Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Identification of Noisy Input-Output Models Using the Least-Squares Based Methods

Wei Xing Zheng
School of Computing and Mathematics
University of Western Sydney
Penrith South DC NSW 1797
Australia

Abstract—**This paper addresses the problem of parameter estimation of noisy input-output models, where the measurements of both the input and the output of the system are corrupted by noise. Motivated by the fact that the Koopmans-Levin method and the maximum likelihood estimation type methods assume the known ratio of the variances of the input noise and the output noise, some key equations are derived by using correlation analysis and the knowledge of the noise variance ratio. An objective function is introduced for the purpose of solely finding the input noise variance. An estimate of the system parameters can then be easily obtained without involving any iteration procedure. This leads to the establishment of an efficient identification algorithm. Performance comparisons with other existing identification methods are made via computer simulations.**

I. INTRODUCTION

Noisy input-output models are a type of mathematical models that are normally used to describe dynamic systems in which the input signal and the output signal are both subjected to noise. Such models have found widespread applications in many areas of engineering [5], [7], [11]. Two recent edited volumes [17] and [18] have documented an ample number of engineering applications, ranging from control engineering to signal processing and to communications. Noisy input-output models are also known as (dynamic) errors-in-variables (EIV) models in statistics and econometrics.

Compared to the conventional system identification problem in which the system input signal is always available noise-free, the problem of identifying noisy input-output models turns out to be a much formidable task. A well-known fact is that the standard least-squares (LS) method and other related methods fail to produce unbiased parameter estimates when applied to noisy input-output models. Research on this important topic can be traced back to about thirty years ago. The continued efforts by many researchers have resulted in a considerable number of different consistent identification algorithms, for instance, the joint-output (JO) method [12], the Koopmans-Levin (KL) method [4], the Frisch scheme method [1], [2], the logarithmic least-squares frequency-domain (LLS-FD) method [6], the combined instrumental variable and weighted subspace fitting (IV-WSF) method [16], and the bias-eliminated least-squares (BELS) methods [19], [20], to just mention a few. Each method may exhibit a mix of advantages and disadvantages, depending upon the system model under investigation and the requirements on the identification task.

Recent years have witnessed a resurgent, strong research interest in identification of noisy input-output models. For example, a direct approach for identifying continuous-time linear systems from noisy input-output measurements is proposed in [9]. The accuracy of the BELS methods in dynamic EIV modeling is analyzed in [8]. A new maximum likelihood (ML) method for identification of linear dynamic EIV systems is introduced in [3]. A frequency domain Gaussian

978-1-4244-3124-3/08/$25.00 ©2008 IEEE

ML algorithm for noisy input-output system identification is developed in [10]. Two recent survey papers [15] and [13] have provided a timely, comprehensive summary of various aspects of dynamic EIV system identification as well as important advances in this area.

This paper is concerned with estimating parameters of linear systems in the presence of white input noise and white output noise. The work is motivated by the interesting observation that several EIV identification methods mentioned above are developed based on the assumption that the ratio of the variances of the input noise and the output noise is known. For example, the KL method is workable only with the known noise variance ratio, while the ML method requires the known noise variance ratio assumption for ensuring estimation consistency in the case of unknown deterministic inputs [3]. As explained in [4], such an assumption is not a very restrictive condition since other estimation techniques, such as Kalman filtering, also need to know this ratio. Further, it is pointed out in [3] that the application of the ML method in various practical situations [17], [18] is not hindered by this ratio assumption.

With this motivation, the objective of the paper is to follow the idea presented in [14] to develop an efficient parameter estimation method for noisy input-output models. Correlation analysis and the knowledge of the noise variance ratio will be utilized to arrive at some key equations such that estimation of the input noise variance will be cast as a minimization problem. The formulated minimization problem will be solved in a straightforward way for a desired estimate of the input noise variance, followed by making the bias correction to get the system parameter estimates. In addition to the algorithmic feature of no need for iterative estimation, the developed method also has an improved performance over the previous BELS type methods in terms of computational complexity.

