

State-Space Parametrizations of Multivariable Linear Systems using Tridiagonal Matrix Forms¹

Tomas McKelvey and Anders Helmersson

Dept. of Electrical Engineering, Linköping University

S-581 83 Linköping, Sweden,

Email: tomas@isy.liu.se, andersh@isy.liu.se.

CDC, Kobe Japan, 1996

Abstract

Tridiagonal parametrizations of linear state-space models are proposed for multivariable system identification. The parametrizations are surjective, i.e. all systems up to a given order can be described. The parametrization is based on the fact that any real square matrix is similar to a real tridiagonal form as well as a compact tridiagonal form. These parametrizations has significantly fewer parameters compared to a full parametrization of the state-space matrices.

1 Introduction

System identification deals with selecting a linear model from a model set using measured data. A set of models is most often described by a parametrized model structure. Hence, a parametrization is a mapping from the parameter space, a subset of \mathbb{R}^d , to the space of linear systems of given input-output configuration and state order. The best model in the set is determined by a parametric optimization of some criterion function. We will here consider parametrizations of multivariable linear systems based on the state-space form

$$\dot{x} = Ax + Bu, \quad y = Cx + Du$$

with n states, m inputs and p outputs. A parametrization describes which matrix elements have fixed values and where the parameters enter into the system matrix elements.

Parametrization of multivariable linear systems is known as a difficult problem if a minimal parametrization is sought. In a minimal parametrization the mapping from parameter space to system space is injective, i.e. each system can be represented by at most one point in the parameter space. It is known [1] that the set of all possible input-output behaviors is a manifold of dimension $n(m+p) + pm$. The number of elements in the system matrices is however $n^2 + n(m+p) + pm$. Conse-

quently, in order to obtain an injective parametrization, n^2 degrees of freedom must be removed.

1.1 Minimal Parametrizations

In system identification injective parametrizations are known as being identifiable. If the parametrization is also bijective the parametrization can describe all systems of given order and input-output configuration. In this case each system corresponds to one and only one point in the parameter space.

If the parametrization is bijective it is sufficient, for each model order, to only consider one parametrization and consequently only have to run one optimization. For single input multiple output systems or multi input single output systems the standard controllable canonical form and the observable canonical form are examples of bijective parametrizations [5].

A classical result tells us that if both the input and output dimensions are larger than one, a bijective parametrization does *not* exist [7, 3]. Partially overlapping, injective parametrizations, based on the companion matrix form, has been used for multivariable system identification [12]. In this type of parametrizations the number of parameters, which is the minimal number, is

$$d_{\min} = n(m+p) + mp$$

where n is the state dimension, m the input dimension and p the output dimension. For the observable form there exist $\binom{n-1}{p-1}$ different structures and for the controllable form $\binom{n-1}{m-1}$ structures. The model sets spanned by these parametrizations are partially overlapping and the union of all these equals the set of all dynamical systems of fixed input-output dimensions and fixed state order. The obvious drawback with such a parametrization is that several, or all, parametrizations should be considered when identifying a model from data without any prior structural knowledge. Even though one structure is capable of describing almost the entire set of systems it is likely that such a parametrization is numerically less favorable for a much larger class near the border of the set.

Based on the balanced realization a minimal (injec-

¹This work was supported in part by the Swedish Research Council for Engineering Sciences (TFR), which is gratefully acknowledged.

tive) parametrization has also been developed [10] and studied in an identification context in [8]. They argue that the balanced realization in general offer a more well conditioned parametrization as compared with the controllable-observable forms.

1.2 Non-minimal Parametrizations

If we drop the requirement of injectiveness of the parametrization and instead consider surjective parametrizations, the full parametrization is one trivial choice. In this parametrization all matrix elements of the state space matrices (A, B, C, D) are parameters. In an identification context such a parametrization was proposed in [9]. An obvious drawback with the full parametrization is the large number of parameters which have to be estimated:

$$d_{\text{full}} = n^2 + n(m + p) + mp.$$

In comparison with the minimal parametrization the full parametrization has n^2 extra, redundant parameters. When using a parametrization with redundant parameters the minimum of the criterion function will not be unique in the parameter space. This leads to a singular Hessian matrix. However, trivial modifications to standard Newton-type optimization algorithms ensure that a minimum is reached in a numerically stable way.

