Identification of noisy linear systems with multiple ARMA inputs

Tigelaar, H.H.

Publication date:
1983

Citation for published version (APA):
Tigelaar, H. H. (1983). *Identification of noisy linear systems with multiple ARMA inputs*. (Research memorandum / Tilburg University, Department of Economics; Vol. FEW 126). Unknown Publisher.

subfaculteit der econometrie

RESEARCH MEMORANDUM

[redacted]

TILBURG UNIVERSITY
DEPARTMENT OF ECONOMICS
Postbus 90153 - 5000 LE Tilburg
Netherlands

[redacted]

IDENTIFICATION OF NOISY LINEAR SYSTEMS WITH
MULTIPLE ARMA INPUTS

by

Harry H. Tigelaar

IDENTIFICATION OF NOISY LINEAR SYSTEMS WITH MULTIPLE ARMA INPUTS

by Harry H. Tigelaar

In this paper we give conditions for identifiability of the parameters of a linear system that is disturbed by moving average noise, where the inputs are generated by a multivariate ARMA-process. The identifiability is based on finite samples from in- and output.

1. INTRODUCTION

IN ORDER to avoid messy notation we shall use the lag-operator L when there is no impact on mathematical rigor. Consider the model

$$(1) \qquad A(L)y_t = B(L)x_t + C(L)\varepsilon_t \qquad (t \in \mathbb{Z}),$$

where $\{x_t\}$ is an m-variate weakly stationary (observable) process of inputs, $\{y_t\}$ is the (observable) scalar output process and $\{\varepsilon_t\}$ is a (non observable) white noise sequence with $E\varepsilon_t = 0$ and $E|\varepsilon_t|^2 = \sigma^2 > 0$. Furthermore, $A(z) = \Sigma_{k=0}^{p} a_k z^k$ and $C(z) = \Sigma_{k=0}^{q} c_k z^k$ are scalar polynomials with $a_0 = c_0 = 1$ and $B(z) = \Sigma_{k=0}^{r} z^k$ is a 1×m matrix of polynomials. Usually the processes $\{x_t\}$ and $\{\varepsilon_t\}$ are supposed to be independent, but since we are mainly concerned with second-order properties, we shall only assume orthogonality.

The integers p, q and r are supposed to be a priori known. We are interested in second-order informative samples for $(A(z), B(z), \sigma^2)$, (see [4, Ch. 1] or [5]). If the samples $(x_a, x_{a+1}, \ldots, x_b)$ and $(y_c, y_{c+1}, \ldots, y_d)$ are jointly second-order informative for a function ψ of the unknown parameters of the system, then we shall say that the *sampling-scheme* $< a,b|c,d >$ is second-order informative for ψ.

Notice, that $\{x_t\}$ and $\{y_t\}$ are jointly weakly stationary processes since $\{x_t\}$ and $\{\varepsilon_t\}$ are orthogonal. Therefore any sampling scheme admits a shift in time .

Although we are only interested in identifiability with respect to finite samples, our main tool will be the analysis of spectral density matrices. Therefore we can, without complicating de discussion, all variables and coefficients allow to take complex values, and we shall transposition of a matrix or vector always combine with complex conjugation, denoted by an asterisk.

2. PRELIMINARY RESULTS

Suppose that $A(z)$ has no zeros on $|z| = 1$.
Let f_x denote the spectral density matrix of the input process $\{x_t\}$. Then it follows from the orthogonality of the processes $\{x_t\}$ and $\{\varepsilon_t\}$ that the joint spectral density matrix of the process $\{(y_t,x_t)\}$ is

$$\begin{bmatrix} f_y & g_{xy} \\ g_{xy}^{*} & f_x \end{bmatrix},$$

where

$$(2) \qquad f_y(\lambda) = \frac{B(1^{-i\lambda})f_x^{(\lambda)}\ B^{*}(e^{-i\lambda})}{|A(e^{-i\lambda})|^2} + \frac{\sigma^2}{2\pi}\left|\frac{C(e^{-i\lambda})}{A(e^{-i\lambda})}\right|^2, \quad (-\pi < \lambda \leq \pi)$$

and

$$(3) \qquad g_{xy}(\lambda) = \frac{1}{A(e^{-i\lambda})} B(e^{-i\lambda}) f_x(\lambda), \qquad (-\pi < \lambda \leq \pi) .$$

