P20

INT. J. SYSTEMS SCI., 1985, VOL. 16, NO. 4, 439–446

Some modifications on the refined instrumental variable method†

WANG XIAN-LAI‡ and M. GEVERS§

Instrumental variable methods are known to produce fairly poor estimates of the noise-model parameters. This paper presents some modifications to Young's refined IV–AML method, which improve the quality of the noise-model estimates. The performance of this modified algorithm is evaluated by Monte Carlo simulations with medium sample sizes.

1. Introduction

The instrumental variable (IV) approach to parameter estimation in time series models has been proposed and developed by many authors (see, for example, Kendall and Stuart 1961, Johnston 1963, Wong and Polak 1967, Rowe 1970, Young 1970, 1976, Young and Jakeman 1979, 1980, Söderström and Stoica 1981, 1983, Stoica and Söderström 1983). In the engineering world, the IV method has been extensively experimented with by Young, and its convergence properties have been established in Söderström *et al.* (1978) and in Söderström and Stoica (1981) for different choices of the instrumental variables.

When it is desired to identify not only an input–output model (parametrized by a vector θ), but also a noise model (parametrized by a vector η), the IV algorithm must be followed by a method for the estimation of the noise-model parameters. This can be done in several ways, using for example, a prediction error method or an extended least-squares method to model the estimated residuals (namely, the equation errors). One of the better known methods for the estimation of both the input–output model parameters θ and the noise-model parameters η is the 'refined' instrumental variable approximate maximum likelihood (IV–AML) method proposed by Young (1976) and further studied in Young and Jakeman (1979, 1980), where extensive simulation studies are reported. The word 'refined' refers to a prefiltering of the data. In Stoica and Söderström (1983), the convergence properties of the IV–AML method and of other related methods are presented for different model sets. In this paper, Stoica and Söderström have derived an optimality criterion for the selection of the prefilters and of the instruments. The prefilters of the refined IV–AML method satisfy this criterion for the particular model set used by Young; however this criterion can also be applied to other model sets. This led Stoica and Söderström to suggest an optimal multistep procedure for the joint estimation of θ and η.

In all cases, the estimation of η is based on a modelling of the equation errors ξ_k (see § 2); these will of course be a function of the parameters θ of the input–output model, which must be estimated first. The estimation methods for η, and the

---

Received 2 May 1984.

† This work was undertaken while both authors were at the Department of Electrical Engineering, Louvain University, B-1348 Louvain la Neuve, Belgium.

‡ Department of Electrical Engineering and Automation, Tianjin University, Tianjin, China.

§ Department of Systems Engineering, Institute of Advanced Studies, Australian National University, Canberra, ACT 2601, Australia.

convergence analysis for $\hat{\eta}$ presented in Stoica and Söderström, rely on the assumption that the estimates $\hat{\theta}$ are close to the true θ, and that the estimated residuals $\hat{\xi}_k$ are therefore good (that is, consistent) estimates of the noise ξ_k. For small and medium sample sizes, these assumptions are of course not valid, and in such case the IV-based methods are known to produce poor estimates of the noise-model parameters. One reason for this is that the residuals ξ_k are estimated with error, which produces a bias in the estimation of the parameters η.

In this paper, two modifications to the refined IV–AML method are suggested. These modifications have resulted in a significant improvement of the noise-model parameter estimates where the noise can be modelled by an AR model. The refined IV–AML method was modified by introducing an additional instrumental variable, and by compensating for bias in the estimated noise model. These modifications are based on heuristic arguments, and are not substantiated by any theoretical analysis. However, all the simulations show that they result in a significant improvement of the noise-model parameter estimates and at the same time provide estimates of the input–output model that are at least as good as those obtained with the refined IV–AML method. Two of these Monte Carlo simulations will be presented. In addition, the choice of instruments in our modified IV–AML method satisfies Stoica and Söderström's optimality criterion, even with the additional IV that we have introduced.

