C22 IFAC World Congress 1984 C 22

STRUCTURE IDENTIFICATION FOR MULTIVARIABLE LINEAR SYSTEMS

Panel Discussion

Chairman: Michel Gevers
Australian National University, Canberra, Australia

Members: Roberto Guidozzi
Universita di Bologna, Bologna, Italy

Gilberto Correa
University of Oxford, Oxford, United Kingdom

Manfred Deistler
Technische Universitat Wien, Vienna, Austria

Keywords: identification, multivariable systems

The objective of the panel discussion was to evaluate the relative merits of different methods of parametrizing linear multivariable systems.

M. Gevers first presented the three alternative approaches to be discussed by the panel members:
1) estimation of the structure indices (i.e. Kronecker indices) from the data, and subsequent use of a canonical form.
2) use of one out of a finite set of candidate pseudocanonical (or overlapping) forms.
3) use of fully parametrized forms.

The panel considered linear finite-dimensional systems mapping input sequences $\{u(t),\ t \in T,\ u(\cdot) \in R^m\}$ into output sequences $\{y(t),\ t \in T,\ y(\cdot) \in R^p\}$. $M(n)$ denotes the set of all such systems of order n. $\bar{M}(n)$ denotes the set of all such systems of order less than or equal to n. ($\bar{M}(n)$ is the closure of $M(n)$).

Two finite dimensional models were considered:
1) State space models:

$$x(t+1) = F\,x(t) + G\,u(t)$$
$$y(t) = H\,x(t) + K\,u(t)$$

2) ARMA models

$$A(z^{-1})\,y(t) = B(z^{-1})\,u(t)$$

where

$$A(z^{-1}) = A_0 + A_1 z^{-1} + \cdots + A_s z^{-s}$$
$$B(z^{-1}) = B_0 + B_1 z^{-1} + \cdots + B_q z^{-q}$$

The transfer function $H(z) \triangleq \sum_{i=0}^{\infty} H_i z^{-i}$ of the system is related to the finite dimensional descriptions by

$$H(z) = H(zI-F)^{-1}G + K = A(z^{-1})^{-1}B(z^{-1}).$$

The problem discussed by the panel was that of finding parametrizations of H,F,G,K (alternatively $A(z^{-1})$, $B(z^{-1})$) that are identifiable from input-output data. Three aproaches were considered.

First approach: Canonical forms.
In this approach, there corresponds to each system in $M(n)$ a unique set of structure indices $n_1, \cdots, n_p$ (also called Kronecker indices). They define a multiindex $\alpha = (n_1, \cdots, n_p)$, with $\Sigma\, n_i = n$, the order of the system. Once the Kronecker indices are known, canonical (state-space or ARMA) forms can be used, whose parameters are identifiable from I/O data. The set of free parameters is denoted $\Delta_\alpha^{(1)}$ for the canonical state space form (observer form), $\Theta_\alpha^{(1)}$ for the canonical ARMA form (echelon form), while the set of systems that has Kronecker indices α is denoted $U_\alpha^{(1)}$. Canonical forms have been studied by Popov (1972), Denham (1974), Rissanen (1974), Guidorzi (1981) and many others. R. Guidorzi pointed out some advantages of the canonical form approach:
- once the structure is estimated, the subsequent parameter estimation is efficient, because the number of parameters is miminal.
- in the procedure developed by Guidorzi (Guidorzi, 1981; Guidorzi and Beghelli, 1982), the structural estimation is performed using tools essentially similar to the single input single output case.

R. Guidorzi illustrated his procedure with some practical applications and pointed out that, even if the procedure is unable to estimate the optimal structure from I/O data, it can at least eliminate some bad structures.
M. Deistler pointed out that the BIC structure selection criterion gives consistent estimators $\hat{\alpha}_T$ of the Kronecker indices α (Hannan, 1981; Hannan and Kavalieris, 1984).
G. Correa explained that with canonical parametrizations one can face numerical problems when the true (or optimal) system lies close to the border of the generic parametrization. This led him to advocate the use of pseudocanonical forms.

Second Approach: Pseudocanonical (or overlapping) forms.
With this approach, the set of structure indices can be selected almost arbitrarily. The reason is that all pseudocanonical parametrizations are generic. This provides maximum flexibility. For a given set $\alpha = (n_1, \cdots, n_p)$ of structure indices, the set of free parameters is denoted $\Delta_\alpha^{(2)}$ for the pseudocanonical state space form, $\Theta_\alpha^{(2)}$ for the pseudocanonical ARMA form. The set of systems representable in these forms is denoted $U_\alpha^{(2)}$. Technical details about pseudocanonical forms can be found in Rissanen and Ljung (1975), Correa and Glover (1982), Wertz, Gevers and Hannan (1982), Deistler (1983a,b), Correa and Glover (1984), Hannan and Kavalieris (1984), Gevers and Wertz (1984).
The structure estimation problem is eliminated

(except for order estimation) but, as G. Correa pointed out, the question now is to choose a parametrization that is most suitable for a given identification set up. This amounts to constructing selection criteria. G. Correa explained that criteria could be based either on statistical properties of the model estimators or on numerical considerations.