Before we can deal with the general case with ARMA-inputs, we shall derive an informative sampling scheme for the case that we have MA-inputs. Thus we have

$$(4) \qquad f_x(\lambda) = \frac{1}{2\pi} Q(e^{-i\lambda})\Omega\, Q^{*}(e^{-i\lambda}), \qquad (-\pi < \lambda \leq \pi),$$

where Q is a polynomial with $Q(0) = I_m$ (the m×m unit matrix), and Ω is some unknown, hermitian positive definite matrix. The degree of Q is a priori known not to exceed some known integer w. It is well known that a sample of size w+1 from the process $\{x_t\}$ is second-order informative for f_x and if det $Q(z)$ is supposed to have no zeros in $|z| < 1$ it is also second-order informative for $(Q(z),\Omega)$ (see [4, th. 3.3.1]).

LEMMA 1: Let $U(z)$ and $V(z)$ be polynomials of degree u and v respectively, and $V(0) \neq 0$. Then the rational function $U(z)/V(z)$ is uniquely determined by its first $u + v + 1$ Taylor coefficients.

PROOF: Suppose there exist $\tilde{U}(z)$ and $\tilde{V}(z)$ satisfying the conditions of the lemma, such that $\tilde{U}(z)/\tilde{V}(z)$ has the same first $u + v + 1$ Taylor coefficients as $U(z)/V(z)$. Then we can write

$$\frac{U(z)}{V(z)} - \frac{\tilde{U}(z)}{\tilde{V}(z)} = \sum_{j=u+v+1}^{\infty} \psi_j z^j$$

for all z in some annulus $\mathcal{A}$ around the origin.

Hence $z^{-u-v-1}[U(z)/V(z) - \tilde{U}(z)/\tilde{V}(z)]$ is holomorphic on $\mathcal{A}$ and so is $z^{-u-v-1}[U(z)\tilde{V}(z) - \tilde{U}(z)V(z)]$.
This, however, is a contradiction since degree $[U(z)\tilde{V}(z) - \tilde{U}(z)V(z)] \leq u + v$. □

We can now state the following theorem.

THEOREM 1: Suppose $A(z) \neq 0$, $|z| \leq 1$ and let the inputs be generated by an m-variante MA(w) process with a.e. positive definite spectral density matrix. Then the sampling scheme $< 0,w|0,w+r+p >$ is second-order informative for the rational function $B(z)/A(z)$.

PROOF: Let the spectral density matrix of the input process be given by (4). Notice that $f_x(\lambda)$ is the boundary value of the function $Q(z)\Omega\, Q^*(\frac{1}{z})$ and that $z^w Q(z)\Omega\, Q^*(\frac{1}{z})$ is a (scalar) polynomial of degree $\leq 2w$.
We have

$$\tau_k := E\, y_t\, x^*_{t-k} = \int_{-\pi}^{\pi} e^{ik\lambda}\, g_{xy}(\lambda)d\lambda =$$

$$= \frac{1}{2\pi i} \int_{|z|=1} z^{-k-1-w} \frac{B(z)}{A(z)} [z^w\, Q(z)\Omega\, Q^*(\frac{1}{z})]\, dz, \qquad (k \in \mathbb{Z}).$$