II. Noisy Input-Output Model

Assume that the transfer function of the underlying linear system is given by

$$G(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} = \frac{b_0 + b_1 q^{-1} + ... + b_m q^{-m}}{1 - a_1 q^{-1} - ... - a_n q^{-n}}. \quad (1)$$

The system is excited by the input signal $r(t)$, so that the output signal $x(t)$ is described by the following input-output dynamic equation

$$x(t) = G(q^{-1}) r(t). \quad (2)$$

In practical situations, both the true input signal $r(t)$ and the true output signal $x(t)$ are often measured in additive noise as

$$u(t) = r(t) + v(t), \qquad y(t) = x(t) + w(t) \quad (3)$$

where $v(t)$ and $w(t)$ represent additive measurement noises at the system input and output terminals, respectively.

Since the model given by (1)-(3) takes into account the input noise $v(t)$ and the output noise $w(t)$ to describe the underlying dynamic system, it is known as the noisy input-output model or the dynamic errors-in-variables model. Figure 1 depicts the linear noisy input-output model under investigation.

The following assumptions that are commonly used in dynamic EIV modeling are adopted.

A1. The transfer function $G(q^{-1})$ is exponentially stable, i.e., all the zeros of the denominator polynomial $A(q^{-1})$ lie strictly inside the unit circle.

A2. The polynomials $A(q^{-1})$ and $B(q^{-1})$ are co-prime.

A3. The orders n and m of the identified model are given.

A4. The driving input signal $r(t)$ is a persistently excitation signal of proper order.

A5. The input noise $v(t)$ and the output noise $w(t)$ are white noises with unknown variances σ_v^2 and σ_w^2, respectively.

A6. $r(t)$, $v(t)$, $w(t)$ are independent of each other statistically.

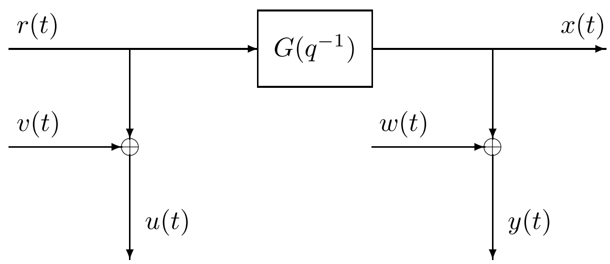

Fig. 1. Linear noisy input-output model.

A7. The ratio α^2 of the output noise variance σ_w^2 to the input noise variance σ_v^2 is known, where α^2 is defined by

$$\alpha^2 = \frac{\sigma_w^2}{\sigma_v^2}. \tag{4}$$

A brief discussion of the above assumptions is in order. Assumptions A1 and A2 are self-explanatory. This paper focuses on parameter estimation, so Assumption A3 is justifiable. Assumption A4 is quite general, and essentially no a priori knowledge of the noiseless input signal $r(t)$ is required. This means that the noiseless input signal $r(t)$ can be either a stationary random sequence modeled by an autoregressive moving average (ARMA) process or an unknown deterministic sequence (e.g., a binary sequence). Since the input noise and the output noise are often measurement errors or sensor noises, Assumption A5 is reasonable. Moreover, the measuring instrumentation noise is usually independent of the signals, and the noise in one instrument is usually independent of that in another, so Assumption A6 is likely to be met in engineering problems. Assumption A7 is the assumption of the known ratio of the noise variances, upon which the KL method and the ML method are built.

The problem of parameter estimation of noisy input-output models is formulated as the one of using noisy input and output measurements $\{u(t), y(t), 1 \leq t \leq N\}$ to obtain an unbiased estimate of the system parameter vector

$$\boldsymbol{\theta}^\top = [\mathbf{a}^\top;\ \mathbf{b}^\top] = [a_1\ ...\ a_n;\ b_0\ b_1\ ...\ b_m] \tag{5}$$

as well as the noise variances σ_v^2 and σ_w^2, where N denotes the number of noisy data points.

