Identification of dynamic errors-in-variables systems with periodic data

Torsten Söderström, Mei Hong

Department of Information Technology
Uppsala University
Box 337, SE-751 05 Uppsala, Sweden

Technical report 2004-037
August 2004
ISSN 1404-3203

Identification of dynamic errors-in-variables systems with periodic data

Torsten Söderström, Mei Hong

Division of Systems and Control, Department of Information Technology,
Uppsala University, P O Box 337, SE-75105 Uppsala, Sweden
email: ts@it.uu.se mhong@it.uu.se

version August 31, 2004

Abstract

Using instrumental variable methods to estimate the parameters of dynamic errors-in-variables systems with a periodic input signal is the focus in this report. How to choose suitable instrumental variable vectors is the key point. Four variants are proposed; all of them can generate consistent estimates. An analysis shows that the best accuracy is achieved by using a specific overdetermined instrumental variable vector. Numerical illustrations demonstrate the effectiveness of the proposed IV3 method for both white and colored measurement noise. It is superior to alternative methods under low signal to noise ratios.

1 Introduction

The problem of identification of dynamic errors-in-variables (EIV) systems appear in very wide scientific areas, such as time series modelling, array signal processing for direction-of-arrival estimation, blind channel equalization, multivariate calibration in analytical chemistry, image processing, astronomical data reduction, etc. A number of methods were elaborated in the past few decades. For example, attempts using instrumental variable estimates [9] or the Frisch scheme [1] and other bias-compensating of the least-squares estimates were developed. Other methods are based on the frequency domain [4], or use higher-order statistics [11]. It is also possible to apply the maximum likelihood and the prediction error approaches [6]. Some comparison between different approaches are given in [8].

Since periodic excitation signals offer interesting advantages not only in frequency domain but also in time domain identification, several methods were proposed to deal EIV problem under this experiment condition [5], [3]. In this report a time domain approach, the instrumental variable (IV) method, will be applied to solve the EIV problem with a periodic input signal, because IV estimators are computationally inexpensive and are applicable under fairly general noise conditions. The main motivation is to analyze what type of instrumental variables should be chosen to maximally utilize the periodic information of the measurement data. Another contribution of this work is that we analyze the optimal weighting matrix and the relevant lower bound of the estimation accuracy. Formulas for the estimate parameter covariance matrices for different IV variants show that using a certain overdetermined instrumental variable vector can reduce the noise contribution. It is shown that for this particular case a simple least squares solution given the same accuracy as using an optimal weighting matrix. This attractive property simplifies the identification process greatly.

The outline of this report is as follows: Section 2 gives some assumptions and preliminary issues regarding errors-in-variables models. Three instrumental variable variants are also proposed in this section. Consistency and accuracy of these IV variants are analyzed in Sections 3 and 4. At the end of Section 4, a forth instrumental variable variant is also mentioned and analyzed. In Section 5, simulations are illustrated in three aspects, numerical result of three IV variants, comparison with Basic IV method and the comparison with CLS and CTLS algorithms [3]. Finally, some conclusions are drawn in Section 6.

2 Preliminaries

Consider the system depicted in Figure 1.

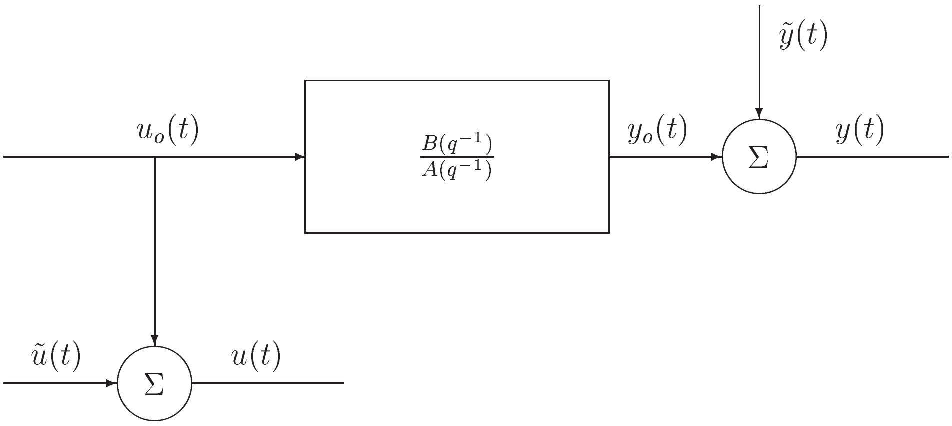


Figure 1: The basic setup for a dynamic errors-in-variables problem.

The system is given by the difference equation

$$A(q^{-1})\, y_o(t) = B(q^{-1})\, u_o(t) \tag{2.1}$$

where

$$\begin{aligned} A(q^{-1}) &= 1 + a_1\, q^{-1} + \ldots + a_n\, q^{-n} \\ B(q^{-1}) &= b_1\, q^{-1} + \ldots + b_n\, q^{-n} \end{aligned} \tag{2.2}$$

are polynomial in the backward shift operator q^{-1}, *i.e.* $q^{-1}\, x(t) = x(t-1)$, and $y_o(t)$, $u_o(t)$ denote the noise free output and input, respectively.

It is generally assumed that the system is asymptotically stable and that the model order n is known. The model can easily be extended by allowing different degrees of A and B and by introducing a further delay. To keep the treatment reasonably simple these extensions are not dealt with here. The input and the output are not directly available, but can be measured with some noise: $\tilde{u}(t)$ and $\tilde{y}(t)$, respectively. It is the noise corrupted measurements $u(t)$ and $y(t)$ that are available.

We next introduce some general assumptions.

A1. The noise free signal $u_o(t)$ is a periodic function. The length of the period is denoted N. It is assumed that m periods of the data $u(t), y(t)$ are available. In each period $u_o(t)$ is a stationary process.

A2. The measurement noise signals $\tilde{u}(t)$ and $\tilde{y}(t)$ are uncorrelated with the noise free input $u_o(s)$ for all t and s. Further, the measurement noise signals within different periods are uncorrelated.

For simplicity we will generally assume that all signals have zero mean values.

Next introduce the regressor vector $\varphi(t)$ as

$$\varphi(t) \quad = \quad [-y(t-1) \ldots - y(t-n) \; u(t-1) \ldots u(t-n)]^T \tag{2.3}$$

We can then form a linear regressive model as

$$y(t) = \varphi^T(t)\theta + \varepsilon(t) \tag{2.4}$$

where the parameter vector θ is

$$\theta = [a_1 \ldots a_n \; b_1 \ldots b_n]^T \tag{2.5}$$

Due to the presence of noise in both y and u, a linear least squares estimate of θ will be biased and not consistent. Instead we will consider instrumental variable estimates, which are obtained by correlating the model with a vector $z(t)$ of instruments:

$$\sum_t z(t)y(t) \cong \left[\sum_t z(t)\varphi^T(t)\right]\theta \tag{2.6}$$

The vector $z(t)$ must have at least as many elements as θ. When $z(t)$ is larger, we solve (2.6) in a weighted least square sense to get

$$\hat{\theta} \quad = \quad \arg\min_{\theta} || \left[\sum_t z(t)\varphi^T(t)\right]\theta - \left[\sum_t z(t)y(t)\right] ||_Q^2 \tag{2.7}$$

leading to

$$\hat{\theta} = (\hat{R}^T Q \hat{R})^{-1} \hat{R}^T Q \hat{r} \tag{2.8}$$

with

$$\hat{R} = \frac{1}{Nm}\sum_t z(t)\varphi^T(t) \quad \hat{r} = \frac{1}{Nm}\sum_t z(t)y(t) \tag{2.9}$$

In (2.7), (2.8) Q is a positive definite weighting matrix. The summations in (2.6)-(2.9) are over all time points (in the m periods), that is, t goes from 1 to Nm.

Next we introduce some more detailed notation, where we exploit the periodicity of the noise free data $u_o(t)$, $y_o(t)$. In period j (where $1 \leq j \leq m$), we write the regressor vector as

$$\varphi_j(t) = \varphi^0(t) + \tilde{\varphi}_j(t) \qquad t = 1, \ldots, N \tag{2.10}$$

where $\varphi^0(t)$ contains the noise free data and $\tilde{\varphi}_j(t)$ denotes the noise contributions. Similarly, $z_j(t)$ (for $1 \leq j \leq m$) will denote the instrumental vector for period j.

We proceed to give some variants of instrumental variable vectors.

Example 1. It is natural to let the vector $z_j(t)$ be formed by regressors other than $\varphi_j(t)$. A simple choice is to take

$$z_j(t) = \begin{cases} \varphi_{j+1}(t) & j = 1, \ldots, m-1 \\ 0 & j = m \end{cases} \tag{2.11}$$

■

Example 2. A slight extension of Example 1 would be to take

$$z_j(t) = \begin{cases} \varphi_{j+1}(t) & j = 1, \ldots, m-1 \\ \varphi_1(t) & j = m \end{cases} \tag{2.12}$$

■

In both the above example z_j and φ_j will have the same dimensions. Then $\hat{R}$ in (2.8), (2.9) will be a square matrix and the weighting Q becomes superfluous.

