FP-5-2 - 5:30

Proceedings of the 29th Conference on Decision and Control
Honolulu, Hawaii • December 1990

BALANCED PARAMETRIZATIONS IN TIME-SERIES IDENTIFICATION

B. Paul M^cGinnie Raimund J. Ober Jan M. Maciejowski
Engineering Dept., Trumpington Rd,
Cambridge, U.K. CB2 1PZ.

Abstract

The paper examines the identification of time-series, following a predictor-error approach using a new parametrization for predictors. This parametrization relies on a canonical form for minimum-phase systems which utilises a generalised notion of a balanced realisation. In this framework an identification scheme involving only unconstrained optimisations is developed, and an approach to the order selection problem is suggested.

1 Introduction

The problem which we shall consider is that of the identification of a p-component, wide sense stationary, purely linearly non-deterministic, full-rank stochastic process $(y_k)_{k\in Z}$ with a innovations representation

$$\left.\begin{aligned} x_{k+1} &= Ax_k + BD\nu_k \\ y_k &= Cx_k + D\nu_k \end{aligned}\right\} \qquad (1)$$

where $\lambda(A) \subseteq \mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$, $D > 0$, $(A,B,C,D) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times p} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{p\times p}$ is minimal, and $(\nu_k)_{k\in Z}$ is a p-component, identity covariance, orthogonal, wide sense stationary process.

Following Caines ([1] Chapter 8) we consider identification to be the procedure of finding the optimal predictor of the process. We shall consider mean square predictor-error optimality. Using a novel form of balancing we shall demonstrate a convenient parametrization of the appropriate predictors. This parametrization will also be shown to have important model reduction and augmentation properties which will lead to an approach to the model order selection problem. Other benefits of this method are described in the last section.

As mentioned, it is the least-square optimal predictor of the process $(y_k)_{k\in Z}$ which is of interest. This is given by the well known Kalman filter which, in this case, has a realization

$$\left.\begin{aligned} \hat{x}_{k+1} &= (A - BC)\hat{x}_k + By_k \\ \hat{y}_k &= C\hat{x}_k, \end{aligned}\right\} \qquad (2)$$

which corresponds to the zero solution of the Kalman filter Ricatti equation. From standard results (e.g. [1] Chapter 4) we also know that this system is stable.

Let us introduce two classes of systems which relate to time-series in innovations form.

•**Minimum-Phase Realizations:** If $(A,B,C,D) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times m}$ is minimal with $D \in \mathrm{GL}_m(\mathbb{R})$ (the group of invertible $m \times m$ matrices) then:

1) (A,B,C,D) is called a *discrete minimum-phase* realization if both $\lambda(A) \subseteq \mathbb{D}$ and $\lambda(A - BD^{-1}C) \subseteq \mathbb{D}$. The set of all such realizations will be denoted DM_n^m.

2) (A,B,C,D) is called a *continuous minimum-phase* realization if both $\lambda(A) \subseteq \mathbb{C}_- := \{z \in \mathbb{C} : \mathcal{R}e(z) < 0\}$ and $\lambda(A - BD^{-1}C) \subseteq \mathbb{C}_-$. The set of all such realizations will be denoted M_n^m. ∎

Thus the class of innovations realizations we are concerned with is the set $\mathcal{I}_n^p := \{(A,B,C,D) \in \mathrm{DM}_n^p : D > 0\}$ and the class of predictors is $\mathcal{P}_n^p := \{(A,B,C,D) \in \mathrm{DM}_n^p : D = I_p\}$. Notice that the D term is a parameter of the systems under consideration but not of the predictor.