Since $A(z)$ has no zeros for $|z| \leq 1$, the rational function

$\varphi(z) = B(z)/A(z)[z^w Q(z)\Omega\, Q^*(\frac{1}{\bar{z}})]$ is holomorphic (component wise) on $|z| < \rho$ for some $\rho > 1$ and so the sequence $\{\tau_k\}_{k=-w}^{\infty}$ is the sequence of Taylor coefficients of $\varphi(z)$.
Hence, by lemma 1, $\varphi(z)$ is uniquely determined by $\tau_{-w}, \tau_{-w+1}, \ldots, \tau_{w+r+p}$. Consequently, the sampling scheme $< 0,w|0,w+r+p >$ is second-order informative for $\varphi(z)$. On the other hand, the sampling scheme $< 0,w|0,w+r+p >$ is second-order informative for $e^{-iw\lambda} f_x(\lambda)$ i.e. the boundary value of $z^w Q(z)\Omega\, Q^*(\frac{1}{\bar{z}})$. But then it is also second-order informative for $B(e^{-i\lambda})/A(e^{-i\lambda})$ (because $f_x > 0$ a.e.) and since $B(z)/A(z)$ is uniquely determined by its values on $|z| = 1$ (since $B(z)/A(z)$ is holomorphic on $|z| < \rho$), it follows that it is second-order informative for $B(z)/A(z)$. □

For the next step we rewrite the model (1) as

$$\text{(5)} \qquad A(L)u_t = C(L)\varepsilon_t, \qquad (t \in \mathbb{Z}),$$

where

$$\text{(6)} \qquad u_t = y_t - A^{-1}(L)\, B(L)x_t, \qquad (t \in \mathbb{Z}).$$

Notice, that the expression $A^{-1}(L)$ only makes sense when $A(z)$ is supposed to have no zeros in $|z| \leq 1$.
If the process $\{u_t\}$ were observable, a sample of size p+q+1 would be sufficient for identifying $(A(z), C(z), \sigma^2)$ since we could apply the theorem on ARMA (p,q) processes in [4, th. 2.5.3]. However, even if $B(z)/A(z)$ is identified, a finite sample from $\{u_t\}$ requires the observation of an infinite sequence of the process $\{x_t\}$.
Let $\gamma_s = E\, u_t\, \bar{u}_{t-s}$, $(t,s \in \mathbb{Z})$.

LEMMA 2: Under the conditions of theorem 1, the sampling scheme $< 0,\max(w,p+q+r)|0,p+r+\max(w,q) >$ is second-order informative for the sequence $\{\gamma_s\}$.

PROOF: Let $\tilde{u}_t = y_t - A^{-1}(L)B(L)\tilde{x}_t$, $(t \in \mathbb{Z})$,

where

$$\tilde{x}_t = \begin{cases} x_t & \text{if} \quad t \geq 0 \\ 0 & \text{if} \quad t < 0 \end{cases} .$$

Then $\quad \tilde{u}_t = u_t + A^{-1}(L)B(L)(x_t - \tilde{x}_t), \quad (t \in \mathbb{Z}),$

and $\quad \tilde{u}_t = u_t$ if $t \geq r$. Hence for $s = 0,1,\ldots,p+q$

$$\nu_{t,s} = E\, \tilde{u}_t\, \bar{\tilde{u}}_{t-s} = \gamma_s,$$

if $t \geq p+q+r$. Thus, if $\{\tilde{u}_t\}$ were observable, the sample $(\tilde{u}_r, \tilde{u}_{r+1},\ldots,\tilde{u}_{p+q+r})$ would be second-order informative for $(\gamma_0, \gamma_1,\ldots,\gamma_{p+q})$. But then it follows that the sampling scheme $< 0,p+q+r|r,p+q+r >$ is second-order informative for $(\gamma_0,\gamma_1,\ldots,\gamma_{p+q})$ conditional on $A^{-1}(z)B(z)$ (see [4, p. 11], or [5] where the concept of conditional identifiability is introduced).
Since the sequence $\{\gamma_s\}$ is uniquely determined by $(\gamma_0,\ldots,\gamma_{p+q})$ (see [4, th. 2.5.2]) it follows that the sampling scheme $< 0,p+q+r|r,p+q+r >$ is second-order informative for $\{\gamma_s\}$ conditional on $A^{-1}(z)B(z)$. Since the sampling scheme $< 0,w|0,w+r+p >$ is second-order informative for $A^{-1}(z)B(z)$ by theorem 1, it follows by the conditional identification theorem ([4, th. 1.2.3]) that the sampling scheme $< 0,\max(w,p+q+r)|0,p+r+\max(w,q) >$ is second-order informative for $\{\gamma_s\}$. □

THEOREM 2: Let the conditions of theorem 1 be satisfied, and suppose $C(z) \neq 0$, $|z| < 1$. *Then the sampling scheme* $< 0,\max(w,p+q+r)|0,p+r+\max(w,q) >$ *is second-order informative for* $(\sigma^2, C(z)/A(z))$.