A somewhat more elaborated example is the following, where we use all available regressors as instruments.

Example 3. Consider an overdetermined instrumental variable vector as

$$\begin{aligned} z_1(t) &= \begin{pmatrix} \varphi_2(t) \\ \vdots \\ \vdots \\ \varphi_m(t) \end{pmatrix}, \qquad z_2(t) = \begin{pmatrix} \varphi_3(t) \\ \vdots \\ \vdots \\ \varphi_m(t) \\ \varphi_1(t) \end{pmatrix}, \ldots, \\ z_j(t) &= \begin{pmatrix} \varphi_{j+1}(t) \\ \vdots \\ \varphi_m(t) \\ \varphi_1(t) \\ \vdots \\ \varphi_{j-1}(t) \end{pmatrix} \end{aligned} \tag{2.13}$$

■

For a general case we introduce the following assumption:

A3. The instrumental vector $z_j(t)$ is uncorrelated to the measurement noise in period j. Further, the covariance matrix of $z_j(t)$ has rank at least equal to dim(θ).

It will be convenient to further introduce the following matrices

$$\begin{aligned} \phi(t) &= [\varphi_1(t)\ldots\varphi_m(t)] \\ Z(t) &= [z_1(t)\ldots z_m(t)] \end{aligned} \qquad t = 1,\ldots,N \tag{2.14}$$

and the vector

$$Y(t) = \begin{pmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{pmatrix} \qquad t = 1,\ldots,N \tag{2.15}$$

where $y_j(t)$ is the output at time t within period j:

$$y_j(t) = y(t + (j-1)N). \tag{2.16}$$

The the basic equation (2.6) can be compactly written as

$$\left[\sum_{t=1}^{N} Z(t)\phi^T(t)\right]\theta \cong \left[\sum_{t=1}^{N} Z(t)Y(t)\right] \tag{2.17}$$

We write the estimate as, see (2.8)

$$\hat{\theta} = (\hat{R}^T Q \hat{R})^{-1}\hat{R}^T Q \hat{r} \tag{2.18}$$

where now

$$\hat{R} = \frac{1}{Nm}\sum_{t=1}^{N} Z(t)\phi^T(t) = \frac{1}{Nm}\sum_{j=1}^{m}\sum_{t=1}^{N} z_j(t)\varphi_j^T(t) \tag{2.19}$$

$$\hat{r} = \frac{1}{Nm}\sum_{t=1}^{N} Z(t)Y(t) = \frac{1}{Nm}\sum_{j=1}^{m}\sum_{t=1}^{N} z_j(t)y_j(t) \tag{2.20}$$

When analyzing the general estimate (2.18), it is worth noticing that the underlying model equation (2.17) take the form of a multivariable instrumental variance estimate [10, page 262]. This observation will be useful when examining the statistical properties of the estimate (2.18).

In the analysis, we generally assume that the model structure captures the true dynamic. More precisely, we assume that there is a true parameter vector θ_0 such that

$$\varphi(t) = \varphi^T(t)\theta_0 + v(t) \tag{2.21}$$

where

$$v(t) = A_0(q^{-1})\tilde{y}(t) - B_0(q^{-1})\tilde{u}(t). \tag{2.22}$$

Introduce the vector $V(t)$ similarly to $Y(t)$, (2.15). The relation (2.21) now simplifies to

$$Y(t) = \phi^T(t)\theta_0 + V(t) \tag{2.23}$$

Inserting (2.23) into (2.20), and combining this with (2.18) gives

$$\hat{\theta} - \theta_0 = (\hat{R}^T Q \hat{R})^{-1} \hat{R}^T Q \hat{\tilde{r}} \tag{2.24}$$

where

$$\hat{\tilde{r}} = \frac{1}{Nm} \sum_{t=1}^{N} Z(t)V(t) \tag{2.25}$$

3 Consistency Analysis

To analyze consistency (*i.e.* to examine if $\lim_{N\to\infty} \hat{\theta} = \theta_0$ holds), it follows from (2.24) and standard conditions, see [9], [10] that consistency is guaranteed if

$$\lim_{N\to\infty} \hat{R} = R \tag{3.1}$$

has full rank and

$$\lim_{N\to\infty} \hat{\tilde{r}} = 0 \tag{3.2}$$

To examine these conditions further, note that the left hand side of (3.2) can be written

$$\begin{aligned} \lim_{N\to\infty} \hat{\tilde{r}} &= \lim_{N\to\infty} \frac{1}{Nm} \sum_{j=1}^{m} \sum_{t=1}^{N} z_j(t) v_j(t) \\ &= \frac{1}{m} \sum_{j=1}^{m} E z_j(t) v_j(t) \end{aligned} \tag{3.3}$$

Hence, (3.2) follows as the instrumental variables are constructed such that $z_j(t)$ is uncorrelated with $v_j(t)$, see Assumption A3. Under Assumption A2 this holds for all the variants of choosing $Z(t)$ which were considered above.

To examine the condition (3.1) introduce the covariance matrix R_0 of the noise-free regressor vector $\varphi^0(t)$:

$$R_0 = \frac{1}{N} \sum_{t=1}^{N} \varphi^0(t) \varphi^0(t)^T \tag{3.4}$$

where R_0 is assumed to be positive definite. Note that this assumption is basically a condition on persistent excitation of the noise-free input signal $u_0(t)$, see [10].

We find from (2.19) that

$$\begin{aligned} R &= \lim_{N\to\infty} \frac{1}{Nm} \sum_{j=1}^{m} \sum_{t=1}^{N} z_j(t) [\varphi^{0^T}(t) + \tilde{\varphi}_j^T(t)] \\ &= \frac{1}{m} \sum_{j=1}^{m} \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} z_j(t) \varphi^{0^T}(t) \end{aligned} \tag{3.5}$$

as $z_j(t)$ is uncorrelated with the noise during period j. (see Assumption A3).

We next evaluate the limit matrix R in (3.5), for the instrumental vectors introduced in Examples 1-3. For Example 1 we get

$$\begin{aligned} R &= \frac{1}{m}\sum_{j=1}^{m-1}\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\varphi_{j+1}(t)\varphi^{0^T}(t) \\ &= \frac{1}{m}\sum_{j=1}^{m-1}\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}[\varphi^0(t)+\tilde{\varphi}_{j+1}(t)]\varphi^{0^T}(t) \\ &= \frac{m-1}{m}R_0 \end{aligned} \tag{3.6}$$

which is nonsingular and of full rank since R_0 is positive definite. In the same way it is established that for Example 2

$$R = R_0 \tag{3.7}$$

holds. Finally, for Example 3 we get

$$\begin{aligned} R &= \frac{1}{m}\sum_{j=1}^{m}\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N}\{\begin{pmatrix}\varphi^0(t)\\ \vdots \\ \varphi^0(t)\end{pmatrix}+\begin{pmatrix}\tilde{\varphi}_{j+1}(t)\\ \vdots \\ \tilde{\varphi}_m(t)\\ \tilde{\varphi}_1(t)\\ \vdots \\ \tilde{\varphi}_{j-1}(t)\end{pmatrix}\}\times\varphi^{0^T}(t) \\ &= \begin{pmatrix}R_0\\ \vdots \\ R_0\end{pmatrix} = e_{m-1}\otimes R_0 \end{aligned} \tag{3.8}$$

where $e_{m-1} = (1\dots1)^T$ has dimension $(m-1)\times 1$, and $\otimes$ denotes Kronecker product. Apparently, the matrix R is of full rank in (3.8).

4 Accuracy Analysis

It follows from the general theory of instrumental variable estimation for multi-variable system [9], [10], that the estimation error is asymptotically Gaussian distributed as

$$\sqrt{mN}(\hat{\theta}-\theta_0) \xrightarrow{dist} \mathcal{N}(0,P) \tag{4.1}$$

where

$$P = P(Q) = (R^TQR)^{-1}R^TQSQR(R^TQR)^{-1} \tag{4.2}$$

and

$$S = E\left[\sum_{j=0}^{\infty}Z(t+j)H_j\right]\Lambda\left[\sum_{k=0}^{\infty}H_k^TZ^T(t+k)\right] \tag{4.3}$$

In (4.3) $\{H_j\}_{j=0}^{\infty}$ and Λ are defined by a spectral factorization:

$$\Phi_V(w) = \mathbb{H}(e^{iw})\Lambda\mathbb{H}^*(e^{iw}) \tag{4.4}$$

along with the condition that $H_0 = I$, $\mathbb{H}(q^{-1}) = \sum_{j=0}^{\infty} H_j q^{-j}$ and $\mathbb{H}^{-1}(q^{-1})$ being asymptotically stable. In (4.4), $\Phi_V(w)$ denotes the spectral density matrix of the vector $V(t)$, see (2.23). Note that all quantities in (4.4) as well as $\{H_j\}$ are $m \times m$ matrices.