2 Parametrisation of Predictors

Before proceeding we need the following:

•**Input-Output Equivalence and Canonical Forms:** If $(A,B,C,D),(\bar{A},\bar{B},\bar{C},\bar{D}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m} \times \mathbb{R}^{p\times n} \times \mathbb{R}^{p\times m}$ are both minimal then they are said to be *input-output equivalent* if there exists $T \in \mathrm{GL}_n(\mathbb{R})$ such that

$$(\bar{A},\bar{B},\bar{C},\bar{D}) = (TAT^{-1}, TB, CT^{-1}, D)$$

and such a T is called a *state space transformation.*

A *canonical form* for a set X with respect to an equivalence relation $\sim$ is a map $\Gamma : X \to X$ such that it is constant on each equivalence class and $\Gamma(x) \in [x]_\sim$, the equivalence class of x under the relation $\sim$. ∎

To produce a parametrization of systems in $\mathcal{P}_n^p$ we proceed to develop a canonical form/parametrization of continuous minimum-phase systems and then induce a parametrization of the discrete minimum-phase systems by means of the bilinear transform. The constraint $D = I_p$ is then imposed upon the induced parametrization to produce the required result.

As previously mentioned, we base our canonical form on the notion of a balanced realisation appropriate to minimum-phase systems. A method for producing balanced canonical forms for classes of linear systems is described in [2]. The following is the appropriate notion of balancing in this case.

•**Minimum-Phase Balanced:** If $(A,B,C,D) \in \mathrm{M}_n^m$ and we have the controllability Gramian W_c and the observability Gramian of the inverse system W_o given by

$$\begin{aligned} AW_c + W_cA^T &= -BB^T \\ (A - BD^{-1}C)^TW_o + W_o(A - BD^{-1}C) &= -C^TD^{-T}D^{-1}C \end{aligned}$$

then (A,B,C,D) is said to be *minimum-phase balanced, or mp-balanced* if $W_c = W_o = \Sigma = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_k I_{n_k})$ where $\sigma_1 > \ldots > \sigma_k > 0$. Σ is the *minimum-phase Gramian* and σ_i are the *minimum-phase characteristic values.* ∎

Proofs and further details of the subsequent results may be found in [3].

Theorem 1 (Existence of MP-Balanced Realisations)
For each $(A,B,C,D) \in \mathrm{M}_n^m$ there exists $T \in \mathrm{GL}_n(\mathbb{R})$ such that $(TAT^{-1}, TB, CT^{-1}, D)$ is minimum-phase balanced with minimum phase Gramian $\Sigma = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_k I_{n_k})$.

The realisation is unique up to state space transformation by an orthogonal matrix of the form $U = \mathrm{diag}(U_1, \ldots, U_k)$, $U_i \in \mathrm{O}_{n_i}(\mathbb{R})$ (the group of $n_i \times n_i$ orthogonal matrices) ∎

Corollary 1.1 *If $(A,B,C,D) \in \mathrm{M}_n^{m,m}$ is a minimum-phase balanced realisation with distinct minimum-phase characteristic values (i.e. $k = n$) then it is unique up to state space transformation by a sign matrix, $S = \mathrm{diag}(\pm 1, \ldots, \pm 1)$.* ∎

We state the following theorem for the single-input-single-output case for simplicity. The result generalises to the multivariable case in a natural way, although the number of parameters increases significantly.

Theorem 2 (Parametrization/Canonical Form)
In each state space equivalence class there exists exactly one realisation $(A(\theta), B(\theta), C(\theta), D(\theta))$ given by a set of parameters

$$\theta = \begin{pmatrix} k, n_1, \ldots, n_k, \sigma_1, \ldots, \sigma_k, b_1, \ldots, b_k, s_1, \ldots, s_n, d, \\ \alpha_1^{(1)}, \ldots, \alpha_{n_1-1}^{(1)}, \ldots, \alpha_1^{(k)}, \ldots, \alpha_{n_k-1}^{(k)} \end{pmatrix}$$

where $n \geq k \in \mathbb{N}$ and

$$\begin{array}{llllll} \sigma_1 > \ldots > \sigma_k > 0 & , & \text{where} & \sigma_i \in \mathbb{R}, \\ n_1, \ldots, n_k & , & " & n_i \in \mathbb{N};\ \sum_{j=1}^k n_j = n, \\ s_1, \ldots, s_k & , & " & s_i \in \{\pm 1\}, \\ \alpha_1^{(i)}, \ldots, \alpha_{n_i-1}^{(i)} & , & " & \alpha_j^{(i)} > 0,\ 1 \leq i \leq k, \\ b_1, \ldots, b_k & , & " & b_i > 0, \\ d \neq 0. \end{array}$$