PROOF: From (5) we see that $\{u_t\}$ is an univariate ARMA (p,q) process with spectral density

$$f_u(\lambda) = \frac{\sigma^2}{2\pi} \left| \frac{C(e^{-i\lambda})}{A(e^{-i\lambda})} \right|^2, \quad (-\pi < \lambda \leq \pi).$$

By lemma 2 we know that the sampling scheme is second-order informative for the covariance function $\{\gamma_s\}$ and so it is also second-order informative for f_u. Since $A(z)$ has no zeros in $|z| \leq 1$, the function $C(z)/A(z)$

is holomorphic on $|z| < \rho$ for some $\rho > 1$. But then f_u determines $\sigma\, C(z)/A(z)$ uniquely (see [4, lemma 2.1.1]) and since $C(0) = A(0) = 1$, both σ^2 and $C(z)/A(z)$ are uniquely determined by f_u. This proves the theorem. □

THEOREM 3: Let the conditions of theorem 2 be satisfied, and suppose $(A(z),B(z),C(z)) \neq 0$ for all z. Then the sampling scheme $< 0, \max(w,p+q+r) | 0, p+r+\max(w,q) >$ is second-order informative for $(\sigma^2, A(z), B(z), C(z))$.

PROOF: From theorems 1 and 2 we obtain, that the sampling scheme given in the theorem is second-order informative for $(\sigma^2, B(z)/A(z), C(z)/A(z))$. Since $(A(z),B(z),C(z)) \neq 0$ for all z, it follows that $A(z)$ cannot have a factor in common with all components of $(B(z),C(z))$. This implies that $A(z)$ can be determined uniquely from $(B(z)/A(z), C(z)/A(z))$, and so it is identified. But then also $B(z)$ and $C(z)$ are identified and the theorem is proved. □

3. ARMA INPUTS

Let the m-variate input process be generated by

$$(7) \qquad P(L)x_t = Q(L)\eta_t \qquad (t \in \mathbb{Z}),$$

where $\{\eta_t\}$ is m-variate white noise with $\Omega_\eta = E\,\eta_t\,\eta_t^* > 0$. The m×m matrix polynomials $P(z) = \Sigma_{k=0}^{n} P_k z^k$ and $Q(z) = \Sigma_{k=0}^{w} Q_k z^k$ are unknown, the integers n and w are known. Although we are primarily interested in the system (1), we shall suppose that a sample of size N from the process $\{x_t\}$ is second-order informative for $P(z)$. For sufficient conditions and the determination of N we refer to [4, Ch. 3]. In the special case of m independent univariate ARMA inputs, the conditions are simpler and N can be smaller; see [4, Ch. 2]. Strictly spoken it is not necessary to have identifiability of $Q(z)$; compare the case with MA-inputs.

Put $z_t = P(L)x_t$, $(t \in \mathbb{Z})$. Then, clearly $\{z_t\}$ is not observable since $P(z)$ is unknown. However, if a sample of size n_0 from $\{z_t\}$ were second-order informative for some unknown parameter, then a sample of size $n_0 + n$ from $\{x_t\}$ is second-order informative conditional on $P(z)$ for that parameter. On the other hand, a sample of size N from $\{x_t\}$ was supposed to be second-order informative for $P(z)$ and so, by the conditional identification theorem ([4, th. 1.2.3]) a sample of size $\max(N, n_0+n)$ of $\{x_t\}$ is second-order informative for that parameter. Thus we may consider $\{z_t\}$ as observable, but we must keep in mind, that the sampling scheme $< a,b|c,d >$ for $\{z_t, y_t\}$ corresponds to $< a-n,b|c,d >$ for $\{x_t, y_t\}$.