III. A Non-Iterative Estimation Method

In this section, like the assumption adopted by the KL method and the ML method, we will make use of the noise variance ratio α^2 to develop an identification algorithm for the noisy input-output system.

Introduce the data regressor vector

$$\begin{aligned}\boldsymbol{\psi}_t^\top &= [\mathbf{y}_t^\top;\ \mathbf{u}_t^\top] \\ &= [y(t-1)\ \cdots\ y(t-n);\ u(t)\ \cdots\ u(t-m)].\end{aligned} \tag{6}$$

The noisy input-output model described by (1)-(3) is recast in a linear regression model

$$y(t) = \boldsymbol{\psi}_t^\top \boldsymbol{\theta} + \varepsilon(t) \tag{7}$$

where the equation error $\varepsilon(t)$ is given by

$$\varepsilon(t) = w(t) - \boldsymbol{\varepsilon}_t^\top \boldsymbol{\theta} \tag{8}$$

and

$$\begin{aligned}\boldsymbol{\varepsilon}_t^\top &= [\mathbf{w}_t^\top;\ \mathbf{v}_t^\top] \\ &= [w(t-1)\ \cdots\ w(t-n);\ v(t)\ \cdots\ v(t-m)].\end{aligned} \tag{9}$$

The standard LS estimate of the system parameter vector $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{LS} = \mathbf{R}_{\psi\psi}^{-1}\mathbf{R}_{\psi y} \tag{10}$$

where

$$\mathbf{R}_{\psi\psi} = E[\boldsymbol{\psi}_t\boldsymbol{\psi}_t^\top] \tag{11}$$

$$\mathbf{R}_{\psi y} = E[\boldsymbol{\psi}_t y(t)]. \tag{12}$$

Following Assumptions A1-A6, it can be shown that $\hat{\boldsymbol{\theta}}_{LS}$ has the following asymptotic expression

$$\hat{\boldsymbol{\theta}}_{LS} = \boldsymbol{\theta} - \mathbf{R}_{\psi\psi}^{-1}\mathbf{D}\boldsymbol{\theta} \tag{13}$$

where

$$\mathbf{D} = \begin{bmatrix} \sigma_w^2 \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \sigma_v^2 \mathbf{I}_{m+1} \end{bmatrix} \tag{14}$$

(see [19], [20] for details). It is clear from (13) and (14) that the bias in the standard LS estimate $\hat{\boldsymbol{\theta}}_{LS}$ is determined by the two noise variances σ_v^2 and σ_w^2.

Now applying Assumption A7 to (14) yields

$$\mathbf{D} = \begin{bmatrix} \alpha^2\sigma_v^2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \sigma_v^2\mathbf{I}_{m+1} \end{bmatrix} = \sigma_v^2 \begin{bmatrix} \alpha^2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m+1} \end{bmatrix} = \sigma_v^2\mathbf{D}_\alpha \tag{15}$$

where

$$\mathbf{D}_\alpha = \begin{bmatrix} \alpha^2\mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m+1} \end{bmatrix}. \tag{16}$$

Then substituting (15) into (13) gives

$$\hat{\boldsymbol{\theta}}_{LS} = \boldsymbol{\theta} - \sigma_v^2\mathbf{R}_{\psi\psi}^{-1}\mathbf{D}_\alpha\boldsymbol{\theta} \tag{17}$$

This clearly shows that the standard LS method is biased, and the estimation bias is determined by the input noise variance σ_v^2.

