & T_{k+1,k-2} & T_{k+1,k-3} & \cdots \\ T_{k+2,k-1} & T_{k+2,k-2} & T_{k+2,k-3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{1}$$

where the individual matrix entries $T_{i,j}$ are given as

$$T_{ij} = \begin{cases} D_i & \text{for} \quad i = j, \\ C_iA_{i-1}\dots A_{j+1}B_j & \text{for} \quad i < j, \\ 0 & \text{for} \quad i > j, \end{cases}$$

Factorization of the Hankel Operator

Figure 3 shows the concept for decomposing the Hankel operator into a sequence of two maps. The first map, denoted by $\mathcal{C}$ represents the map of the past input signals u_p into the internal states of the system, i.e. the internal signals x, which are stored in the registers Z. The second map is denoted by the symbol $\mathcal{O}$ describes how the internal states x are mapped into the future output signals y_f. We see that the states x sit between the input signals that lie in the past and the output signals that lie in the future, or in other words, the states store the information from the past that is relevant for the future. The information about the state-structure of the linear system is embedded in the Hankel operator.

According to these considerations we factor the Hankel operator in Equation 1 as the product

$$\mathcal{H}_k = \begin{bmatrix} C_kB_{k-1} & C_kA_{k-1}B_{k-2} & C_kA_{k-1}A_{k-2}B_{k-3} & \cdots \\ C_{k+1}A_kB_{k-1} & C_{k+1}A_kA_{k-1}B_{k-2} & C_{k+1}A_kA_{k-1}A_{k-2}B_{k-3} & \cdots \\ C_{k+2}A_{k+1}A_kB_{k-1} & C_{k+2}A_{k+1}A_kA_{k-1}B_{k-2} & C_{k+2}A_{k+1}A_kA_{k-1}A_{k-2}B_{k-2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \mathcal{O}_k \cdot \mathcal{C}_k$$

where we have

$$\mathcal{O}_k = \begin{bmatrix} C_k \\ C_{k+1}A_k \\ C_{k+2}A_{k+1}A_k \\ \vdots \end{bmatrix}, \quad \mathcal{C}_k = \begin{bmatrix} B_{k-1} & A_{k-1}B_{k-2} & A_{k-1}A_{k-2}B_{k-3} & \cdots \end{bmatrix}. \tag{2}$$

We call $\mathcal{O}_k$ the *observability* matrix, because it describes how the states are mapped into the output signal, which allows us to infer the sequence of states by just observing the output signal y. Observe that the observability matrix depends only on the parameters A_k and C_k.

We call the matrix $\mathcal{C}_k$ the *controllability* matrix, which describes the mapping of the input signals into the states, which allows us to control the system's internal states using the input signal u. Observe that the controllability matrix depends only on the parameters A_k and B_k.

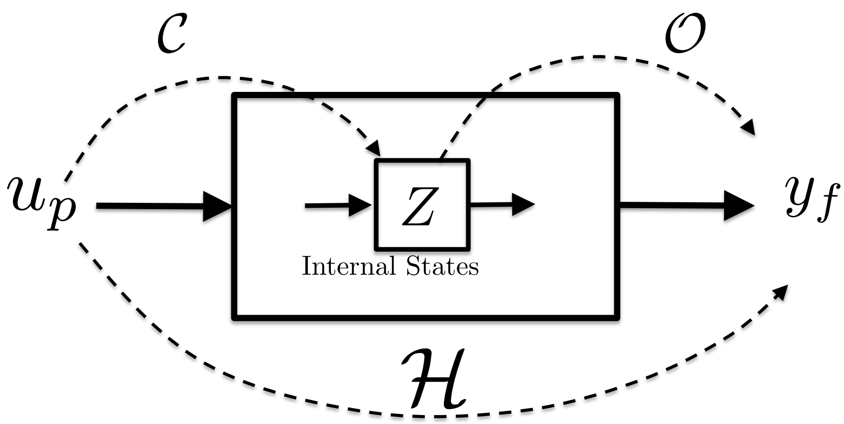

Figure 3: Partioning of the Hankel operator into a sequence of two maps.