Concerning statistical properties, G. Correa made the following remarks:

1) Consistency - like considerations will not discriminate between different parameterizations, because one can approximate the true sytem as close as desired using any generic parameterization.

2) As for asymptotic accuracy:
   - if an I/O invariant function is minimized (e.g. a prediction error (PE) method), then for any two generic pseudocanonical parameterizations, any continuous I/O function of the model estimates will converge in probability to the same function, independently of the chosen parameterization.
   - if a non-optimal instrumental variable (IV) method is used, the choice of parameterization can affect the asymptotic covariance matrix of the Markov parameter estimates.

Since, for PE methods, all generic parameterizations are equally preferable from a statistical point of view, numerical considerations should prevail. In this respect, G. Correa made the following remarks:

- for PE methods, one should look at the conditioning of the Hessian of the criterion estimated at each iteration. Van Overbeek and Ljung (1982) have shown that this is related to the steady-state state-covariance matrix.
- for IV methods, one should consider the conditioning of the Gramian of the instrumental product moment matrices.

Third Approach: Fully parametrized ARMA forms. M. Deistler presented two fully parametrized ARMA forms and their identifiability properties. For brevity, we mention just one. Let $\Theta_{s,q}$ be the set of all pairs $(A(z^{-1}),B(z^{-1}))$ such that

- s,q are the maximum lag lengths of $A(z^{-1})$ and $B(z^{-1})$ respectively
- $A_0 = B_0 = I$
- A and B are left prime
- rank $[A_s,B_q] = p$

Let π: $\Theta_{s,q} \to U_{s,q}$: $\pi(A,B) = A^{-1}(z^{-1})B(z^{-1})$, where $U_{s,q}$ is the image of $\Theta_{s,q}$.

M.Deistler explained that $\Theta_{s,q}$ is identifiable. There are systems H(z) in M(n) for which there are no s,q with $H(z) \in U_{s,q} = \pi(\Theta_{s,q})$. However, such H(z) can be approximated with arbitrary accuracy in $U_{s,q}$ for some s and q. More details about fully parametrized forms can be found in Deistler (1983a,b) and P. Stoica (1982).

M. Deistler also gave a number of structure estimation results which apply to all three parametrizations. These results were given for the case where u(t) is unmeasurable zero mean white noise with covariance matrix Σ. They can be briefly summarized as follows.

Identification procedure: Given a finite family $(\Theta_\alpha \mid \alpha \in I)$, $\Theta_\alpha \in \Theta_A$ of identifiable classes and a set of observations {y(t)}, then the identification procedure computes, for every y(t), t=1, ..., T, an estimate $\hat{\alpha}_T$ of the true α_o and estimates $\hat{\theta}_T$ and $\hat{\Sigma}_T$ of the true $\theta_o \in \Theta_\alpha$ and Σ_o.

The usual identification procedures consist in 3 steps:

i) Estimation of α_o by $\hat{\alpha}_T$.

ii) Estimation of the true transfer function H_o over $\bar{U}_{\hat{\alpha}T}$ by $\hat{H}_T$, where

$U_\alpha = \pi(\Theta_\alpha)$, and of Σ_o by $\hat{\Sigma}_T$.

iii) Parametrization of $\hat{H}_T$.

The prototype estimation procedure for $\hat{H}_T$ and $\hat{\Sigma}_T$, whatever the parametrization, is maximum likelihood estimation (MLE). Under mild assumptions, the MLE $\hat{H}_T$ and $\hat{\Sigma}_T$ are strongly consistent and asymptotically normal (Hannan, Dunsmuir and Deistler, 1980).

The prototype estimation procedure for α is to minimize an expression of the form

$$A(\alpha) = \log \det \hat{\Sigma}_T(\alpha) + d_\alpha \frac{c(T)}{T}$$

over a finite number of α's (Akaike 1977, Rissanen 1978, Hannan and Kavalieris 1984), where $\hat{\Sigma}_T(\alpha)$ is the MLE of Σ_o over $\bar{U}_\alpha \times \{\Sigma \mid \Sigma > o\}$, d_α is the dimension of Θ_α. For C(T) = 2, $A(\alpha)$ is called AIC, for C(T) = c log T (with $c \geq o$), $A(\alpha)$ is called BIC.