Next we examine the autocovariance function

$$r_{yy}(j) = E[y(t)y(t-j)] \tag{18}$$

at $j = 0$. It follows from (7) and (12) and the imposed assumptions that

$$r_{yy}(0) = E[y(t)y(t)] = \mathbf{R}_{\psi y}^\top\boldsymbol{\theta} + \sigma_w^2 \tag{19}$$

which, using (4), may be expressed as

$$r_{yy}(0) = \mathbf{R}_{\psi y}^\top\boldsymbol{\theta} + \alpha^2\sigma_v^2. \tag{20}$$

Combining (17) and (20) together and making some rearrangements yields

$$\alpha^2\sigma_v^2 + \sigma_v^2\hat{\boldsymbol{\theta}}_{LS}^\top\mathbf{D}_\alpha\boldsymbol{\theta} = r_{yy}(0) - \mathbf{R}_{\psi y}^\top\hat{\boldsymbol{\theta}}_{LS}. \tag{21}$$

Hence, we have derived two key equations (17) and (20). Using (10), expression (17) can be rearranged as

$$(\mathbf{R}_{\psi\psi} - \sigma_v^2\mathbf{D}_\alpha)\boldsymbol{\theta} = \mathbf{R}_{\psi y}. \tag{22}$$

Then an unbiased BELS estimate of the system parameter vector $\boldsymbol{\theta}$ can be obtained from (22) as

$$\hat{\boldsymbol{\theta}}_{BELS} = (\mathbf{R}_{\psi\psi} - \sigma_v^2\mathbf{D}_\alpha)^{-1}\mathbf{R}_{\psi y} \tag{23}$$

provided that the input noise variance σ_v^2 is known or an estimate of it is available. Furthermore, replacing $\boldsymbol{\theta}$ in (20) by $\hat{\boldsymbol{\theta}}_{BELS}$ given in (23) yields

$$\alpha^2\sigma_v^2 + \mathbf{R}_{\psi y}^\top(\mathbf{R}_{\psi\psi} - \sigma_v^2\mathbf{D}_\alpha)^{-1}\mathbf{R}_{\psi y} = r_{yy}(0). \tag{24}$$

It is important to note that in the above equation (24), all the variables are given or computable except for σ_v^2 that remains as the only unknown variable.

Introduce a function of the independent variable s as follows:

$$f(s) = s\alpha^2 + \mathbf{R}_{\psi y}^\top(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)^{-1}\mathbf{R}_{\psi y} - r_{yy}(0). \tag{25}$$

Apparently, $f(s)$ is a known function of s. Since $f(\sigma_v^2) = 0$, the estimate $\hat{\sigma}_v^2$ of the input noise variance σ_v^2 should be a zero of the function $f(s)$. In order to get an insight into the function $f(s)$, we write the inverse matrix $(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)^{-1}$ as

$$(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)^{-1} = \frac{\operatorname{adj}(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)}{\det(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)} \tag{26}$$

where $\det(\cdot)$ and $\operatorname{adj}(\cdot)$ denote the determinant and the adjoint of a square matrix, respectively. Substituting (26) into (25) produces

$$f(s) = s\alpha^2 + \frac{\mathbf{R}_{\psi y}^\top\operatorname{adj}(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)\mathbf{R}_{\psi y}}{\det(\mathbf{R}_{\psi\psi} - s\mathbf{D}_\alpha)} - r_{yy}(0). \tag{27}$$

The above equation clearly indicates that $f(s)$ is a rational function of s.

On the basis of the above analysis, in order to obtain an estimate of the input noise variance σ_v^2, it is now proposed to minimize the nonnegative function $f^2(s)$ with respect to s, i.e.

$$\min_s f^2(s). \tag{28}$$

Then the positive minimum point of the objective function $f^2(s)$ may be taken as the desired estimate $\hat{\sigma}_v^2$.

Once $\hat{\sigma}_v^2$ is obtained, we will be able to use (23) and (4) to compute $\hat{\boldsymbol{\theta}}_{BELS}$ and $\hat{\sigma}_w^2$, respectively. The proposed non-iterative version of the BELS method with the variance ratio is called BELSR-NI for short.