The realization $(A(\theta), B(\theta), C(\theta), D(\theta))$ is mp-balanced and is given by:

1) $$A(\theta) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix} \quad \textit{where}$$

$$A_{ij} = \begin{bmatrix} \frac{-\sqrt{1+\sigma_i^2}\, b_i b_j}{\sigma_j\sqrt{1+\sigma_i^2}+\sigma_i s_i s_j\sqrt{1+\sigma_j^2}} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times n_j},\ \textit{if } i \neq j,$$

$$\textit{and } A_{ii} = \begin{bmatrix} \frac{-b_i^2}{2\sigma_i} & \alpha_1^{(i)} & 0 & \cdots & \cdots & 0 \\ -\alpha_1^{(i)} & 0 & \alpha_2^{(i)} & \cdots & \cdots & 0 \\ 0 & -\alpha_2^{(i)} & 0 & \ddots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 & \alpha_{n_i-1}^{(i)} \\ 0 & 0 & 0 & \cdots & -\alpha_{n_i-1}^{(i)} & 0 \end{bmatrix} \in \mathbb{R}^{n_i \times n_i}.$$

2) $B(\theta)^T = (\underbrace{b_1, 0, \ldots, 0}_{n_1}, \ldots, \underbrace{b_j, 0, \ldots, 0}_{n_j}, \ldots, \underbrace{b_k, 0, \ldots, 0}_{n_k}).$

3) $C(\theta) = d\left(\overbrace{b_1(\sigma_1 + s_1\sqrt{1+\sigma_1^2}), 0, \ldots, 0}^{n_1}, \ldots, \underbrace{b_k(\sigma_k + s_k\sqrt{1+\sigma_k^2}), 0, \ldots, 0}_{n_k}\right)$ *4)* $D(\theta) = d.$ ∎

Theorem 3 (Model Reduction Property)
If $\left(\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}, \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \begin{bmatrix} C_1 & C_2 \end{bmatrix}, D\right) \in M^{\mathrm{m}}_{n_1+n_2}$ *is a mp-balanced realisation in canonical form with minimum-phase Gramian* $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_k I_{n_k}) \in \mathbb{R}^{n_1+n_2 \times n_1+n_2}$ *then* (A_{11}, B_1, C_1, D) *is a mp-balanced realisation in canonical form with minimum-phase Gramian* Σ_1. ∎

Remark 3.1 *This canonical form also has an augmentation property, that is, by adding new parameters* σ_i, b_i, s_i *and/or* $\alpha_i^{(j)}$ *we may increase the order of the model without effecting the other states. It is this property which we hope to take advantage of in the identification scheme to be discussed in §3.*

We have a parametrization of $\mathrm{M}^{\mathrm{m}}_{\mathrm{n}}$ and now we induce one on $\mathrm{DM}^{\mathrm{m}}_{\mathrm{n}}$ via the following Möebius transform of the complex plane.

Proposition 4 *The map* $\mathcal{B}^{\mathrm{m}}_{\mathrm{n}} : \mathrm{M}^{\mathrm{m}}_{\mathrm{n}} \to \mathrm{DM}^{\mathrm{m}}_{\mathrm{n}}$ *given by*

$$\mathcal{B}^{\mathrm{m}}_{\mathrm{n}}(A,B,C,D) = \begin{pmatrix} (I_n - A)^{-1}(I_n + A), \sqrt{2}(I_n - A)^{-1}B, \\ \sqrt{2}C(I_n - A)^{-1}, D + C(I_n - A)^{-1}B \end{pmatrix} \quad (3)$$

is a bijection preserving input-output equivalence. ∎

We have thus produced a canonical form for $\mathrm{DM}^{\mathrm{m}}_{\mathrm{n}}$ by the obvious application of $\mathcal{B}^{\mathrm{m}}_{\mathrm{n}}$.