Without dropping the surjective property it is desirable to find a parametrization using fewer parameters. In this paper we show that every real square matrix is similar to a tridiagonal real matrix. Any matrix can be reduced to this quite compact form using similarity transformations without any assumptions about the eigenvalues or the eigensystem structure of the matrix. We propose a surjective parametrization of state-space models based on this tridiagonal form.

In Section 2, we prove that any real square matrix is similar to a real tridiagonal matrix form as well as a real tridiagonal compact form. Based on these tridiagonal forms two new parametrizations of multivariable state-space models are introduced in section 3. In section 4, a small example illustrates the numerical advantages with the proposed structure.

2 Tridiagonal Matrix Forms

A square matrix A is similar to another matrix \tilde{A} if there exists a similarity transformation defined by the non-singular transformation matrix T such that $T^{-1}AT = \tilde{A}$.

Definition 1 A real tridiagonal matrix is a square real matrix having non-zero elements only on the main, first super and first sub diagonals.

Lemma 1 (The Real Tridiagonal Lemma) Any square real matrix is similar to a real tridiagonal matrix form.

Proof. This lemma is straightforward to show assuming distinct eigenvalues. As a preparation for the proof of general case, we first recap some results on the Jordan form.

It is a well-known fact, see e.g. [4], that any square matrix is similar to a block-diagonal matrix

$$\tilde{A} = \text{diag} [[J_{\nu_1}(\lambda_1), J_{\nu_2}(\lambda_2), \dots, J_{\nu_n}(\lambda_n)], \quad (1)$$

containing Jordan blocks. Each Jordan block,

$$J_{\nu}(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda & \lambda \end{bmatrix} \in \mathbb{C}^{\nu \times \nu}, \quad (2)$$

is associated with the eigenvalue λ . The eigenvalues can be either real or complex. For real eigenvalues the Jordan block is already on a real tridiagonal form. It remains to prove Lemma 1 for complex eigenvalues.

Jordan blocks associated with a complex eigenvalue and its complex conjugate can be rewritten into a real Jordan block $C_{\nu}(a, b) =$

$$\begin{bmatrix} a & b & 1 & 0 & \cdots & \cdots & \cdots & 0 & 0 \\ -b & a & 0 & 1 & \cdots & \cdots & \cdots & 0 & 0 \\ 0 & 0 & a & b & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -b & a & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 & a & b & 1 & 0 \\ \vdots & \vdots & & 0 & 0 & -b & a & 0 & 1 \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & a & b \\ 0 & 0 & \cdots & \cdots & \cdots & 0 & 0 & -b & a \end{bmatrix} \in \mathbb{R}^{2\nu \times 2\nu}$$

where $\lambda = a + jb$ satisfies the equation $\lambda^2 - 2a\lambda + a^2 + b^2 = 0$. The real Jordan block is not tridiagonal since it contains two super-diagonals. There exists a real tridiagonal form that is similar to the real Jordan form for complex eigenvalues. This is now shown by using minimal polynomials [4].

Consider a tridiagonal matrix of the form $A_{\nu}(a, b) =$

$$\begin{bmatrix} a & b & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ -b & a & 1 & 0 & & & & & \vdots \\ 0 & 0 & a & b & 0 & & & & \vdots \\ \vdots & 0 & -b & a & 1 & 0 & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & 0 & 0 & a & b & 0 & \vdots \\ \vdots & & & & 0 & -b & a & 1 & 0 \\ \vdots & & & & & 0 & 0 & a & b \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & -b & a \end{bmatrix} \in \mathbb{R}^{2\nu \times 2\nu}. \quad (3)$$

The compact tridiagonal form has

$$d_{\text{c-tri}} = 2n - 1 + n(m + p) + pm$$

number of parameters. The price paid for surjectivity is the estimation of $2n - 1$ additional parameters, which is considerably less than n^2 for the full parametrization. Although more parameters have to be adjusted it does not necessary mean that the actual computational load during an identification has to increase. As we will see in the example in the next section the number of necessary evaluations of the identification criterion in order to reach a minimum actually decreases for the non-minimal forms.