Due to Assumption A2, the measurement noise sequences in different periods are uncorrelated. Hence the components $v_j(t)$ of $V(t)$ are uncorrelated. Therefore the spectral density matrix $\Phi_V(w)$ is diagonal, and in fact, its diagonal elements are equal. We write this as

$$\Phi_V(w) = \phi_v(w)I \tag{4.5}$$

It follows that the spectral factorization on (4.4) can be substituted by a scalar spectral factorization:

$$\begin{aligned} \phi_v(w) &= H(e^{iw})\lambda^2 H(e^{-iw}) && (4.6) \\ H(q^{-1}) &= \sum_{k=0}^{\infty} h_k q^{-k}, \qquad h_0 = 1 && (4.7) \end{aligned}$$

It then follows that

$$\begin{aligned} H_k &= h_k I \\ \Lambda &= \lambda^2 I \end{aligned} \tag{4.8}$$

Therefore, the matrix S in (4.3) can be simplified:

$$\begin{aligned} S &= E\left[\sum_{j=0}^{\infty} h_j Z(t+j)\right] \lambda^2 I \left[\sum_{k=0}^{\infty} h_k Z^T(t+k)\right] \\ &= E\lambda^2 \sum_{j=0}^{\infty}\sum_{k=0}^{\infty} \left[h_k Z(t'-k) h_j Z^T(t'-j)\right] \\ &= \lambda^2 E\left[H(q^{-1})Z(t)\right]\left[H(q^{-1})Z(t)\right]^T \end{aligned} \tag{4.9}$$

4.1 Optimal weighting

The covariance matrix P in (4.2) apparently depends on the weighting matrix Q. In the examples treated so far, Q is irrelevant in Example 1 and 2, but does appear explicitly for Example 3. There is in fact an optimal choice of the weighting matrix. It is shown in [10] that

$$P \geq P^{opt} \tag{4.10}$$

meaning that the difference $P - P^{opt}$ is nonnegative definite where

$$P^{opt} = (R^T S^{-1} R)^{-1} \tag{4.11}$$

Further, equality holds in (4.10) if

$$Q = S^{-1} \tag{4.12}$$

We next examine how to compute the matrix S when the instruments are chosen as in Example 3. First the noise $V(t)$, see (2.22), has to be considered. Its spectral density $\phi_v(w)$ is factorized to give $H(q^{-1})$ and λ^2, see (4.6). This is a standard procedure, see for example [7]. Next we find from (2.14) and (4.9)

$$S_{IV3} = \lambda^2 E\left[H(q^{-1})\left[z_1(t)\dots z_m(t)\right]\right]\left[H(q^{-1})\left[z_1(t)\dots z_m(t)\right]\right]^T \quad (4.13)$$

Split the instrumental vector $z_j(t)$ into a noise-free part and a noise contribution as, see (2.13),

$$\begin{aligned} z_j(t) &= z_j^0(t) + \tilde{z}_j(t) \\ &= e_{m-1} \otimes \varphi^0(t) + \begin{pmatrix} \tilde{\varphi}_{j+1}(t) \\ \vdots \\ \tilde{\varphi}_m(t) \\ \tilde{\varphi}_1(t) \\ \vdots \\ \tilde{\varphi}_{j-1}(t) \end{pmatrix} \end{aligned} \quad (4.14)$$

This gives

$$\begin{aligned} S_{IV3} &= \lim_{N\to\infty} \frac{\lambda^2}{Nm} \sum_{j=1}^{m} \sum_{t=1}^{N} \left[H(q^{-1})z_j(t)\right]\left[H(q^{-1})z_j(t)\right]^T \\ &= \frac{\lambda^2}{m} \sum_{j=1}^{m} E\left[H(q^{-1})z_j^0(t)\right]\left[H(q^{-1})z_j^0(t)\right]^T \\ &\quad + \frac{\lambda^2}{m} \sum_{j=1}^{m} E\left[H(q^{-1})\tilde{z}_j(t)\right]\left[H(q^{-1})\tilde{z}_j(t)\right]^T \\ &= \lambda^2 E\left[e_{m-1} \otimes H(q^{-1})\varphi^0(t)\right]\left[e_{m-1} \otimes H(q^{-1})\varphi^0(t)\right]^T \\ &\quad + \lambda^2 I_{m-1} \otimes E\left[H(q^{-1})\tilde{\varphi}_j(t)\right]\left[H(q^{-1})\tilde{\varphi}_j(t)\right]^T \\ &= \lambda^2 e_{m-1}e_{m-1}^T \otimes E\left[H(q^{-1})\varphi^0(t)\right]\left[H(q^{-1})\varphi^0(t)\right]^T \\ &\quad + \lambda^2 I_{m-1} \otimes E\left[H(q^{-1})\tilde{\varphi}_j(t)\right]\left[H(q^{-1})\tilde{\varphi}_j(t)\right]^T \end{aligned} \quad (4.15)$$

In (4.15) I_{m-1} denotes the identity matrix of dimension $m-1$. To simplify the expression, introduce the notations:

$$\begin{aligned} C_1 &= E\left[H(q^{-1})\varphi^0(t)\right]\left[H(q^{-1})\varphi^0(t)\right]^T \\ C_2 &= E\left[H(q^{-1})\tilde{\varphi}_j(t)\right]\left[H(q^{-1})\tilde{\varphi}_j(t)\right]^T \end{aligned} \quad (4.16)$$

Note that both C_1 and C_2 are positive definite.

Hence, by (4.15)

$$S_{IV3} = \lambda^2 e_{m-1}e_{m-1}^T \otimes C_1 + \lambda^2 I_{m-1} \otimes C_2 \quad (4.17)$$

It is of course of interest to find the inverse matrix S^{-1} in order to apply optimal weighting. We have the following result:

Lemma Consider the instrumental vectors chosen in Example 3. It holds that

$$S_{IV3}^{-1} = \frac{1}{\lambda^2}[I_{m-1} \otimes C_2^{-1} - e_{m-1}e_{m-1}^T \otimes (C_2C_1^{-1}C_2 + (m-1)C_2)^{-1}] \quad (4.18)$$

Furthermore, the optimal covariance matrix becomes in this case

$$P_{IV3}^{opt} = \lambda^2 R_0^{-1} \left(C_1 + \frac{C_2}{m-1} \right) R_0^{-1} \quad (4.19)$$

Proof. See Appendix.

The expression (4.19) gives the asymptotic covariance matrix of the parameters estimates when the optimal weighting $Q = S^{-1}$is applied. It hence provides a lower bound on the achievable accuracy for a large class of estimators.

Next we examine whether the lower bound can be achieved for other Q. Let us evaluate the degradation of using $Q = I$. Due to the general form of the covariance matrix P in (4.2), it follows that

$$P(I) = (R^TR)^{-1}R^TSR(R^TR)^{-1}. \quad (4.20)$$

By inserting (3.8) and (4.17) in the (4.20), we find

$$P_{IV3}(I) = \lambda^2 R_0^{-1} \left(C_1 + \frac{C_2}{m-1} \right) R_0^{-1} \quad (4.21)$$

(see Appendix for details). Comparing with (4.21), we find that the optimal performance is in fact achieved also with $Q = I$. The choice $Q = I$ is to be preferred, as it leads to significantly simpler computations than the choice (4.12).

4.2 Comparing the asymptotic covariance matrix of the parameters for different estimates

For Example 1, the covariance matrix of $\hat{\theta}$ is:

$$P_{IV1} = \frac{\lambda^2 m}{m-1} R_0^{-1}(C_1 + C_2)R_0^{-1} \quad (4.22)$$

while for Example 2 it holds

$$P_{IV2} = \lambda^2 R_0^{-1}(C_1 + C_2)R_0^{-1} \quad (4.23)$$

(see Appendix).

From the explicit expressions (4.19), 4.19), (4.22) and (4.23), we easily get

$$P_{IV1} \geq P_{IV2} \geq P_{IV3} = P^{opt} \quad (4.24)$$

which partly re-establishes (4.10). The instrumental variable method of Example 3 gives the best accuracy compared to variants 1 and 2. The noise contribution in P_{IV3} is reduced a factor $1/(m-1)$ by using the over–determined instrumental variable vector.