The discussion above enables us to reduce the problem with ARMA-inputs to the problem with MA-inputs. Before we state and prove the final theorem, we shall list the restrictions on unknown parameters, which will be used to identify the system (1) with inputs generated by (7).

IDENTIFIABILITY CONDITIONS

(A) $A(z) \neq 0, \quad |z| \leq 1,$

(B) $C(z) \neq 0, \quad |z| < 1,$

(C) $(A(z), B(z), C(z)) \neq 0, \quad z \in \mathbb{C},$

(D) $\det P(z) \neq 0, \quad |z| \leq 1,$

(E) Any set of conditions guaranteeing a sample of size N from $\{x_t\}$ to be second-order informative for $P(z)$

(F) $Q(e^{-i\lambda})\Omega_n Q^*(e^{-i\lambda}) > 0$ a.e.

We can now prove

THEOREM 4: Under the conditions (A)-(F) the sampling scheme $< -n, \max(N-n-1, w, p+q+r+n(3m-1)) | 0, p+r+n(2m-1)+\max(w, q+nm) >$ *is second-order informative for* $(A(z), B(z), C(z), \sigma^2)$.

PROOF: Let Adj $P(z)$ denote the adjoint matrix polynomial of $P(z)$, i.e. Adj $P(z) = P^{-1}(z)\det P(z)$. Since its elements are determinants of

$(m-1)\times(m-1)$ sub-matrices of $P(z)$, it follows that

$$\text{degree}[\text{Adj } P(z)] \leq (m-1)n.$$

Put $T(z) = \det P(z)$. Then degree $[T(z)] \leq mn$ and substituting $z_t = P(L)x_t$ we can write

$$(8) \qquad T(L)A(L)y_t = B(L)\text{Adj } P(L)z_t + T(L)C(L)\varepsilon_t, \quad (t \in \mathbb{Z})$$

Put $\quad \tilde{A}(z) = T(z)A(z)$, degree $[\tilde{A}(z)] = \tilde{p}$

$\tilde{B}(z) = B(z)\text{Adj } P(z)$, degree $[\tilde{B}(z)] = \tilde{r}$

$\tilde{C}(z) = T(z)C(z)$, degree $[\tilde{C}(z)] = \tilde{q}$.

We have $\tilde{p} \leq p+nm$, $\tilde{r} \leq r+(m-1)n$ and $\tilde{q} \leq q+nm$. Furthermore, $\{z_t\}$ is a m-variate MA(w) process, which may be considered as observable by the discussion given above. Since also $\tilde{A}(z) \neq 0$, $|z| \leq 1$ by conditions (A) and (D), we can apply theorem 1, which for the process $\{z_t,y_t\}$ implies that the sampling scheme $< 0,w|0,w+r+(m-1)n+p+nm >$ is second-order informative for $\tilde{B}(z)/\tilde{A}(z)$. Hence, for the process $\{x_t,y_t\}$ the sampling scheme $< -n,\max(N-n-1,w)|0,w+r+p+n(2m-1) >$ is second-order informative for $\tilde{B}(z)/\tilde{A}(z)$ and so, by condition (E) for $B(z)/A(z)$.
In a similar way it follows from theorem 2 that (for the process $\{x_t,y_t\}$) the sampling scheme

$$< -n,\max(N-n-1,w,p+q+r+n(3m-1)|0,p+r+n(2m-1)+\max(w,q+nm) >$$

is second-order informative for $(\sigma^2,\tilde{C}(z)/\tilde{A}(z)) = (\sigma^2,C(z)/A(z))$. Since it includes the sampling scheme obtained for $B(z)/A(z)$, it follows that it is also second-order informative for $(\sigma^2,B(z)/A(z),C(z)/A(z))$. But then, by condition (C), it follows as in theorem 3, that it is second-order informative for $(\sigma^2,A(z),B(z),C(z))$. □

REMARK: For $n = 0$, also $N = 0$ and the result of theorem 3 is obtained.