The paper is organized as follows. In § 2 the basic IV method and the refined IV–AML procedure of Young (1976) are introduced. Suggested modifications are then presented in § 3, together with a heuristic justification. Some simulation results are presented in § 4.

2. The instrumental variable method

A brief description of the IV–AML method in its general form follows, along with the more specialized refined method of Young. Consider the discrete-time stochastic system described by

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_k + \xi_k \tag{1 a}$$

$$\xi_k = \frac{D(z^{-1})}{C(z^{-1})} e_k \tag{1 b}$$

where the subscript k denotes the sampling time; y_k and u_k are the output and input signal samples, respectively; z^{-1} is the backward shift operator (that is, $z^{-1}u_k = u_{k-1}$). The polynomials in eqns. (1) are

$$\left.\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + \ldots + a_{na} z^{-na} \\ B(z^{-1}) &= b_0 + b_1 z^{-1} + \ldots + b_{nb} z^{-nb} \\ C(z^{-1}) &= 1 + c_1 z^{-1} + \ldots + c_{nc} z^{-nc} \\ D(z^{-1}) &= 1 + d_1 z^{-1} + \ldots + d_{nd} z^{-nd} \end{aligned}\right\} \tag{2}$$

e_k is white noise with the following properties

$$E\{e_k\} = 0; \quad E\{e_j e_k\} = \sigma^2 \delta_{jk}; \quad E\{e_j u_k\} = 0 \quad \text{for all } j, k$$

where δ_{jk} is the Kronecker delta function. It is further assumed that the system is asymptotically stable, and that the polynomials $A(z)$ and $B(z)$, resp. $C(z)$ and $D(z)$, are coprime.

The various non-recursive IV methods can now be described by the following general equations:

$$\hat{\theta} = [\Sigma \hat{X}_k Z_k^{\mathrm{T}}]^{-1} \Sigma \hat{X}_k y_k \tag{3}$$

where $\hat{\theta}$ is the vector of estimates of the parameters in $A(z^{-1})$ and $B(z^{-1})$:

$$\hat{\theta} = [\hat{a}_1 \ldots \hat{a}_{\hat{n}a} \quad \hat{b}_0 \ldots \hat{b}_{\hat{n}b}]^{\mathrm{T}} \tag{4}$$

$\hat{n}a$ and $\hat{n}b$ are the estimates of na and nb,

$$Z_k^{\mathrm{T}} = [-y_{k-1} \ldots -y_{k-\hat{n}a} u_k \ldots u_{k-\hat{n}b}] \tag{5}$$

while $\hat{X}_k$ is the vector of instrumental variables. Different IV methods are obtained by different choices of instrumental variables (Söderström and Stoica 1981), for example, filtered inputs or outputs, delayed inputs or outputs, or other combinations.

The noise-model parameters can be estimated in an analogous manner by approximate maximum likelihood methods:

$$\hat{\eta} = [\Sigma \hat{W}_k V_k^{\mathrm{T}}]^{-1} \Sigma \hat{W}_k \hat{\xi}_k \tag{6}$$

where $\hat{\eta}$ is the vector of estimates in the parameters of $C(z^{-1})$ and $D(z^{-1})$:

$$\hat{\eta} = [\hat{c}_1 \ldots \hat{c}_{\hat{n}c} \quad \hat{d}_1 \ldots \hat{d}_{\hat{n}d}] \tag{7}$$

$\hat{n}c$ and $\hat{n}d$ are the estimates of nc and nd, and

$$V_k^{\mathrm{T}} = [-\hat{\xi}_{k-1} \ldots -\hat{\xi}_{k-\hat{n}c} \quad \hat{e}_{k-1} \ldots \hat{e}_{k-\hat{n}d}] \tag{8}$$

Here $\hat{\xi}_k$ and $\hat{e}_k$ are estimates of ξ_k and e_k. Finally, $\hat{W}_k$ is a vector of instruments; we refer to the refined IV–AML method below for a particular choice of $\hat{W}_k$.