4.3 IV Example 4

By analyzing the covariance matrix of the estimated data, the advantage of using over-determined instrumental vectors is obvious. For m periods data, as long as the Assumption A3 is met, there are totally $(m-1)^m$ different combinations of $\varphi_k(t)$ that can be used for forming the instruments $z_j(t)$. We may choose instrumental vector $Z(t)$ as gathering all these $(m-1)^m$ corporations of data. As an illustration, for $m=3$,

$$\begin{aligned}
\phi(t) &= \begin{bmatrix} \varphi_1(t) & \varphi_2(t) & \varphi_3(t) \end{bmatrix} \\
Z(t) &= \begin{bmatrix}
\varphi_2(t) & \varphi_3(t) & \varphi_1(t) \\
\varphi_2(t) & \varphi_3(t) & \varphi_2(t) \\
\varphi_2(t) & \varphi_1(t) & \varphi_1(t) \\
\varphi_2(t) & \varphi_1(t) & \varphi_2(t) \\
\varphi_3(t) & \varphi_3(t) & \varphi_1(t) \\
\varphi_3(t) & \varphi_3(t) & \varphi_2(t) \\
\varphi_3(t) & \varphi_1(t) & \varphi_1(t) \\
\varphi_3(t) & \varphi_1(t) & \varphi_2(t)
\end{bmatrix}
\end{aligned}$$

We will call this kind of instrumental vectors as IV4. It is interesting to analyze whether there are any benefits we can get from it.

For simplicity, we set

$$g = (m-1)^m. \tag{4.25}$$

Similarly to the instrumental variant IV3, we get

$$R_{IV4} = e_g \otimes R_0 \tag{4.26}$$

where $e_g = (1 \ldots 1)^T$ has dimension $g \times 1$. Note that here the matrix $Z(t)$ has dimension $g \times m$. In this case it holds

$$S_{IV4} = \lambda^2 (e_g e_g^T \otimes C_1 + M_g \otimes C_2) \tag{4.27}$$

where, $M_g = [m_{i,j}] \in R^{g \times g}$ with $m_{i,j} = 1$ if $i = j$, and $m_{i,j} \in [0, 1/m, ..., (m-1)/m]$ if $i \neq j$.

Proof. See Appendix.

As an illustration for $m = 3$, and $Z(t)$ as above, the matrix M_g reads

$$M_g = \frac{1}{3}\begin{bmatrix}
3 & 2 & 2 & 1 & 2 & 1 & 1 & 0 \\
2 & 3 & 1 & 2 & 1 & 2 & 0 & 1 \\
2 & 1 & 3 & 2 & 1 & 0 & 2 & 1 \\
1 & 2 & 2 & 3 & 0 & 1 & 1 & 2 \\
2 & 1 & 1 & 0 & 3 & 2 & 2 & 1 \\
1 & 2 & 0 & 1 & 2 & 3 & 1 & 2 \\
1 & 0 & 2 & 1 & 2 & 1 & 3 & 2 \\
0 & 1 & 1 & 2 & 1 & 2 & 2 & 3
\end{bmatrix}$$

When $Q = I$, we have

$$P_{IV4}(I) \;=\; \lambda^2 R_0^{-1}\left(C_1 + \frac{e_g^T M_g e_g}{g^2} C_2\right) R_0^{-1} \tag{4.28}$$

Proof. See Appendix.

Considering $M_g e_g = \alpha e_g$, where $\alpha = (m-1)^{(m-1)}$ (see Appendix),

$$\begin{aligned}
\frac{e_g^T M_g e_g}{g^2} &= \frac{(m-1)^{(m-1)} e_g^T e_g}{g^2} \\
&= \frac{(m-1)^{(m-1)}}{g} \\
&= \frac{1}{m-1}
\end{aligned}$$

Hence,

$$P_{IV4}(I) \;=\; \lambda^2 R_0^{-1}\left(C_1 + \frac{C_2}{m-1}\right) R_0^{-1}. \tag{4.29}$$

We can further prove (see Appendix) that

$$P_{IV4}^{opt} \;=\; P_{IV4}(I)$$

i.e.

$$P_{IV4}^{opt} = P_{IV4}(I) = \lambda^2 R_0^{-1}\left(C_1 + \frac{C_2}{m-1}\right) R_0^{-1}. \tag{4.30}$$

Compare (4.30) with (4.19) and (4.21), we can see using the instrumental vector IV4 gives the same estimation accuracy as that of using IV3. As IV3 is computationally the simplest alternative, it is the appropriate choice (with $Q = I$) for achieving optimal accuracy.

5 Numerical illustrations

5.1 Simulation results of proposed IV methods

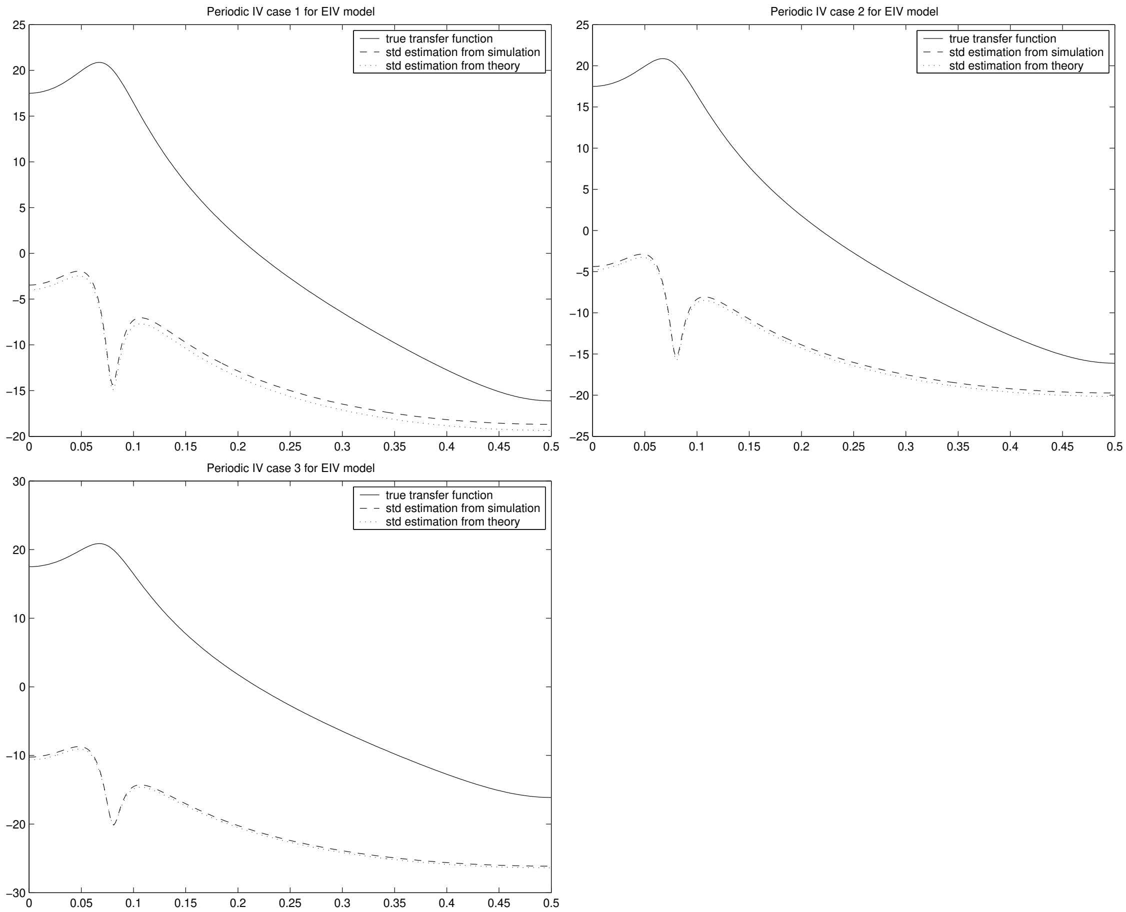
To illustrate numerically the identification methods introduced in the previous section, we consider a second order system with

$$\begin{aligned}
A(q^{-1}) &= 1 - 1.5q^{-1} + 0.7q^{-2} \\
B(q^{-1}) &= 1.0q^{-1} + 0.5q^{-2}
\end{aligned} \tag{5.1}$$

The true input $u_0(t)$ is an ARMA process given by

$$u_0(t) = \frac{1 + 2q^{-1} + q^{-2}}{1 - 1.8q^{-1} + 0.9q^{-2}} e(t) \tag{5.2}$$

Figure 2: The transfer functions based on IV estimates using instrumental variables as Examples 1, 2 and 3. Horizontal axis is the normalized frequency, and the unit of the vertical axis is dB. Simulation with white measurement noise, $\mathrm{SNR}_{input} = 13dB$.



where $e(t)$ is a zero mean white noise sequence with variance 0.25. The number of data points N in each period is 1024, and the number of periods m is 6. For each simulation 500 realizations are done.

As long as Assumption A2 is met, there is no specified restriction on the correlation of $\tilde{u}(t)$ and $\tilde{y}(t)$ within a period. We first let the noise signals $\tilde{u}(t)$ and $\tilde{y}(t)$ be mutually uncorrelated white noise signals. The variances of the measurement noise sequences $\tilde{u}(t)$ and $\tilde{y}(t)$ are both equal to 10.

Simulation results of the IV estimates using the instrumental variable vectors as in Examples 1–3 are illustrated in Figure 2. The simulation results are numerically summarized in Table 1. As showed both in Figure 2 and Table 1, it is clear that the theoretical analysis previous is well supported by the Monte-Carlo simulation.