3.2 Finding the Tridiagonal Form

Prior to a parametric identification an initial estimate is derived, for example by some subspace method. In order to utilize the tridiagonal parametrization a similarity transformation has to be found that tridiagonalize the A -matrix. If all eigenvalues are distinct a standard eigen-decomposition of A will result in a similarity transformation that diagonalizes the system. From the complex diagonal form it is trivial to find the transformation to the real compact tridiagonal form. If multiple eigenvalues are at hand and the eigenvectors do not form a complete basis, the diagonal form does not exist. Then a non-symmetric Lanczos tridiagonalization procedure can be used, see [11, 13].

4 Tridiagonal Parametrization Improves Optimization Performance

The eigenvalues of a matrix in the companion form is known to be very sensitive to perturbations in the non-fixed elements except for a certain eigenvalue distribution [13]. The sensitivity of the eigenvalues of parameter perturbations of matrices in the tridiagonal form is in general much smaller [13]. Consequently one could argue that the tridiagonal form is more favorable from a numerical point of view.

By a simple example we will substantiate this claim by showing that the probability of convergence to the global optimum is larger and that less criterion function evaluations are necessary. We do this by estimating systems of increasing complexity using three different parametrizations:

- tridiagonal form
- compact tridiagonal form
- observable canonical form

The systems have the transfer function

$$G_0(s) = \sum_{k=1}^M \frac{1}{s^2 + 2\zeta_k \omega_k s + \omega_k^2}$$

Table 1: Estimation results of Monte Carlo simulations using different model orders and fixed perturbation level of the initial parameters.

M		Parametrization		
		Tridiagonal	Compact	Observable
4	max V	3.1×10^{-9}	3.7×10^{-9}	9.6×10^{-9}
	min V	1.5×10^{-11}	4.7×10^{-11}	5.9×10^{-9}
	# fail	0	0	0
	eval.	306	266	609
5	max V	7.3×10^{-9}	6.6×10^{-2}	5.9×10^{-4}
	min V	9.7×10^{-11}	2.5×10^{-10}	2.4×10^{-5}
	# fail	0	2	10
	eval.	448	400	-
6	max V	8.0×10^{-9}	8.1×10^{-9}	4.8×10^{-1}
	min V	1.1×10^{-10}	1.3×10^{-10}	1.4×10^{-2}
	# fail	0	0	10
	eval.	647	606	-
7	max V	9.4×10^{-9}	7.9×10^{-9}	6.2×10^{-1}
	min V	1.7×10^{-9}	8.7×10^{-10}	2.3×10^{-2}
	# fail	0	0	10
	eval.	1070	979	-

with natural frequency $\omega_k = 1 + 4(k - 1)/(M - 1)$ and damping ratio $\zeta_k = 0.1$. Denote by $g_0(t)$ the impulse response of G_0 . The data used for the identification are 200 samples of the noise-free impulse response $g_0(t)$ of the system sampled with a period time $T=0.1$.

The parametric optimization minimizes the nonlinear-least squares criterion function

$$V(\theta) = \sum_{k=1}^N |g_0(kT) - \hat{g}(k, \theta)|^2 \quad (7)$$

using the Levenberg-Marquardt method `leastsq` in the Matlab Optimization Toolbox [2]. In (7), $\hat{g}(k, \theta)$ is the impulse response of a parametrized discrete time system. Since the data is noise-free the criterion $V(\theta)$ should be zero at a minimum. The minimization is terminated if the criterion function decreases below 10^{-8} or if more than 2000 criterion evaluations has occurred. The optimization is considered successful if the algorithm terminates with $V(\theta) < 10^{-8}$.

The optimization algorithm is initiated with a parameter value derived from a perturbed system. The perturbation is constructed by forming the sample and hold equivalence of $G_0(s)$ in balanced form and perturbing all matrix elements with zero mean Gaussian random noise with variance σ^2 . The perturbed balanced form is converted to the form associated with the particular parametrization which defines the initial point in the parameter space.

Tables 1 and 2 report the results of Monte Carlo simulations. For each experimental setup, 10 optimizations are performed using different initial parameters. In Table 1 the results for different system orders ($M = 4, \dots, 7$) and constant perturbation level $\sigma^2 = 0.01$ are reported. The results when keeping the

Table 2: Estimation results of Monte Carlo simulations with different perturbation levels of the initial parameters.