The convergence properties of the basic recursive IV algorithm have been studied in Söderström *et al.* (1978). In Söderström and Stoica (1981), several choices of the instrumental variables have been examined, for the special case of an ARMAX model structure, such as $C(z^{-1}) = A(z^{-1})$ in eqn. (1). Finally, Stoica and Söderström (1983) have derived an optimality criterion for the selection of the instruments; it should be noted that this criterion does not lead to a unique choice of optimal instruments.

The estimation of $\hat{\eta}$ can be interweaved with the estimation of $\hat{\theta}$. This is done, for example, in the multistep procedure of Stoica and Söderström (1983), or in the IV–AML algorithm of Young (1976), which we describe below. The convergence properties of $\hat{\eta}$ will then depend on the convergence properties of $\hat{\theta}$. From the convergence analysis of Stoica and Söderström (1983), the following conclusions may be deduced about the convergence of the noise-model parameters $\hat{\eta}$.

2.1. *Convergence of* $\hat{\eta}$

Consider the system (1) with the following assumptions:

(i) The process operates in open loop, and certain conditions on the polynomials $A(z^{-1})$, $B(z^{-1})$, $C(z^{-1})$, $D(z^{-1})$, on the choice of inputs and of the basic instrumental variables $\hat{X}_k$ are fulfilled.

(ii) $\hat{\theta}_N \to \theta^*$ as $N \to \infty$, where θ^* is the true θ, and N is the number of samples.

(iii) $\hat{X}_k$ is asymptotically uncorrelated with $\hat{W}_k$.

Then $\hat{\xi}_k$ is a consistent estimate of the noise signal ξ_k, and the estimate $\hat{\eta}_N$ converges to the true η^* as $N \to \infty$.

The refined IV–AML method of Young (1976), which, for the particular choice of model (1), is identical to the prediction error method of Ljung and Söderström (1983) and the optimal IV method of Stoica and Söderström (1983), is obtained by replacing y_k, Z_k and $\hat{X}_k$ in (3) by y_k^*, Z_k^* and $\hat{X}_k^*$, respectively, where

$$\hat{X}_k^* = [-\hat{x}_{k-1}^*, \ldots, -\hat{x}_{k-\hat{n}a}^*, \quad u_k^*, \ldots, u_{k-\hat{n}b}^*]^{\mathrm{T}} \tag{9 a}$$

$$Z_k^* = [-y_{k-1}^*, \ldots, -y_{k-\hat{n}a}^*, \quad u_k^*, \ldots, u_{k-\hat{n}b}^*]^{\mathrm{T}} \tag{9 b}$$

$$\hat{x}_k = \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})} u_k \tag{10}$$

and

$$\hat{x}_k^* = \hat{F}(z^{-1})\hat{x}_k, \quad y_k^* = \hat{F}(z^{-1})y_k, \quad u_k^* = \hat{F}(z^{-1})u_k \tag{11 a}$$

with

$$\hat{F}(z^{-1}) = \frac{\hat{C}(z^{-1})}{\hat{A}(z^{-1})\hat{D}(z^{-1})} \tag{11 b}$$

For the computation of $\hat{\eta}$ by (6), V_k is defined by (8), while $\hat{\xi}_k$, $\hat{e}_k$ and $\hat{W}_k$ are recursively computed as follows:

$$\hat{\xi}_k = y_k - \hat{x}_k \tag{12}$$

$$\hat{e}_k = \hat{\xi}_k - V_k^{\mathrm{T}}\hat{\eta}_k \tag{13}$$

$$\hat{W}_k = [-\hat{\xi}_{k-1}^*, \ldots, -\hat{\xi}_{k-\hat{n}c}^*, \hat{e}_{k-1}^*, \ldots, \hat{e}_{k-\hat{n}d}^*]^{\mathrm{T}} \tag{14}$$

where

$$\hat{\xi}_k^* = \frac{1}{\hat{D}(z^{-1})}\hat{\xi}_k, \quad \hat{e}_k^* = \frac{1}{\hat{D}(z^{-1})}\hat{e}_k \tag{15}$$

The vector $\hat{\eta}_k$ in (13) is the estimate of η at the kth sampling instant. Notice that all the variables defined in eqns. (12) to (15) depend upon the estimate $\hat{\theta}$ obtained by the IV-algorithm via the variable $\hat{x}_k$ in (12). The two algorithms (IV and AML) are therefore clearly interweaved (for further details, see Young 1976 or Young and Jakeman 1979, 1980).