Figure 3 gives a comparison of the three IV estimates (Examples 1, 2 and 3) for both small and large amounts of measurement noise. It shows that the benefit of using the overdetermined instrumental variable vector (IV3) is more pronounced when the measurement noise is large (*i.e.* in small SNR conditions).

Figure 3: For white measurement noise condition, comparison of different IV estimates (with instrumental variables as Example 1, 2 and 3) for $\mathrm{SNR}_{input} = 26dB$ and $\mathrm{SNR}_{input} = 13dB$, respectively. Horizontal axis is the normalized frequency, and the unit of the vertical axis is dB.

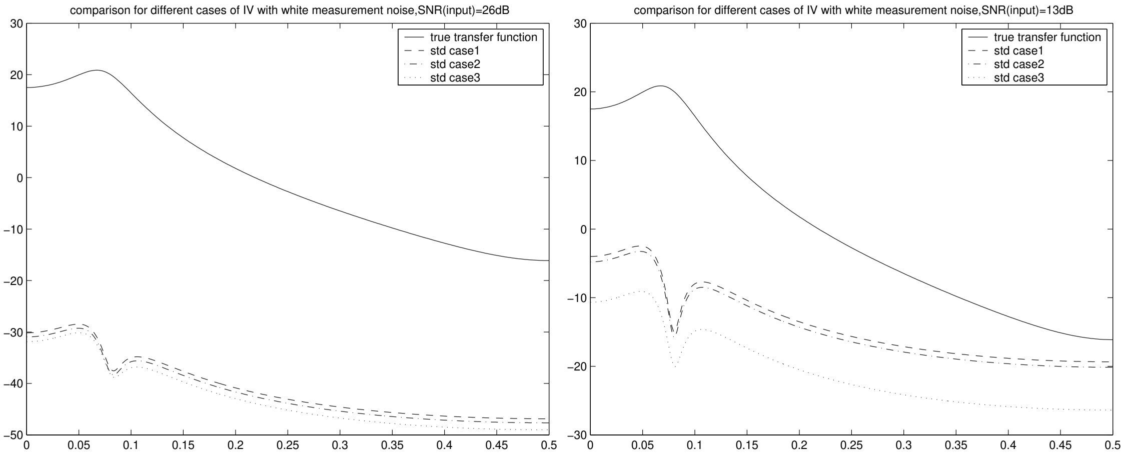


Table 1: Simulation results with the white measurement noise, $\mathrm{SNR}_{input} = 13dB$. (s)= simulation, (t)= theory

Estimate		a_1	a_2	b_1	b_2
$\hat{\theta}_{IV1}$	Mean	-1.501	0.700	1.006	0.490
	std(s)	±0.011	±0.0036	±0.148	±0.214
	std(t)	±0.010	±0.0035	±0.141	±0.203
$\hat{\theta}_{IV2}$	Mean	-1.500	0.700	1.001	0.499
	std(s)	±0.0099	±0.0033	± 0.137	± 0.197
	std(t)	±0.0091	±0.0032	±0.129	±0.185
$\hat{\theta}_{IV3}$	Mean	-1.499	0.700	0.996	0.507
	std(s)	±0.0058	±0.0030	±0.064	± 0.093
	std(t)	±0.0054	±0.0028	±0.063	±0.090

Next, we consider a case where $\tilde{u}(t)$ and $\tilde{y}(t)$ are mutually uncorrelated, but colored measurement noise.

$$\tilde{u}(t) = (1 + q^{-1})w_u(t) \tag{5.3}$$

$$\tilde{y}(t) = \frac{1}{1 - 0.5q^{-1}} w_y(t) \tag{5.4}$$

where $w_u(t)$ and $w_y(t)$ are uncorrelated zero mean white noise sequences with variance 5 and 7.5, respectively (to keep the SNR at the input and output sides at the same values as in white measurement noise case).

We show in Figure 4 and Table 2 that all the three IV estimators considered can also work well in the colored noise case.

Figure 4: For colored measurement noise condition, comparison of different IV estimates (with instrumental variables as Example 1, 2 and 3) for $\mathrm{SNR}_{input} = 23dB$ and $\mathrm{SNR}_{input} = 13dB$, respectively. Horizontal axis is the normalized frequency, and the unit of the vertical axis is dB.

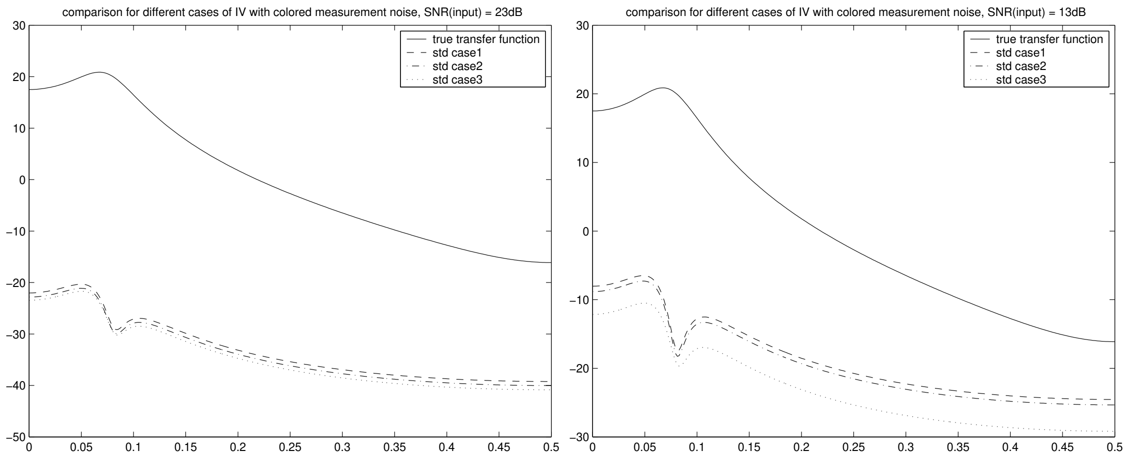


Table 2: Simulation results with the colored measurement noise, $\mathrm{SNR}_{input} = 13dB$. (s)= simulation, (t)= theory

Estimate		a_1	a_2	b_1	b_2
$\hat{\theta}_{IV1}$	Mean	-1.500	0.699	0.999	0.505
	std(s)	± 0.012	± 0.0065	$\pm$ 0.123	± 0.183
	std(t)	± 0.012	± 0.0062	± 0.116	± 0.142
$\hat{\theta}_{IV2}$	Mean	-1.500	0.700	1.004	0.508
	std(s)	± 0.011	$\pm$ 0.0058	± 0.112	± 0.166
	std(t)	± 0.010	± 0.0056	± 0.106	± 0.157
$\hat{\theta}_{IV3}$	Mean	-1.500	0.700	1.005	0.493
	std(s)	± 0.0058	± 0.0053	± 0.060	$\pm$ 0.084
	std(t)	± 0.0058	± 0.0049	± 0.064	± 0.092

5.2 Comparison with the Basic IV method

For comparison, suppose we do not have the knowledge that the data are periodic, and use $Z(t) = \phi(t - L)$ as the instrumental vector. We denote this method as the Basic IV in this report. Under the white noise conditions, both the Basic IV and the IV methods treated earlier will give consistent estimates. When the input and output noises are MA(k) process, Basic IV will be consistent only $L \geq n + k$, where n and k are the orders of system and noise, respectively. If the noise is an ARMA process, then Basic IV will be inconsistent except for some bias compensation algorithms. The IV methods IV1, IV2, IV3 all result in consistent estimates under all these noise conditions, as long as Assumption A2 is satisfied.

Here, the Basic IV method was simulated in different noise conditions. In first case, $\tilde{u}(t)$ and $\tilde{y}(t)$ are white noise with variance 10. In the second noise case, $\tilde{u}(t)$ and $\tilde{y}(t)$

are mutually uncorrelated $MA(1)$ process.

$$\tilde{u}(t) = (1+q^{-1})w_u(t) \tag{5.5}$$
$$\tilde{y}(t) = (1-q^{-1})w_y(t) \tag{5.6}$$

where $w_u(t)$ and $w_y(t)$ are uncorrelated zero mean white noise sequences with variance 5 (to keep the SNR at the input and output sides at the same values as in the white measurement noise case). Basic IV estimation is run with both $L = 4$ (*i.e.* $L \geq n+k$), and $L = 2$ (*i.e.* $L < n+k$). Finally, Basic IV was also applied with ARMA noise for $L = 4$. Detailed estimation results are listed in Table 3. Comparison of Basic IV and IV1, IV2, IV3 are shown in Figure 5. From Table 3, it can be seen that Basic IV will give consistent estimates only in the white measurement noise or when $L \geq n+k$ in case the noise is an MA process. Just as expected, in Figure 5 we see in the consistent cases that Basic IV gives the worst accuracy in comparison with the other IV estimates which use the knowledge of periodic data.