As can be seen, the major algorithmic difference between the BELS methods in [19] and [20] and the new BELSR-NI method is that the former need to conduct iterative estimation between the noise variances σ_v^2 and σ_w^2 and the system parameter vector $\boldsymbol{\theta}$, whereas the latter provides a closed-form solution without alternating iterations. Thus, the developed BELSR-NI method has a much simpler and more compact structure. It is conceivable that the BELSR-NI method requires fewer computations than the BELS methods in [19] and [20]. On the other hand, it is noted that in adaptive estimation the BELS methods in [19] and [20] are superior to the BELSR-NI method as they can be readily implemented via a recursive scheme.

IV. SIMULATION RESULTS AND CONCLUDING REMARKS

To examine the performance of the BELSR-NI method for short sample lengths and unknown deterministic inputs, we consider estimating a second-order linear EIV system described by (1)-(3) and with the following transfer function

$$G(q^{-1}) = \frac{2 - 1.2q^{-1} - 0.6q^{-2}}{1 - 0.5q^{-1} + 0.3q^{-2}}. \tag{29}$$

The noiseless input $r(t)$ is now a piecewise constant binary sequence with unit variance. The variances of the white measurement noises $v(t)$ and $w(t)$ are given by

$$\sigma_v^2 = 0.1, \qquad \sigma_w^2 = 0.25 \tag{30}$$

which yields that the signal-to-noise ratios (SNR) at the input (SNRI) and the output (SNRO) are approximately 10dB, that is,

$$\text{SNRI} = 10\log_{10}\frac{E[r(t)^2]}{E[v(t)^2} \approx 10\text{dB}, \tag{31a}$$

$$\text{SNRO} = 10\log_{10}\frac{E[x(t)^2]}{E[w(t)^2} \approx 10\text{dB}. \tag{31b}$$

Noisy input-output measurements with a short sample length $N = 250$ are used in identification over $M = 100$ Monte-Carlo runs. Note that this example was studied in [3]. The BELSR-NI method is applied to this noisy input-output model. The corresponding sample means and standard deviations of the estimates of the system parameters and the noise variances are displayed in Table 1. The results by the KL method and the ML method as shown in [3] are also included in Table 1 for convenience of comparison. Note that in Table 1 the RE is the abbreviation of the relative error which is defined by

$$\text{RE} = \frac{\|\mathbf{m}(\hat{\boldsymbol{\theta}}) - \boldsymbol{\theta}\|}{\|\boldsymbol{\theta}\|}, \tag{32}$$

where $\mathbf{m}(\hat{\boldsymbol{\theta}})$ denotes the sample mean of an estimator $\hat{\boldsymbol{\theta}}$.

As can be seen from Table 1, in the case of short sample lengths and unknown deterministic inputs, the BELSR-NI method can work very satisfactorily and their performances are also very comparable to those of the ML method [3]. The attractiveness of the BELSR-NI method lies in that they can be more easily implemented at a lower computational cost than the ML method since the latter usually involves dealing with some highly non-convex optimization problem.

Hence, when a partial information of the input noise and the output noise (such as their variance ratio) becomes available in realistic situations, the use of the developed BELSR-NI method can be very appealing.

REFERENCES

[1] S. Beghelli, R. P. Guidorzi and U. Soverini, "The Frisch scheme in dynamic system identification," *Automatica*, vol. 26, no. 1, pp. 171-176, 1990.

[2] S. Beghelli, P. Castaldi, R. P. Guidorzi and U. Soverini, "A comparison between different model selection criteria in Frisch scheme identification," *Systems Science Journal*, vol. 20, no. 1, pp. 77-84, 1994.

[3] R. Diversi, R. Guidorzi and U. Soverini, "Maximum likelihood identification of noisy input-output models," *Automatica*, vol. 43, no. 3, pp. 464-472, 2007.

[4] K. V. Fernando and H. Nicholson, "Identification of linear systems with input and output noise: the Koopmans-Levin method," *IEE Proceedings–Control Theory Applications*, vol. 132, no. 1, pp. 30-36, 1985.

[5] C. W. J. Granger and P. Newbold, *Forecasting Economic Time Series*. Princeton, NJ: Academic Press, 1986.