Gramians

Observability Gramian

We start out with recalling the definition of the observability matrix

$$\mathcal{O}_k = \begin{bmatrix} C_k \\ C_{k+1}A_k \\ C_{k+2}A_{k+1}A_k \\ \vdots \end{bmatrix}.$$

In order to evaluate the characteristics of vectors which span row space of $\mathcal{O}_k$ we calculate the Gramian matrix according to

$$\mathcal{O}_k^T\mathcal{O}_k = \begin{bmatrix} C_k^T & A_k^TC_{k+1}^T & A_k^TA_{k+1}^TC_{k+2}^T & \cdots \end{bmatrix} \begin{bmatrix} C_k \\ C_{k+1}A_k \\ C_{k+2}A_{k+1}A_k \\ \vdots \end{bmatrix} =$$

$$= C_k^TC_k + A_k^TC_{k+1}^TC_{k+1}A_k + A_k^TA_{k+1}^TC_{k+2}^TC_{k+2}A_{k+1}A_k + \cdots,$$

which we will call for obvious reasons the *Observability Gramian.* Using the shorthand $K_k = \mathcal{O}_k^T\mathcal{O}_k \geq 0$ for the Observability Gramian we can re-write the lengthy expression recursively as

$$K_k = C_k^TC_k + A_k^TK_{k+1}A_k.$$

This equation is commonly referred to as a Lyapunov equation. The dynamic degree of a minimal state-space realization corresponds with the rank of the observability Gramian, i.e. with the number or linear independent rows of the observability matrix. We can compute this rank and the associated basis for the space by applying an SVD to the Hankel matrix.

We know from control engineering that a system is called completely observable, if the observability matrices $\mathcal{O}_k$ has full row rank, or, equivalently that all observability Gramians $\mathcal{K}_k$ is positive definite.

Controllability Gramian

We start out with recalling the definition of controllability matrix

$$\mathcal{C}_k = \begin{bmatrix} B_{k-1} & A_{k-1}B_{k-2} & A_{k-1}A_{k-2}B_{k-3} & \cdots \end{bmatrix}.$$

In order to evaluate the characteristics of the vectors, which span the column space $\mathcal{C}_k$ we calculate the Gramian matrix

$$\mathcal{C}_k\mathcal{C}_k^T = \begin{bmatrix} B_{k-1} & A_{k-1}B_{k-2} & A_{k-1}A_{k-2}B_{k-3} & \cdots \end{bmatrix} \begin{bmatrix} B_{k-1}^T \\ B_{k-2}^TA_{k-1}^T \\ B_{k-3}^TA_{k-2}^TA_{k-1}^T \\ \cdots \end{bmatrix}$$

$$= B_{k-1}B_{k-1}^T + A_{k-1}B_{k-2}B_{k-2}^TA_{k-1}^T + A_{k-1}A_{k-2}B_{k-3}B_{k-3}^TA_{k-2}^TA_{k-1}^T + \cdots,$$

which we will call for obvious reasons the *Controllabilty Gramian.* Using the shorthand $W_k = \mathcal{C}_k\mathcal{C}_k^T \geq 0$ for the Controllability Gramian we can re-write the lengthy expression recursively as

$$W_k = B_{k-1}B_{k-1}^T + A_{k-1}W_{k-1}A_{k-1}^T.$$

This equation is commonly referred to as a Lyapunov equation. The dynamic degree of a minimal state-space realization corresponds with the rank of the controllability Gramian, i.e. with the number or linear independent columns of the controllability matrix. We can compute this rank and the associated basis for the space by applying an SVD to the Hankel matrix.

We know from control engineering that a system is called completely controllable, if the Controllability matrices $\mathcal{C}_k$ has full row rank, or, equivalently that all controllability Gramians $\mathcal{W}_k$ is positive definite.

A system is called stable if it is completely observable and completely controllable.