Figure 5: Comparison of Basic IV and different IV estimates with instrumental variable as in Example 1, 2 and 3 for $\mathrm{SNR}_{input} = 13dB$, in white (left diagram) and colored (right diagram) measurement noise MA(1) conditions. Horizontal axis is the normalized frequency, and the unit of the vertical axis is dB.

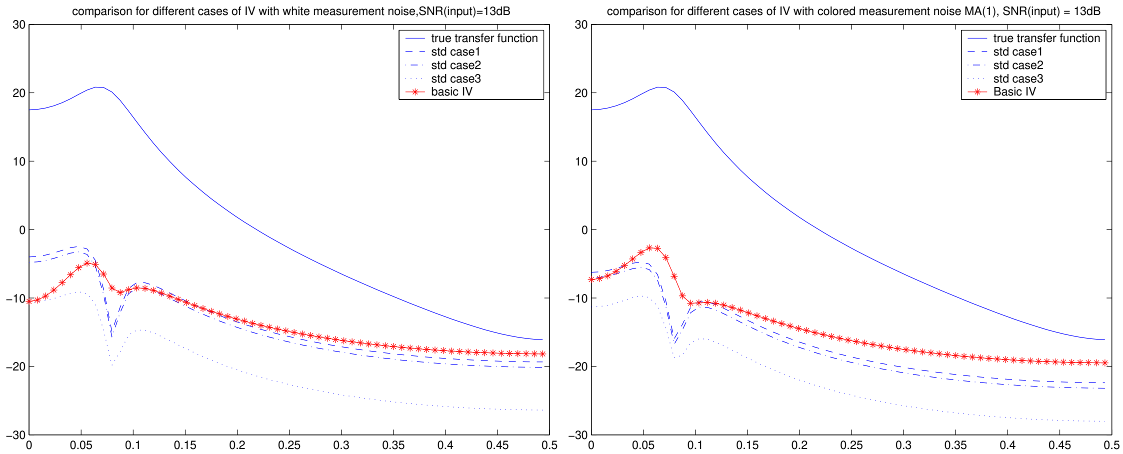


Table 3: Basic IV estimation results with different noise conditions.

Noise type		a_1	a_2	b_1	b_2
True value		-1.5000	0.7000	1.0000	0.5000
white noise	Mean	-1.4978	0.6948	1.1080	0.3880
	std(s)	±0.0067	±0.0101	±0.1878	±0.1889
MA(1), L=4	Mean	-1.5014	0.7011	0.9941	0.5002
	std(s)	±0.0189	±0.0186	±0.2281	±0.1947
MA(1), L=2	Mean	-1.5647	0.7676	0.0623	1.1152
	std(s)	±6.8318	±7.9721	±109.11	±91.181
ARMA, L=4	Mean	0.3770	-1.7357	3.8087	-3.2484
	std(s)	±36.413	±47.872	±744.72	±664.25

5.3 Comparison with other EIV methods with periodic data

Compensated least squares (CLS) and Compensated total least squares (CTLS) [3] are two algorithms for dynamic EIV system identification using periodic excitation signals. To remove the bias in the least squares estimation, these two methods first subtract away the disturbances by a non-parametric noise model, estimated directly from the measured data (based on averaging or FFT idea), and then use the least squares or total least squares estimation. For comparison, we run the IV3 method of this report under the same conditions as in [3]. The results of Monte-Carlo simulations are shown in the Table 4. We note that the IV3 method works better than CLS and CTLS.

It is also shown that, under this example, the improvement in the accuracy of IV3 with pre-filtered data (IV3-F) is not so obvious as that of CLS-F and CTLS-F (CLS and CTLS with pre-filtered data). Here we use the same methods to pre-whitening the output noise as mentioned in [3], *i.e.* the data were pre-filtered in the frequency domain by multiplying the Fourier square-root factor of the estimated output noise spectrum. The corresponding time-domain signals are obtained through IFFT.

Table 4: Comparing simulation results with other algorithms.

Estimate		a_1	a_2	b_1	b_2
$\hat{\theta}_{CLS}$	Mean	-17.4991	0.6992	0.9970	0.5058
	std(s)	±0.0246	±0.0238	±0.0586	±0.0647
$\hat{\theta}_{CTLS}$	Mean	-1.5102	0.7131	1.0170	0.4698
	std(s)	±0.0255	±0.0235	±0.0636	±0.0634
$\hat{\theta}_{IV3}$	Mean	-1.5005	0.7006	1.0015	0.4983
	std(s)	±0.0145	±0.0141	±0.0318	±0.0367
$\hat{\theta}_{CLS_F}$	Mean	-1.5006	0.7006	0.9988	0.4984
	std(s)	±0.0032	±0.0029	±0.0118	±0.0127
$\hat{\theta}_{CTLS_F}$	Mean	-1.5013	0.7012	1.0013	0.4944
	std(s)	±0.0032	±0.0029	±0.0118	±0.0126
$\hat{\theta}_{IV3_F}$	Mean	-1.4927	0.6913	0.9965	0.5052
	std(s)	±0.0111	±0.0121	±0.0194	±0.0200

6 Conclusions

The motivation of this report was to show how the information of periodic measurement data could be used in identifying systems with errors-in-variables models by using the instrumental variable estimates. Four different instrumental variable vectors are considered. Consistency and accuracy analysis shows that all these IV methods would give consistent estimates.

The best accuracy is achieved when using an overdetermined instrumental variable vector. The optimal weighting matrix and the lower bound on the covariance matrix of the parameter estimates were discussed. The theoretical results are further supported by Monte-Carlo simulation for both white and colored measurement noise conditions. For low SNR, the advantages of using the overdetermined instrumental variable vector is more manifest. The overdetermined IV method proposed in this report achieves better estimation accuracy in examples than CLS and CTLS methods previously proposed in the literature.

A Appendix

A.1 Proof of Lemma

By using properties of the Kronecker product and the matrix inversion lemma

$$\begin{aligned}
S_{IV3}^{-1} &= \left[\lambda^2 e_{m-1}e_{m-1}^T \otimes C_1 + \lambda^2 I_{m-1} \otimes C_2\right]^{-1} \\
&= \frac{1}{\lambda^2}[I_{m-1} \otimes C_2 + (e_{m-1} \otimes I)(1 \otimes C_1)(e_{m-1}^T \otimes I)]^{-1} \\
&= \frac{1}{\lambda^2}[(I_{m-1} \otimes C_2^{-1}) - (I_{m-1} \otimes C_2^{-1})(e_{m-1} \otimes I) \\
&\quad \times[(1 \otimes C_1)^{-1} + (e_{m-1}^T \otimes I)(I_{m-1} \otimes C_2^{-1})(e_{m-1} \otimes I)]^{-1} \\
&\quad \times(e_{m-1}^T \otimes I)(I_{m-1} \otimes C_2^{-1})] \\
&= \frac{1}{\lambda^2}[(I_{m-1} \otimes C_2^{-1}) - (e_{m-1} \otimes C_2^{-1}) \\
&\quad \times[C_1^{-1} + (m-1)C_2^{-1}]^{-1}(e_{m-1}^T \otimes C_2^{-1})] \\
&= \frac{1}{\lambda^2}[I_{m-1} \otimes C_2^{-1} - e_{m-1}e_{m-1}^T \otimes C_2^{-1}[C_1^{-1} + (m-1)C_2^{-1}]^{-1}C_2^{-1}] \\
&= \frac{1}{\lambda^2}[I_{m-1} \otimes C_2^{-1} - e_{m-1}e_{m-1}^T \otimes [C_2C_1^{-1}C_2 + (m-1)C_2]^{-1}]
\end{aligned}$$

which proves (4.18).

Using (3.8) and (4.17) in (4.11) gives

$$\begin{aligned}
(P_{IV3}^{opt})^{-1} &= (e_{m-1}^T \otimes R_0)[S^{-1}](e_{m-1} \otimes R_0) \\
&= \frac{1}{\lambda^2}[e_{m-1}^T e_{m-1} \otimes (R_0C_2^{-1}R_0) - e_{m-1}^T e_{m-1}e_{m-1}^T e_{m-1} \otimes R_0 \\
&\quad \times[C_2C_1^{-1}C_2 + (m-1)C_2]^{-1}R_0] \\
&= \frac{1}{\lambda^2}[(m-1)(R_0C_2^{-1}R_0) - (m-1)^2R_0 \\
&\quad \times[C_2C_1^{-1}C_2 + (m-1)C_2]^{-1}R_0] \\
&= \frac{(m-1)}{\lambda^2}R_0[C_2^{-1} - (m-1)[C_2C_1^{-1}C_2 + (m-1)C_2]^{-1}]R_0 \\
&= \frac{(m-1)}{\lambda^2}R_0C_2^{-1}[C_2C_1^{-1}C_2 + (m-1)C_2 - (m-1)C_2] \\
&\quad \times[C_2C_1^{-1}C_2 + (m-1)C_2]^{-1}R_0 \\
&= \frac{(m-1)}{\lambda^2}R_0[(C_2C_1^{-1}C_2 + (m-1)C_2)(C_1^{-1}C_2)^{-1}]^{-1}R_0 \\
&= \frac{(m-1)}{\lambda^2}R_0[C_2 + (m-1)C_1]^{-1}R_0 \\
&= \frac{1}{\lambda^2}R_0\left[C_1 + \frac{C_2}{(m-1)}\right]^{-1}R_0
\end{aligned}$$

By inverting this we get directly (4.19).