The analysis and the Monte Carlo simulations of Young and Jakeman (1979, 1980) and Stoica and Söderström (1983) show that the introduction of the filters $\hat{C}(z^{-1})/\hat{A}(z^{-1})\hat{D}(z^{-1})$ and $1/\hat{D}(z^{-1})$ improves the quality of the estimates, compared to the basic IV methods, where no filtering is performed. However, for low and moderate sample size, the noise model estimates c_i and d_i are often poor, because the convergence conditions for $\hat{\eta}$ are not satisfied (for instance, the estimate $\hat{\xi}_k$ is biased, $\hat{X}_k$ is correlated with $\hat{W}_k$, and so on). From the simulation results of Young and Jakeman, it can be seen that the noise-parameter estimates are of good quality only when $C(z^{-1}) = A(z^{-1})$, that is, in the case of an ARMAX model.

Thus, if the objective is to estimate the system parameters a_i and b_i (as is often the case), then the refined IV method can be considered good enough. For the case where the noise dynamics are also required, we suggest some modifications to the refined IV–AML method, which have proved to improve the quality of the noise-model estimates. These modifications are described in the next section.

3. Some modifications to the refined IV–AML method

Some modifications may be introduced that aim to reduce the correlation between $\hat{X}_k$ and $\hat{W}_k$, and eliminate the bias of $\hat{\xi}_k$, which has been observed for small and moderate sample sizes. These modifications consist of introducing an additional instrumental variable and compensating for bias in the noise model. In addition, a weighting factor should be used in the algorithm, but this is by now standard practice. In order to apply these modifications, we adopt a purely autoregressive (AR) model for ξ_k. This is of course a restrictive assumption; however most noise processes can be modelled by long AR models, even if this is at the expense of parsimony. The advantage with AR models is that the parameters can be estimated by least-squares methods. The reason for using an AR model is that we can then eliminate the bias on $\hat{\eta}$ caused by the errors $\tilde{\theta}$ (see § 3.3).

3.1. *Model*

We suggest that the following model be used for the representation of the process

$$y_k = \frac{B(z^{-1})}{A(z^{-1})} u_k + \xi_k \tag{16 a}$$

$$\xi_k = \frac{1}{C(z^{-1})} e_k \tag{16 b}$$

The polynomials $A(z^{-1})$, $B(z^{-1})$ and $C(z^{-1})$ have the same form as in eqn. (2).

3.2. *Additional instrumental variable*

The estimate $\hat{\theta}$ is again given, in non-recursive form, by

$$\hat{\theta} = [\Sigma \hat{X}_k^* Z_k^{*\mathrm{T}}]^{-1} \Sigma \hat{X}_k^* y_k^* \tag{17}$$

where $\hat{X}_k^*$, Z_k^* and y_k^* are defined by eqns. (9) to (11), but with $\hat{D}(z^{-1}) = 1$ in (11 *b*). The estimate $\hat{\eta} = [\hat{c}_1 \dots \hat{c}_{\hat{n}c}]^{\mathrm{T}}$ of the noise model is given, as before, by

$$\hat{\eta} = [\Sigma \hat{W}_k V_k^{\mathrm{T}}]^{-1} \Sigma \hat{W}_k \hat{\xi}_k \tag{18}$$

where $\hat{\xi}_k$ is defined as in (12) and V_k^{T} is defined as before:

$$V_k^{\mathrm{T}} = [-\hat{\xi}_{k-1} \dots -\hat{\xi}_{k-\hat{n}c}] \tag{19}$$

But now $\hat{W}_k$ is defined as follows:

$$\hat{W}_k = [-\hat{w}_{k-1} \dots -\hat{w}_{k-\hat{n}c}]^{\mathrm{T}} \tag{20}$$

where $\hat{w}_k$ is an additional instrumental variable generated by

$$\hat{w}_k = \frac{1}{\hat{C}(z^{-1})} \hat{\xi}_k \tag{21}$$

The idea of the filter $1/\hat{C}(z^{-1})$ is to filter out unwanted frequency components in the error between $\hat{\xi}_k$ and ξ_k; these components are generated by the procedure of estimating ξ_k. It should be noted that this choice for $\hat{W}_k$ also satisfies the optimality criterion of Stoica and Söderström (1983).

In the recursive version, the estimation of $\hat{\theta}$ and $\hat{\eta}$ is performed by the following recursions (see Young and Jakeman (1979))

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \hat{P}_{k-1}\hat{X}_k^*[\hat{\sigma}^2 + Z_k^{*T}\hat{P}_{k-1}\hat{X}_k^*]^{-1}(Z_k^{*T}\hat{\theta}_{k-1} - y_k^*) \tag{22 a}$$

$$\hat{P}_k = \hat{P}_{k-1} - \hat{P}_{k-1}\hat{X}_k^*[\hat{\sigma}^2 + Z_k^{*T}\hat{P}_{k-1}\hat{X}_k^*]^{-1}Z_k^{*T}\hat{P}_{k-1} \tag{22 b}$$

$$\hat{\eta}_k = \hat{\eta}_{k-1} - P_{k-1}\hat{W}_k[\hat{\sigma}^2 + V_k^T P_{k-1}\hat{W}_k]^{-1}(V_k^T\hat{\eta}_{k-1} - \hat{\xi}_k) \tag{23 a}$$

$$P_k = P_{k-1} - P_{k-1}\hat{W}_k[\hat{\sigma}^2 + V_k^T P_{k-1}\hat{W}_k]^{-1}V_k^T P_{k-1} \tag{23 b}$$

where $\hat{\sigma}^2$ is an estimate of the variance σ^2 of e_k.

3.3. *Compensation of bias for the AR noise model parameter estimates*

The AR model (16 *b*) has been assumed for ξ_k. However, ξ_k is not directly measured; only the estimate $\hat{\xi}_k$ is available, which is computed by (12) and is of course a function of $\hat{\theta}$. For simplicity of analysis, we shall assume that

$$\hat{\xi}_k = \xi_k + \varepsilon_k \tag{24}$$

with

$$E\{\varepsilon_k\} = 0, \quad E\{\varepsilon_j \varepsilon_k\} = \gamma^2 \delta_{jk}, \quad E\{e_j \varepsilon_k\} = 0 \quad \forall j, k \tag{25}$$

The AR model (16 *b*) is then replaced by

$$C(z^{-1})\hat{\xi}_k = C(z^{-1})\varepsilon_k + e_k \tag{26}$$

As a consequence, the estimation of the AR parameter vector $\eta = (c_1, \ldots, c_{nc})$ based on the noise corrupted data $\hat{\xi}_1, \ldots, \hat{\xi}_N$ leads to biased estimates. Sakai and Arase (1979) have described a recursive modified least-squares method for the estimation of model (26), which compensates for this bias. We have adopted this method in our modified IV–AML procedure. Following Sakai and Arase, the estimate $\tilde{\eta}$ of η is computed recursively as follows

$$\tilde{\eta}_k = \hat{\eta}_k + kP_k\hat{\gamma}_k^2\tilde{\eta}_{k-1} \tag{27}$$

where $\hat{\eta}_k$ and P_k are defined by (23), and where $\hat{\gamma}_k^2$ is a recursive estimate of the variance γ^2 of ε_k. This estimate can be obtained by

$$\hat{\gamma}_k^2 = \frac{k^{-1}R_k - \hat{\sigma}^2}{1 + \tilde{\eta}_{k-1}\hat{\eta}_k} \tag{28}$$

where

$$R_k = R_{k-1} + (1 + V_{k-1}^T P_{k-1} V_{k-1})^{-1}(\hat{\xi}_k - V_k^T\hat{\eta}_{k-1})^2 \tag{29}$$