Special Case: Time Invariant System

Remember that for the case of time-invariant systems the infinite dimensional *Toeplitz* Operator $\boldsymbol{T}$ has *Toeplitz* structure

$$\left[\begin{array}{c} \vdots \\ y_{-2} \\ y_{-1} \\ \hline y_0 \\ y_1 \\ \vdots \end{array}\right] = \left[\begin{array}{ccc|ccc} \ddots & \vdots & \vdots & \vdots & \vdots & \\ & T_0 & T_{-1} & T_{-2} & T_{-3} & \dots \\ \dots & T_1 & T_0 & T_{-1} & T_{-2} & \dots \\ \hline \dots & T_2 & T_1 & T_0 & T_{-1} & \dots \\ \dots & T_3 & T_2 & T_1 & T_0 & \dots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] \left[\begin{array}{c} \vdots \\ u_{-2} \\ u_{-1} \\ \hline u_0 \\ u_1 \\ \vdots \end{array}\right],$$

where all matrix entries along the diagonals are identical. We can express the entries of the Toeplitz operator by using the representation of the impulse response in terms of the state-space realization $\{A, B, C, D\}$

$$T = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \\ \ddots & 0 & \vdots & \vdots & \\ \ddots & T_0 & 0 & \vdots & \\ \ddots & T_1 & T_0 & 0 & \\ \ddots & T_2 & T_1 & T_0 & \ddots \\ \ddots & T_3 & T_2 & T_1 & \ddots \\ & T_4 & T_3 & T_2 & \ddots \\ & \vdots & T_4 & T_3 & \ddots \\ & \vdots & \vdots & T_4 & \ddots \\ & \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \\ \ddots & 0 & \vdots & \vdots & \\ \ddots & D & 0 & \vdots & \\ \ddots & CB & D & 0 & \\ \ddots & CAB & CB & D & \ddots \\ \ddots & CA^2B & CAB & CB & \ddots \\ & CA^3B & CA^2B & CAB & \ddots \\ & \vdots & CA^3B & CA^2B & \ddots \\ & \vdots & \vdots & CA^3B & \ddots \\ & \vdots & \vdots & \vdots & \end{bmatrix}.$$

One can see in Figure 4 that in case of a LTI system all the infinite dimensional *Hankel* Operators are identical, and that this $\mathcal{H}$ has matrix entries, which are identical along the anti-diagonals

$$\mathcal{H} = \begin{bmatrix} T_1 & T_2 & T_3 & \dots \\ T_2 & T_3 & T_4 & \dots \\ T_3 & T_4 & T_5 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} CB & CAB & CA^2B & \dots \\ CAB & CA^2B & CA^3B & \dots \\ CA^2B & CA^3B & CA^4B & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

This is a matrix structure which is commonly referred to as *Hankel* structure. For LTI systems we have only one Hankel Operator, because all $\mathcal{H}_k$ are identical due to the time-invariant property.

$$\begin{matrix} \\ \mathcal{H} \\ \mathcal{H} \\ \mathcal{H} \\ \vdots \\ \\ \\ \\ \end{matrix} \begin{bmatrix} \ddots & D & 0 & \vdots & \\ \ddots & CB & D & 0 & \\ \ddots & CAB & CB & D & \ddots \\ \ddots & CA^2B & CAB & CB & \ddots \\ & CA^3B & CA^2B & CAB & \ddots \\ & \vdots & CA^3B & CA^2B & \ddots \\ & \vdots & \vdots & CA^3B & \ddots \\ & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Figure 4: Schematic representation of time-invariant Hankel matrices for a causal system.

We can factor the Hankel operator into the product $\mathcal{H} = \mathcal{O}\mathcal{C}$, i.e. into the product of observability and controllability, similar to our discussion in the previous section. This factorization looks like

$$\mathcal{H} = \begin{bmatrix} CB & CAB & CA^2B & \dots \\ CAB & CA^2B & CA^3B & \dots \\ CA^2B & CA^3B & CA^4B & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \cdot \begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix} = \mathcal{O}\mathcal{C},$$

such that we have

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix}.$$

The observability Gramian for the realization of a LTI system turns out as

$$W = \mathcal{C}\mathcal{C}^T = \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} \begin{bmatrix} B^T \\ B^TA^T \\ B^T(A^2)^T \\ \vdots \end{bmatrix} = BB^T + ABB^TA^T + A^2BB^T(A^2)T + \cdots$$