A.2 Covariance matrices for the different IV estimators

Since the R is a square matrix, for both Example 1 and 2, the weighting matrix Q is unnecessary in calculating the covariance matrix of estimate parameters. We then have:

$$P = (R^T R)^{-1} R^T S R (R^T R)^{-1} \tag{A.1}$$

For Example 1, from (2.11), (4.9) and (4.16), we find

$$\begin{aligned}
S &= \frac{\lambda^2}{m}\sum_{j=1}^{m-1} E[H(q^{-1})z_j(t)][H(q^{-1})z_j(t)]^T \\
&= \frac{\lambda^2}{m}\sum_{j=1}^{m-1} E[H(q^{-1})z_j^0(t)][H(q^{-1})z_j^0(t)]^T \\
&\quad +\frac{\lambda^2}{m}\sum_{j=1}^{m-1} E[H(q^{-1})\tilde{z}_j(t)][H(q^{-1})\tilde{z}_j(t)]^T \\
&= \frac{\lambda^2}{m}(m-1)E[H(q^{-1})\varphi^0(t)][H(q^{-1})\varphi^0(t)]^T \\
&\quad +\frac{\lambda^2}{m}(m-1)E[H(q^{-1})\tilde{\varphi}_j(t)][H(q^{-1})\tilde{\varphi}_j(t)]^T \\
&= \frac{\lambda^2}{m}(m-1)(C_1 + C_2)
\end{aligned} \tag{A.2}$$

Inserting (3.6) and (5.2) into (5.1) yields:

$$\begin{aligned}
P_{IV1} &= \left[\left(\frac{m-1}{m}\right)^2 R_0^T R_0\right]^{-1} \left(\frac{m-1}{m}\right) R_0^T \\
&\quad \times\lambda^2 \left(\frac{m-1}{m}\right)(C_1 + C_2)\left(\frac{m-1}{m}\right) R_0 \left[\left(\frac{m-1}{m}\right)^2 R_0^T R_0\right]^{-1} \\
&= \frac{\lambda^2 m}{m-1} R_0^{-1}(C_1 + C_2)R_0^{-1}
\end{aligned}$$

which is (4.22).

Similarly, for Example 2, from (2.12), (4.9) and (4.16), we get

$$S = \lambda^2 (C_1 + C_2) \tag{A.3}$$

and

$$\begin{aligned}
P_{IV2} &= (R_0^T R_0)^{-1} R_0^T \lambda^2 (C_1 + C_2) R_0 (R_0^T R_0)^{-1} \\
&= \lambda^2 R_0^{-1}(C_1 + C_2) R_0^{-1}
\end{aligned}$$

For Example 3, when $Q = I$, using (3.8) and (4.17), equation (4.20) shows

$$\begin{aligned}
P_{IV3}(I) &= \left[(e_{m-1} \otimes R_0)^T (e_{m-1} \otimes R_0)\right]^{-1} (e_{m-1} \otimes R_0)^T \\
&\quad \times \lambda^2 (e_{m-1} e_{m-1}^T \otimes C_1 + I_{m-1} \otimes C_2) \\
&\quad \times (e_{m-1} \otimes R_0) \left[(e_{m-1} \otimes R_0)^T (e_{m-1} \otimes R_0)\right]^{-1} \\
&= \lambda^2 \left[(e_{m-1}^T e_{m-1})^{-1} e_{m-1}^T \otimes (R_0^T R_0)^{-1} R_0^T\right] \\
&\quad \times (e_{m-1} e_{m-1}^T \otimes C_1 + I_{m-1} \otimes C_2) \\
&\quad \times \left[e_{m-1} (e_{m-1}^T e_{m-1})^{-1} \otimes R_0 (R_0^T R_0)^{-1}\right] \\
&= \frac{\lambda^2}{(m-1)^2} [e_{m-1}^T e_{m-1} e_{m-1}^T e_{m-1} \otimes R_0^{-1} C_1 R_0^{-1} \\
&\quad + e_{m-1}^T I_{m-1} e_{m-1} \otimes R_0^{-1} C_2 R_0^{-1}] \\
&= \frac{\lambda^2}{(m-1)^2} [(m-1)^2 R_0^{-1} C_1 R_0^{-1} + (m-1) R_0^{-1} C_2 R_0^{-1}] \\
&= \lambda^2 R_0^{-1} \left(C_1 + \frac{C_2}{m-1}\right) R_0^{-1} \qquad \text{(A.4)}
\end{aligned}$$

Hence, (4.21) is proved.

For Example 4, S_{IV4} can be evaluated in a matter similar to S_{IV3}, *cf.* (4.15):

$$\begin{aligned}
S_{IV4} &= \lim_{N \to \infty} \frac{\lambda^2}{Nm} \sum_{j=1}^{m} \sum_{t=1}^{N} \left[H(q^{-1}) z_j(t)\right] \left[H(q^{-1}) z_j(t)\right]^T \\
&= \frac{\lambda^2}{m} \sum_{j=1}^{m} E\left[H(q^{-1}) z_j^0(t)\right] \left[H(q^{-1}) z_j^0(t)\right]^T \\
&\quad + \frac{\lambda^2}{m} \sum_{j=1}^{m} E\left[H(q^{-1}) \tilde{z}_j(t)\right] \left[H(q^{-1}) \tilde{z}_j(t)\right]^T \\
&= \lambda^2 (e_g e_g^T \otimes C_1 + M_g \otimes C_2)
\end{aligned}$$

This proves (4.27).

When $Q = I$, inserting (4.26) and (4.27) into (4.20), we have, *cf.* (A.4):

$$\begin{aligned}
P_{IV4}(I) &= \left[(e_g \otimes R_0)^T (e_g \otimes R_0)\right]^{-1} (e_g \otimes R_0)^T \\
&\quad \times \lambda^2 (e_g e_g^T \otimes C_1 + M_g \otimes C_2) \\
&\quad \times (e_g \otimes R_0) \left[(e_g \otimes R_0)^T (e_g \otimes R_0)\right]^{-1} \\
&= \lambda^2 \left[(e_g^T e_g)^{-1} e_g^T \otimes ((R_0^T R_0)^{-1} R_0^T)\right] \left[(e_g e_g^T \otimes C_1 + M_g \otimes C_2)\right] \\
&\quad \times \left[e_g (e_g^T e_g)^{-1} \otimes (R_0 (R_0^T R_0)^{-1})\right] \\
&= \lambda^2 [1 \otimes (R_0^{-1} C_1 R_0^{-1}) + \frac{e_g^T M_g e_g}{g^2} \otimes (R_0^{-1} C_2 R_0^{-1})] \\
&= \lambda^2 R_0^{-1} \left(C_1 + \frac{e_g^T M_g e_g}{g^2} C_2\right) R_0^{-1}
\end{aligned}$$

which is (4.28).

A.3 Proof of $P^{opt}_{IV4} = P_{IV4}(I)$

Consider first the difference

$$\begin{aligned} P(Q) - P(S^{-1}) &= (R^TQR)^{-1}[R^TQSQR - R^TQR \\ &\quad \times(R^TS^{-1}R)^{-1}R^TQR](R^TQR)^{-1} \\ &= (R^TQR)^{-1}R^TQ[S - R(R^TS^{-1}R)^{-1}R^T]QR(R^TQR)^{-1} \end{aligned}$$

Set

$$\begin{aligned} X &= S^{-1/2}R \\ A &= S^{1/2}[I - X(X^TX)^{-1}X^T]S^{1/2}, \end{aligned}$$

Then

$$P(Q) - P(S^{-1}) = (R^TQR)^{-1}R^TQAQR(R^TQR)^{-1}$$

Since A is a symmetric, positive semidefinite, and idempotent matrix,

$$P(Q) - P(S^{-1}) \geq 0. \tag{A.5}$$

It is of interest to examine when *strict equality* holds in (A.5). First note that $I - X(X^TX)^{-1}X^T$ is the orthogonal projection onto $R(X)^\perp$. Hence we have the following sequence of equivalences:

$$\begin{aligned} P(Q) \equiv P(S^{-1}) &\Leftrightarrow \\ R^TQAQR = 0 &\Leftrightarrow \\ R^TQS^{1/2}[I - X(X^TX)^{-1}X^T]S^{1/2}QR = 0 &\Leftrightarrow \\ S^{1/2}QR \in R(X) &\Leftrightarrow \\ S^{1/2}QR = XC(\text{some nonsingular matrix C}) &\Leftrightarrow \\ SQR = RC & \end{aligned} \tag{A.6}$$

It is trivial that $P(Q)|_{Q=S^{-1}} = P^{opt}$. This can also be seen from (A.6) by taking $C = I$ in this case.