M. Deistler presented the following consistency results, due to Hannan (1981) and Hannan and Kavalieris (1984). Under general assumptions:

1) For $\Theta_\alpha = \Theta_\alpha^{(1)}$ (i.e. canonical forms), BIC gives strongly consistent estimates $\hat{\alpha}_T$ for the Kronecker indices.

2) For $\Theta_\alpha = \Theta_\alpha^{(2)}$ (i.e. pseudocanonical forms), BIC gives a strongly consistent estimate $\hat{n}_T$ of the order n_o.

3) AIC is not consistent in the cases 1) and 2), but is optimal if the objective is spectral estimation.

In the discussion that followed these presentations, J. Bokor (Hungary) questioned whether the product moment test can estimate the Kronecker indices when the data are noisy. He advocated an alternative method based on the decomposition of a MIMO system into elementary subsystems, each subsystem being associated with one pole (or pair of complex-conjugate poles): see Bokor and Keviczky (1984) for details. P. Jansen (Netherlands) questioned the practical usefulness of methods which require to do a parameter estimation for a large class of models with an a posteriori model selection criterion. He thinks that structure selection procedures based on a prior analysis of I/O data are the only practical ones. M. Deistler replied that Hannan and Kavalieris (1984) have proposed a procedure which does not require to scan all admissible structures. In addition he pointed out that most a priori selection procedures require a prior assumption on the order of the system; in practice the order must also be estimated. Finally, I. Vajk (Hungary) suggested the use of multivariable lattice filters.

References

Akaike, H. (1977). On Entropy Maximisation Principle. In Applications of Statistics (P.R. Krishnaiah Ed.). North Holland.

Bokor, J. and Keviczky, L. (1984). Structure and parameter estimation of MIMO systems using elementary sub-system representation. *Int. J. Control*. Vol. 39, No. 5, pp. 965-986.

Correa, G. and Glover, K (1982). Multivariable identification using pseudo-canonical forms, *Proc. 6th IFAC Syump. on Identification and System Parameter Estimation* Washington D.C. pp. 1110-1115.

Correa, G. and Glover,K. (1984). Pseudo-canonical forms, identifiable parametrizations and simple parameter estimation for linear multivariable systems (2 parts). Accepted for publication in *Automatica.*

Deistler, M. (1983a). The properties of the parametrization of ARMAX systems and their relevance for structural èstimation. *Econometrica*, Vol.51, p. 1187-1207.

Deistler, M. (1983b). General Structure and Parametrizations of ARMA and State Space Systems and its Relation to Statistical Problems. In *Handbook of Statistics*, Vol. 5. North Holland.

Denham, M. (1974). Canonical forms for the identification of multivariable systems. *IEEE Trans. Aut. Contr* , AC-19, p. 646.

Gevers, M. and Wertz, V. (1982). On the problems of structure selection for the identification of stationary stochastic processes. *Proc. 6th IFAC Symp. on Identification and System Parameter Estimation*. Washington D.C., pp. 287-292.

Gevers, M. and Wertz, V. (1984). Uniquely identifiable state-space and ARMA parametrizations for multivariable linear systems. *Automatica*, Vol. 20.3.

Guidorzi (1981). Invariants and canonical forms for systems structural and parametric identification. *Automatica* Vol 17, p.117.

Guidorzi, R. and Beghelli, S. (1982). Input - output multistructural models in multivariable systems identification. *Proc. 6th IFAC Symp. on Identification and System Parameter Estimation.* Washington D.C., pp. 461-466.

Hannan, E.J. (1981). Estimating the dimension of a linear system. *J. Multivariate Analysis* Vol 11, pp. 459-473.

Hannan, E.J., Dunsmuir, W. and Deistler, M. (1980). Estimation of Vector ARMAX models. *J. Multivariate Analysis* Vol 10, pp. 275-295.

Hannan, E.J. and Kavalieris, L. (1984). Multivariate linear time series models. To appear in *Advances in Applied Probability*

Popov, V.M. (1972). Invariant description of linear, time-invariant controllable systems. *SIAM J. Control* Vol. 10, p. 252.

Rissanen, J. (1974). Basis of invariants and canonical forms for linear dynamic systems. *Automatica* Vol. 10, p. 175.

Rissanen, J. and Ljung, L. (1975). Estimation of optimum structures and parameters for linear systems. *Proc. CNR CISM Symp. Algebraic Syst. Theory*, Udine.

Stoica, P. (1982). On full matrix fraction descriptions. *Bull. Inst. Politch. Buc* ser. Electro.

Van Overbeek, A.J.M. and Ljung, L. (1982). On line structure selection for multivariable state space models. *Automatica* Vol 18, p. 529.

Wertz, V., Gevers, M. and Hannan, E.J. (1982). The determination of optimum structures for the state-space representation of multivariate stochastic processes. *IEEE Trans. Auto Contr.* AC-27, p. 1200.