The quantities $\hat{\sigma}^2$, $\hat{\xi}_k$ and V_k have been defined before. The derivation of the formulas (27) to (29) and the asymptotic properties of the estimate $\tilde{\eta}_k$ can be found in Sakai and Arase (1979).

3.4. *Time-varying weighting factor*

In our modified IV–AML algorithm, we always use a time-varying weighting factor λ_k in the computation of P_k, as suggested, for example, in Söderström *et al.* (1978), in order to increase the convergence rate of the algorithm. This weighting factor is generated by

$$\lambda_k = \lambda^0 \lambda_{k-1} + (1 - \lambda^0) \tag{30}$$

where λ^0 and the initial value λ_0 of λ_k are chosen close to 1 for example, $\lambda^0 = 0{\cdot}99$, $\lambda_0 = 0{\cdot}95$). The recursive equation (23 *b*) for P_k is then replaced by

$$P_k = [P_{k-1} - P_{k-1}\hat{W}_k(\hat{\sigma}^2\lambda(k) + V_k^{\mathrm{T}}P_{k-1}\hat{W}_k)^{-1}V_k^{\mathrm{T}}P_{k-1}]/\lambda(k) \tag{31}$$

4. Simulation experiments and discussion of results

Different models have been simulated and the parameters have been estimated using our recursive modified IV–AML method. The results were always better than with the original IV–AML method, in the sense that the noise-model parameters were always estimated with greater accuracy for low and moderate sample sizes, while the parameters a_i and b_i were at least as good as with the refined IV–AML method. We now show two simulation experiments for models with the same $A(z^{-1})$ and $B(z^{-1})$ polynomials, but with different noise models. Data were generated from the following two models:

Model 1

$$y_k = \frac{[0{\cdot}8z^{-1} 0{\cdot}2z^{-1}]}{1 - 1{\cdot}3z^{-1} + 0{\cdot}6z^{-2}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_k + \frac{1}{1 - 0{\cdot}8z^{-1} + 0{\cdot}4z^{-2}} e_k$$

Model 2

$$y_k = \frac{[0{\cdot}8z^{-1} 0{\cdot}2z^{-1}]}{1 - 1{\cdot}3z^{-1} + 0{\cdot}6z^{-2}} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_k + \frac{1}{1 - 0{\cdot}527z^{-1} + 0{\cdot}0695z^{-2}} e_k$$

where u_{1k} and u_{2k} are sequences of independent random binary signals, and e_k is gaussian white noise with zero mean.

Tables 1 and 2 show the results obtained with a sample size of $N = 500$, a signal-to-noise ratio $S/N = 9$ and a weighting factor generated by (30) with $\lambda^0 = 0{\cdot}99$ and $\lambda_0 = 0{\cdot}998$. The estimated values of the parameters are the average obtained from 10 Monte Carlo simulations, together with the experimental standard deviations. The estimated values and the experimental standard deviations obtained under the same conditions with the original recursive IV–AML method of Young are given for comparison purposes.

Tables 1 and 2 show that the noise-model parameter estimates obtained with our modified IV–AML method are significantly better than those obtained with the original version of the algorithm. The original IV–AML method does not provide the prefilter introduced by (21), nor does it compensate for the bias caused by the error on the estimation of ξ_k. As mentioned earlier, the deviation ε_k between ξ_k and $\hat{\xi}_k$ is a