σ^2		Parametrization		
		Tridiagonal	Compact	Observable
0.01	max V	5.2×10^{-9}	6.4×10^{-9}	9.6×10^{-9}
	min V	7.6×10^{-11}	1.5×10^{-10}	4.9×10^{-9}
	# fail	0	0	0
	eval.	297	233	568
0.02	max V	6.3×10^{-9}	8.3×10^{-9}	8.7×10^{-4}
	min V	2.3×10^{-11}	3.9×10^{-11}	4.1×10^{-9}
	# fail	0	0	2
	eval.	454	368	790
0.03	max V	8.8×10^{-2}	5.2×10^{-9}	1.5×10^{-2}
	min V	4.0×10^{-11}	5.9×10^{-11}	4.1×10^{-9}
	# fail	2	0	2
	eval.	557	647	1192
0.04	max V	1.6×10^{-1}	1.4×10^3	6.1×10^{-1}
	min V	2.1×10^{-10}	4.0×10^{-11}	3.9×10^{-9}
	# fail	4	3	6
	eval.	758	557	1035

system order constant $M = 4$ and increasing the variance $\sigma^2 = 0.01, \dots, 0.04$ of the initial parameter perturbations are reported in Table 2. For each experimental setup the maximum and minimum value of the criterion function for the 10 runs are given. The number of failed identifications ($V(\theta) > 10^{-8}$) are reported together with the average number of criterion function evaluations used by the optimization algorithm where the average is formed over the successful identifications.

From the results it is clear that both the tridiagonal and the compact tridiagonal parametrizations lead to a higher probability to reach the global minimum when the model complexity increases. It is also interesting to see that although more parameters are estimated the number of evaluations are less for the tridiagonal forms. Since the evaluation of the criterion function is a large part of the computational load, the non-minimal tridiagonal parametrizations are more efficient.

5 Conclusions

Various parametrizations of multivariable state-space models for system identification are studied. Examples of minimal parametrizations are observable and controllable canonical forms. These have the important drawback that different structures must be used to cover all possible systems of a given order. Hence a priori structural knowledge is needed. To avoid these problems non-minimal parametrizations must be employed. One such is the full parametrization of the state-space matrices. This form has the drawback that it requires a large number of redundant parameters ($\sim n^2$).

A real tridiagonal and a compact tridiagonal form are

proposed as alternative parametrization forms. It is shown that any real square matrix is similar to a real (compact) tridiagonal form. The number of redundant parameters ($\sim 2n$) is in this parametrization case significantly less than in the full parametrization. When the number of states are large it is important, for efficiency, to reduce the number of parameters. Monte Carlo simulations indicate that the introduced parametrizations improve the optimization performance compared to observable or controllable canonical forms.

References

- [1] J. M. C. Clark. The consistent selection of parametrizations in system identification. In *Proc. Joint Automatic Control Conference (JACC)*, pages 576–580, Purdue University, Lafayette, Indiana, 1976.
- [2] A. Grace. *Optimization Toolbox*. The Mathworks, Inc, 1992.
- [3] R. Guidorzi. Canonical structures in the identification of multivariable systems. *Automatica*, 11:361–374, 1975.
- [4] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, NY, 1985.
- [5] T. Kailath. *Linear Systems*. Prentice-Hall, Englewood Cliffs, New Jersey, 1980.
- [6] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, Orlando, Florida 32887, 1985.
- [7] D. G. Luenberger. Canonical forms for linear multivariable systems. *IEEE Trans. on Automatic Control*, AC-12:290, 1967.
- [8] J. M. Maciejowski and C. T. Chou. Applications of estimation using balanced parametrizations. In *Proc. Third European Control Conference*, volume 2, pages 1348–1351, Rome, Italy, 1995.
- [9] T. McKelvey. Fully parametrized state-space models in system identification. In *Proc. of the 10th IFAC Symposium on System Identification*, volume 2, pages 373–378, Copenhagen, Denmark, July 1994.
- [10] R. J. Ober. Balanced realizations: Canonical form, parametrization, model reduction. *Int. J. Control*, 46(2):643–670, 1987.
- [11] B. N. Parlett. Reduction to tridiagonal form and minimal realizations. *SIAM J. Matrix Anal. Appl.*, 13(2):567–593, April 1992.
- [12] A. J. M. van Overbeek and L. Ljung. On-line structure selection for multivariable state space models. *Automatica*, 18(5):529–543, 1982.
- [13] J. H. Wilkinson. *The Algebraic Eigenvalue Problem*. Oxford University Press, London, 1965.