[6] P. Guillaume, R. Pintelon and J. Schoukens, "Robust parametric transfer function estimation using complex logarithmic frequency

Table 1. Computer Simulation Results
(N = 250 data points, M = 100 Monte-Carlo runs, SNRI ≈ 10dB, SNRO ≈ 10dB)

method	a_1	a_2	b_0	b_1	b_2	σ_v^2	σ_w^2	RE
LS	0.3545 ±0.0327	−0.2575 ±0.0351	1.8164 ±0.0444	−0.8226 ±0.0732	−0.6524 ±0.0887			18.13%
KL [3]	0.5193 ±0.1633	−0.3138 ±0.1119	2.0012 ±0.1090	−1.2445 ±0.3827	−0.5496 ±0.3774	0.0975 ±0.0059	0.2438 ±0.0148	2.88%
ML [3]	0.5024 ±0.1001	−0.3005 ±0.0574	1.9936 ±0.0855	−1.2029 ±0.2360	−0.5917 ±0.2200	0.0975 ±0.0044	0.2436 ±0.0110	0.45%
BELSR-NI	0.4848 ±0.0520	−0.2861 ±0.0567	1.9979 ±0.0535	−1.1669 ±0.1174	−0.6357 ±0.1514	0.0991 ±0.0096	0.2478 ±0.0240	2.13%
true value	0.5	−0.3	2.0	−1.2	−0.6	0.1	0.25	

response data," *IEEE Transactions on Automatic Control*, vol. 40, no. 7, pp. 1180-1190, 1995.

[7] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall, 1996.

[8] M. Hong, T. Söderström and W. X. Zheng, "Accuracy analysis of bias-eliminating least squares estimates for errors-in-variables systems," *Automatica*, vol. 43, no. 9, pp. 1590-1596, 2007.

[9] K. Mahata and H. Garnier, "Identification of continuous-time errors-in-variables models," *Automatica*, vol. 42, no. 9, pp. 1477-1490, 2006.

[10] R. Pintelon and J. Schoukens, "Frequency domain maximum likelihood estimation of linear dynamic errors-in-variables models," *Automatica*, vol. 43, no. 4, pp. 621-630, 2007.

[11] J. Schoukens and R. Pintelon, *Identification of Linear Systems: A Practical Guideline to Accurate Modeling*. Oxford, U.K.: Pergamon Press, 1991.

[12] T. Söderström, "Identification of stochastic linear systems in presence of input noise," *Automatica*, vol. 17, no. 5, pp. 713-725, 1981.

[13] T. Söderström, "Errors-in-variables methods in system identification," *Automatica*, vol. 43, no. 6, pp. 939-958, 2007.

[14] T. Söderström, M. Hong and W. X. Zheng, "Convergence properties of bias-eliminating algorithms for errors-in-variables identification," *International Journal of Adaptive Control and Signal Processing*, vol. 19, no. 9, pp. 703-722, 2005.

[15] T. Söderström, U. Soverini and K. Mahataa, "Perspectives on errors-in-variables estimation for dynamic systems," *Signal Processing*, vol. 82, no. 8, pp. 1139-1154, 2002.

[16] P. Stoica, M. Cedervall and A. Eriksson, "Combined instrumental variable and subspace fitting approach to parameter estimation of noisy input-output systems," *IEEE Transactions on Signal Processing*, vol. 43, no. 10, pp. 2386-2397, 1995.

[17] S. Van Huffel, Ed., *Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling*. Philadelphia, PA: SIAM Press, 1997.

[18] S. Van Huffel and P. Lemmerling, Eds, *Total Least Squares and Errors-in-Variables Modeling: Analysis, Algorithms and Applications*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

[19] W. X. Zheng, "Transfer function estimation from noisy input and output data," *International Journal of Adaptive Control and Signal Processing*, vol. 12, no. 4, pp. 365-380, 1998.

[20] W. X. Zheng, "On least-squares identification of stochastic linear systems with noisy input-output data," *International Journal of Adaptive Control and Signal Processing*, vol. 13, no. 3, pp. 131-143, 1999.