Parameter	True values	Modified IV–AML estimated values	IV–AML estimated values
a_1	$-1{\cdot}3$	$-1{\cdot}3573 \pm 0{\cdot}0434$	$-1{\cdot}3710 \pm 0{\cdot}0299$
a_2	$0{\cdot}6$	$0{\cdot}6502 \pm 0{\cdot}0363$	$0{\cdot}6816 \pm 0{\cdot}0308$
b_{11}	$0{\cdot}8$	$0{\cdot}7760 \pm 0{\cdot}0102$	$0{\cdot}8061 \pm 0{\cdot}0246$
b_{21}	$0{\cdot}2$	$0{\cdot}1971 \pm 0{\cdot}0301$	$0{\cdot}2039 \pm 0{\cdot}0437$
c_1	$-0{\cdot}8$	$-0{\cdot}7547 \pm 0{\cdot}1346$	$-1{\cdot}4892 \pm 0{\cdot}0065$
c_2	$0{\cdot}4$	$0{\cdot}3926 \pm 0{\cdot}0760$	$0{\cdot}8374 \pm 0{\cdot}0053$

Table 1. Estimated values and experimental standard deviations for model 1.

Parameter	True values	Modified IV–AML estimated values	IV–AML estimated values
a_1	−1·3	−1·3382 ± 0·0429	−1·3626 ± 0·0307
a_2	0·6	0·6176 ± 0·0299	0·6414 ± 0·0224
b_{11}	0·8	0·7696 ± 0·0096	0·7879 ± 0·0073
b_{21}	0·2	0·1792 ± 0·0403	0·1809 ± 0·0466
c_1	−0·527	−0·4345 ± 0·1694	−1·4949 ± 0·0016
c_2	0·0695	0·1148 ± 0·0797	0·7614 ± 0·0042

Table 2. Estimated values and experimental standard deviations for model 2.

complex function of $\hat{x}_k$ and therefore causes a correlation between $\hat{X}_k$ and $\hat{W}_k$, for low and medium sample sizes. This violates one of the convergence conditions on $\hat{\eta}$ (see § 2). The modifications 3.2 and 3.3 are one way of alleviating this problem. For large sample sizes, $\hat{\theta}$ is closer to the true θ and these modifications become less necessary.

5. Conclusions

Some modifications to the basic IV–AML algorithm have been proposed which have been shown to be effective in improving the quality of the noise-model parameter estimates, particularly for low and medium sample sizes. These modifications are based on heuristic arguments. The authors do not claim to have come up with an earthshaking new algorithm, and it is certainly difficult to draw general conclusions from simulation results. However, it appears that the suggested modifications are worth considering, since they have given systematically better results than the basic IV–AML in our simulation experiments.

References

Johnston, J., 1963, *Econometric Methods* (New York: McGraw Hill).

Kendall, M. G., and Stuart, A., 1961, *Advanced Theory of Statistics* (London: Griffin).

Ljung, L., 1976, On the consistency of prediction error identification methods. In R. K. Mehra and D. G. Lainoitis, *System Identification—Advances and Case Studies* (New York: Academic Press).

Ljung, L., and Söderström, T., 1983, *Theory and Practice of Recursive Identification* (Massachusetts: MIT Press).

Rowe, J. H., 1970, *Int. J. Control*, **12,** 721.

Sakai, H., and Arase, M., 1979, *Int. J. Control*, **30,** 949.

Söderström, T., Ljung, L., and Gustavsson, I., 1978, *Automatica*, **14,** 231.

Söderström, T., and Stoica, P., 1981, *Automatica*, **17,** 101; 1983, *Instrumental Variable Methods for System Identification* (Berlin: Springer Verlag).

Stoica, P., and Söderström, T., 1983, *I.E.E.E. Trans. autom. Control*, **28, 7,** 757.

Wong, K. Y., and Polak, E., 1967, *I.E.E.E. Trans. autom. Control*, **12,** 707.

Young, P. C., 1970, *Automatica*, **6,** 271; 1976, *Int. J. Control*, **23,** 593.

Young, P. C., and Jakeman, A., 1979, 1980, *Int. J. Control*, **29,** 1; **29,** 621; **31,** 741.