It only remains to restrict the canonical form to the set $\mathcal{P}^{\mathrm{p}}_{\mathrm{n}}$. In the simple case given in Theorem 2 this reduces to the restriction that the canonical form for $\mathrm{DM}^{\mathrm{p}}_{\mathrm{n}}$ has D term equal to 1. From this and the expression for D in (3) we may calculate the parameter d. Thus we have a parametrization of $\mathcal{P}^1_{\mathrm{n}}$ in terms of the parameters of the canonical form for M^1_{n}. This has an obvious extention to the multivariable case.

3 Identification

We have seen that the model classes $\mathcal{I}^{\mathrm{p}}_{\mathrm{n}}$ are appropriate classes of stochastic realisation to consider for time-series identification. These classes give rise to sets $\mathcal{P}^{\mathrm{p}}_{\mathrm{n}}$ of steady state predictors and in §2 we have demonstrated a parametrization for $\mathcal{P}^1_{\mathrm{n}}$.

Assuming that our time series does have an innovations representation and that the assumption of steady state for the Kalman filter predictor is well founded then our identification scheme is composed of the following two steps:

1) Find the parameters (both discrete and continuous) of the optimal predictor. For the case described by 1 we know that an optimum exists.
2) Calculate the actual realisation. For our class of systems this is a simple process as all but the D term are given by the predictor. The value of the D term may me calculated from the observed variance of y_k and the expression for the covariance of the state process x_k.

The optimisation involved is step 1) is composed of two distinct parts: the selection of structural discrete parameters, most notably the system order, and the optimisation of the continuous parameters.

The optimisation of the continuous parameters may be handled by any standard method. As expressed in Theorem 2 the optimisation is constrained. However, by working with the logarithms of the positive variables and the logarithms of the differences of adjacent characteristic values, we may transform this to an unconstrained optimisation.

As a result of the parametrization we have analytic expressions for the gradients of the predictor error. A current area of research is that of using a symbolic manipulation package to provide some of the calculations necessary for the identification.

The structural selection is more complicated. Our approach to order selection is to start to identify a first order system and repeatedly increment the order of the system. The parameters are initialised from the optimum of order one less. This is repeated until a satisfactory combination of predictor error and model complexity is achieved. This is a further topic fo study as the complexity of the models is dependent on the structural parameters.

We have two new problems due to the use of the parametrization: how to choose s_i at each stage; and how to change the $(n_i)_{i=1}^k$ during the identification. Work is underway to develop a consistent estimator for the s_i. Our approach to the $(n_i)_{i=1}^k$ is to assume that the n_i all take value 1 (this is the "generic" case) and to alter this if two characteristic values move together.

4 Conclusions

Due to limitations of space we have only sketched the development of the material in the simplest case. However, we have derived an identification scheme which is composed of a set of unconstrained optimisations with a systematic choice of structural parameters.

The method has the following advantages:

i. We optimise with respect to a minimal set of parameters.
ii. The optimisation is unconstrained and at each step stability of the system and predictor is guaranteed.
iii. We have analytic expressions for the gradient and hessian of the predictor error.
iv. We are able to utilise the augmentation properties to give an approach to the selection of model order.

The method suffers from the disadvantages common to all identification schemes which use a stationary Kalman filter. Also, as we parametrize all equivalence classes of realisations in $\mathcal{P}^{\mathrm{p}}_{\mathrm{n}}$ we include some badly conditioned realisations. This however is not a flaw in the parametrization, and in fact it is hoped to show that analogously to the balanced parametrization for stable systems this method produces least badly conditioned realisations. The main disadvantage of this method is the proliferation of structural parameters.

This approach is currently being implemented.

References

[1] Peter E. Caines. *Linear Stochastic Systems.* Series in Probability and Mathematical Statistics. Wiley, New York, first edition, 1988.

[2] B. Paul M^cGinnie and Raimund J. Ober. Balanced canonical forms of classes of symmetric transfer functions with applications, 1990. Submitted to Linear Algebra and Its Applications.

[3] B.Paul McGinnie and Raimund J. Ober. Balanced realisations and time-series. To appear as a departmental report, Dept. of Engineering, University of Cambridge, Trumpington St., Cambridge CB2 1PZ, England.