which is a solution of the Lyapunov equation

$$W = BB^T + AWA^T.$$

The corresponding formulation of the observability Gramian turns out to be

$$K = \mathcal{O}^T\mathcal{O} = \begin{bmatrix} C^T & A^TC^T & (A^2)^TC^T & \cdots \end{bmatrix} \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} = C^TC + A^TC^TCA + (A^2)^TC^TCA^2 + \cdots,$$

which is a solution to the Lyapunov equation

$$K = C^TC + A^TKA.$$

Finite Lower Triangular Matrices

We consider a lower triangular matrix T corresponding to a causal system. We can interpret the columns of the matrix $\boldsymbol{T}$ as the impulse responses of a time-varying system. We hence write

$$\boldsymbol{T} = \begin{bmatrix} T_{11} & 0 & 0 & 0 & 0 \\ T_{21} & T_{22} & 0 & 0 & 0 \\ T_{31} & T_{32} & T_{33} & 0 & 0 \\ T_{41} & T_{42} & T_{43} & T_{44} & 0 \\ T_{51} & T_{52} & T_{53} & T_{54} & T_{55} \end{bmatrix} = \begin{bmatrix} D_1 & 0 & 0 & 0 & 0 \\ C_2B_1 & D_2 & 0 & 0 & 0 \\ C_3A_2B_1 & C_3B_2 & D_3 & 0 & 0 \\ C_4A_3A_2B_1 & C_4A_3B_2 & C_4B_3 & D_4 & 0 \\ C_5A_4A_3A_2B_1 & C_5A_4A_3B_2 & C_5A_4B_3 & C_5B_4 & D_5 \end{bmatrix}.$$

We can read off the corresponding Hankel operators as well as the observability and controllability matrices

$$\mathcal{H}_1 = [\cdot]$$

$$\mathcal{H}_2 = \begin{bmatrix} C_2B_1 \\ C_3A_2B_1 \\ C_4A_3A_2B_1 \\ C_5A_4A_3A_2B_1 \end{bmatrix} = \begin{bmatrix} C_2 \\ C_3A_2 \\ C_4A_3A_2 \\ C_5A_4A_3A_2 \end{bmatrix} \cdot \begin{bmatrix} B_1 \end{bmatrix} = \mathcal{O}_2\mathcal{C}_2$$

$$\mathcal{H}_3 = \begin{bmatrix} C_3B_2 & C_3A_2B_1 \\ C_4A_3B_2 & C_4A_3A_2B_1 \\ C_5A_4A_3B_2 & C_5A_4A_3A_2B_1 \end{bmatrix} = \begin{bmatrix} C_3 \\ C_4A_3 \\ C_5A_4A_3 \end{bmatrix} \cdot \begin{bmatrix} B_2 & A_2B_1 \end{bmatrix} = \mathcal{O}_3\mathcal{C}_3$$

$$\mathcal{H}_4 = \begin{bmatrix} C_4B_3 & C_4A_3B_2 & C_4A_3A_2B_1 \\ C_5A_4B_3 & C_5A_4A_3B_2 & C_5A_4A_3A_2B_1 \end{bmatrix} = \begin{bmatrix} C_4 \\ C_5A_4 \end{bmatrix} \cdot \begin{bmatrix} B_3 & A_3B_2 & A_3A_2B_1 \end{bmatrix} = \mathcal{O}_4\mathcal{C}_4$$

$$\mathcal{H}_5 = \begin{bmatrix} C_5B_4 & C_5A_4B_3 & C_5A_4A_3B_2 & C_5A_4A_3A_2B_1 \end{bmatrix} = \begin{bmatrix} C_5 \end{bmatrix} \cdot \begin{bmatrix} B_4 & A_4B_3 & A_4A_3B_2 & A_4A_3A_2B_1 \end{bmatrix} = \mathcal{O}_5\mathcal{C}_5$$

Literatur

[1] G. Strang. *Computational Science and Engineering*. Wellesley-Cambridge Press, 2007.

[2] P. Regalia. *Adaptive IIR Filtering in Signal Processing and Control*. CRC Press, 1995.

[3] T. Kailath. *Linear Systems*. Prentice Hall, 1980.

[4] P. Dewilde, A.-J. van derVeen. *Time-Varying Systems and Computations*. Kluwer Academic Publishers, 1998.

[5] A.C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Advances in Design and Control. SIAM, Philadelphia, 2005.