In particular we shall consider the case $n \geq 1$, where $N = w+(m+1)n$, which is the informative sample size for the parameters of the input-process obtained in [4, p. 93]. In that case we have $w \leq N-n-1 = w+mn-1$. If $w \leq q+nm$ then also $N-n-1 \leq q+2mn-1 \leq p+q+r+n(3m-1)$ and the sampling scheme reduces to

$$< -n, p+q+r+n(3m-1) \,|\, 0, p+q+r+n(3m-1) >$$

Thus, roughly speaking one can say that under the conditions (A)...(E), for $\{x_t, y_t\}$ the sample size p+q+r+3nm+1 is second-order informative for the system.

4. DISCUSSION

There are several ways of generalizing the model given by (1). We shall discuss three of them in a nutshell. At first, we can allow for errors in the variables. Such models were treated by MARAVALL [3] in 1979, who gives conditions for local identifiability based on infinite samples. Except for $C(z)$ and σ^2, there seems to be no real problem in obtaining global results for finite samples. Second, we can allow more general inputs, e.g. non-stationary inputs. Generally, difficult problems may arise, but the results in this paper can easily be generalized to the case of ARIMA inputs. Finally, consider the multivariate analogue of (1). The infinite sample case was treated by HANNAN [1] in 1971, and some results on local identifiability were given by KOHN [2] in 1980. However, the finite sample case presents some problems requiring further research.

REFERENCES

1. HANNAN, E.J. (1971) The identification problem for multiple equation systems with moving average errors. Econometrica, 39.5.

2. KOHN, R. (1980) Local identification of ARMAX structures subject to nonlinear constraints. Metrika, 27.

3. MARAVALL, A. (1979) Identification in dynamic shock error models, Springer, Berlin.

4. TIGELAAR, H.H. (1982) Identification and informative sample size. Mathematical Centre Tracts 147, Amsterdam.

5. TIGELAAR, H.H. (1982) The informative sample size for dynamic multiple equation systems with moving average errors. Research Memorandum 113. Tilburg.

IN 1982 REEDS VERSCHENEN:

107 Aart J. de Zeeuw
Hierarchical decentralized optimal control in econometric policy models.

108 Arie Kapteyn en Tom Wansbeek
Identification in Factor Analysis.

109 G. van der Laan en A.J.J. Talman
Simplical Algorithms for finding Stationary Points, a unifying description

110 Pieter Boot
Ekonomiese betrekkingen tussen Oost en West Europa.

111 B.B. van der Genugten
The asymptotic behaviour of the estimated generalized least squares method in the linear regression model.

112 Jack P.C. Kleijnen / Anton J. van Reeken
Principles of computer charging in a university-like organization.

113 H. Tigelaar.
The informative sample size for dynamic multiple equation systems with moving average errors.

114 Drs. H.G. van Gemert, Dr. R.J. de Groof en Ir. A.J. Markink
Sektorstruktuur en Economische Ontwikkeling

115 P.H.M. Ruys
The tripolar model: a unifying approach to change.

116 Aart J. de Zeeuw
Policy solutions for Mini-Interplay, a Linked Model for two Common Market Countries.

IN 1982 REEDS VERSCHENEN (vervolg):

117 F.J.M. van Doorne en P.H.M. Ruys
Die Struktur einer Sprechhandlung in Habermas' Forschungsprogramm. Formale Analyse mit den Mitteln des tripolaren Modells.

118 A.J.J. Talman en G. van der Laan
Simplicial Approximation of Solutions to the Non-Linear Complementarity Problem

119 Pieter Boot
The East German Planning System Reconsidered

120 J.P.C. Kleijnen
Regression Metamodel Summarization of Model Behaviour

121 P.G.H. Mulder en A.L. Hempenius
Evaluating life time in a competing risks model for a chronic desease.

122 A.J.J. Talman en G. van der Laan
From Fixed Point to Equilibrium

123 J.P.C. Kleijnen
Design of simulation experiments.

124 H.L. Theuns en A.M.L. Passier-Grootjans
Internationaal toerisme; een gids in de algemene basisliteratuur en het bronnenmateriaal.

125 J.H.F. Schilderinck
Part I: Imports and Exports, Sub-divided by Countries Aggregated According the Branches of the European Community Interregional Input-Outputtables 1959, 1965, 1970 and 1975.