To examine if $P(I) = P(S^{-1})$ for IV4, it is necessary to exploit the specific structure of R and S. For simplicity we can neglect the factor λ^2 of S in (4.27). Then we have in this case,

$$\begin{aligned} R &= e_g \otimes R_0 \\ S &= e_ge_g^T \otimes C_1 + M_g \otimes C_2 \end{aligned}$$

where C_1, C_2, R_0 are symmetric matrices and R_0 is nonsingular.

To proceed we need to explore some properties of the matrix M_g. Note from (4.28) that the quadratic form $e_g^TM_ge_g$ will be of interest. We have the following result

Lemma A.1. It holds that

$$M_g e_g \quad = \quad \alpha e_g \tag{A.7}$$

where

$$\alpha = (m-1)^{(m-1)} \tag{A.8}$$

proof. For each row of M_g, how many elements take the value $(m-k)/m$? This number turns out to

$$\gamma_k \quad = \quad \binom{m}{m-k}(m-2)^k \qquad (k = 0, 1, ..., m) \tag{A.9}$$

The relation (A.9) can be seen as follows, [2]. Consider two arbitrary rows of the matrix $Z(t)$. The probability that, for a fixed position, the corresponding elements are equal is

$$p = \frac{1}{m-1}$$

as there are $(m-1)$ different values allowed (and one value already taken in the vector $\phi(t)$). The probability that the elements are different is hence

$$q = 1 - p = \frac{m-2}{m-1}$$

The number of equal elements in the two rows of $Z(t)$ is distributed as Bin(n,p), *i.e.*, it has a pdf

$$p(k) = \binom{m}{k} p^k q^{m-k} \qquad (k = 0, 1, ..., m)$$

This probability can also be expressed as the number of outcomes leading to k equal elements divided by the number of all outcomes. Hence

$$p(k) = \frac{\gamma_{m-k}}{(m-1)^m}$$

Hence

$$\begin{aligned}
\gamma_k \quad &= \quad (m-1)^m p(m-k) \\
&= \quad (m-1)^m \binom{m}{m-k} \frac{1}{(m-1)^{m-k}} \left(\frac{m-2}{m-1}\right)^k \\
&= \quad \binom{m}{m-k}(m-2)^k
\end{aligned}$$

which proves (A.9).

As all elements of e_g are equal to one, an arbitrary element of $M_g e_g$ turns out to be

$$\eta \quad = \quad 1 \times \gamma_0 + \frac{(m-1)}{m} \times \gamma_1 + \ldots + \frac{(m-m)}{m} \times \gamma_m$$

$$
\begin{aligned}
&= \sum_{k=0}^{m} \frac{(m-k)}{m}\gamma_k \\
&= \sum_{k=0}^{m} \frac{(m-k)}{m}\binom{m}{m-k}(m-2)^k \\
&= \sum_{k=0}^{m-1} \frac{(m-1)!}{k!(m-1-k)!}(m-2)^k \\
&= \sum_{k=0}^{m-1} \binom{m-1}{k}(m-2)^k 1^{m-1-k} \\
&= [(m-2)+1]^{m-1} \\
&= \alpha
\end{aligned}
$$

and the lemma is proved. ■

Consider again (A.6), and now for the case $Q = I$.

$$
\begin{aligned}
SQR - RC|_{Q=I} &= SR - RC \\
&= e_g e_g^T e_g \otimes C_1 R_0 + M_g e_g \otimes C_2 R_0 - e_g \otimes R_0 C \\
&= e_g \otimes gC_1 R_0 + e_g \otimes \alpha C_2 R_0 - e_g \otimes R_0 C \\
&= e_g \otimes (gC_1 R_0 + \alpha C_2 R_0 - R_0 C)
\end{aligned}
$$

If we choose

$$
C = R_0^{-1}(gC_1 + \alpha C_2)R_0
$$

then $SR = RC$.

Hence

$$
P_{IV4}^{opt} = P_{IV4}^{S^{-1}} = P_{IV4}(I)
$$

References

[1] S. Beghelli, R.P. Guidorzi, and U. Soverini. The Frisch scheme in dynamic system identification. *Automatica*, 26:171–176, 1990.

[2] D. Söderström. Private communication.

[3] U. Forsell, F. Gustafsson, and T. McKelvey. Time-domain identification of dynamic errors-in-variables systems using periodic excitation signals. In *Proc. IFAC 14th World Congress*, Beijing, P.R.China, 1999.

[4] R. Pintelon and J. Schoukens. *System Identification. A Frequency Domain Approach*. IEEE Press, New York, NY, USA, 2001.

[5] J. Schoukens, R. Pintelon, G. Vandersteen, and P. Guillaume. Frequency domain system identification using non-parametric noise models estimated from a small number of data sets. *Automatica*, 33:1073–1086, 1997.

[6] T. Söderström. Identification of stochastic linear systems in presence of input noise. *Automatica*, 17:713–725, 1981.

[7] T Söderström. *Discrete-Time Stochastic Systems - Estimation and Control, 2nd edition*. Springer-Verlag, London, UK, 2002.

[8] T. Söderström, U. Soverini, and K. Mahata. Perspectives on errors-in-variables estimation for dynamic systems. *Signal Processing*, 82(8):1139–1154, August 2002.

[9] T. Söderström and P. Stoica. *Instrumental Variable Methods for System Identification*. Springer-Verlag, Berlin, 1983.

[10] T. Söderström and P. Stoica. *System Identification*. Prentice Hall International, Hemel Hempstead, UK, 1989.

[11] J. K. Tugnait and Y. Ye. Stochastic system identification with noisy input-output measurement using polyspectra. *IEEE Transactions on Automatic Control*, AC-40:670–683, 1995.

Recent technical reports from the Department of Information Technology

2004-018	Emmanuel Papaioannou, Erik Borälv, Athanasios Demiris, Niklas Johansson, and Nikolaos Ioannidis: *User Interface Design for Multi-platform Interactive Sports Content Broadcasting*
2004-019	Owe Axelsson and Maya Neytcheva: *Eigenvalue Estimates for Preconditioned Saddle Point Matrices*
2004-020	Malin Ljungberg and Krister Åhlander: *Generic Programming Aspects of Symmetry Exploiting Numerical Software*
2004-021	Erik Berg and Erik Hagersten: *Efficient Data-Locality Analysis of Long-Running Applications*
2004-022	Pascal Van Hentenryck, Pierre Flener, Justin Pearson, and Magnus Ågren: *Compositional Derivation of Symmetries for Constraint Satisfaction*
2004-023	Jarmo Rantakokko: *Interactive Learning of Algorithms*
2004-024	Mathias Spjuth, Martin Karlsson, and Erik Hagersten: *Low Power and Conflict Tolerant Cache Design*
2004-025	David Lundberg: *Feasibility Study of WLAN Technology for the Uppsala - Stockholm Commuter Train*
2004-026	David Lundberg: *Ad Hoc Protocol Evaluation and Experiences of Real World Ad Hoc Networking*
2004-027	Thorild Selén: *Reorganization in the Skewed-Associative TLB*
2004-028	Lars-Henrik Eriksson: *Using Formal Methods in a Retrospective Safety Case*
2004-029	Krister Åhlander and Hans Munthe-Kaas: *On Applications of the Generalized Fourier Transform in Numerical Linear Algebra*
2004-030	Wendy Kress and Per Lötstedt: *Time Step Restrictions using Semi-Implicit Methods for the Incompressible Navier-Stokes Equations*
2004-031	Malin Ljungberg: *Curvilinear Coordinates in a PDE Solver Framework; Analysis*
2004-032	Malin Ljungberg and Kurt Otto: *Curvilinear Coordinates in a PDE Solver Framework; Validation*
2004-033	Parosh Abdulla, Johann Deneux, and Pritha Mahata: *Closed, Open and Robust Timed Networks*
2004-034	Parosh Abdulla, Pritha Mahata, and Richard Mayr: *Decidability of Zenoness, Token Liveness and Boundedness of Dense-Timed Petri Nets*
2004-035	Emad Abd-Elrady and Torsten Söderström: *Bias Analysis in Least Squares Estimation of Periodic Signals Using Nonlinear ODE's*
2004-036	Stefan Johansson: *High Order Finite Difference Operators with the Summation by Parts Property Based on DRP Schemes*
2004-037	Torsten Söderström and Mei Hong: *Identification of Dynamic Errors-in-Variables Systems with Periodic Data*

UPPSALA
UNIVERSITET

August 2004
ISSN 1404-3203
http://www.it.uu.se/