RM-75-38

IDENTIFICATION AND ESTIMATION OF THE ERROR-IN-VARIABLES MODEL (EVM) IN STRUCTURAL FORM

R.K. Mehra

July 1975

Identification and Estimation of the Error-In-Variables Model (EVM) in Structural Form*

R.K. Mehra**

Abstract

It is shown that the EVM in structural form is identifiable if serial correlation is present in the independent variables. Least Squares, Instrumental Variable and Maximum Likelihood techniques for the identification and estimation of serial correlations and other EVM parameters are given. The techniques used are based on State Vector Models, Kalman Filtering and Innovation representations. Generalizations to EVM involving multiple regressions and randomly time-varying coefficients are also discussed.

Introduction

The Error-In-Variables Model (EVM) is a regression model in which the independent variables are only measured with errors. It has been investigated extensively in the statistical and the econometric literature for over thirty years ([1-10]; for further references see the recent paper by Florens et al. [10]). However, as some of these authors point out, the proposed

---

*This paper was prepared for presentation at the Symposium on Stochastic Systems, University of Kentucky, Lexington, Kentucky, June 10-14, 1975. The work reported here was made possible through a grant from IIASA and through US Joint Services Contract No. N00014-67-A-0298-0006 to the Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts.

**International Institute of Applied Systems Analysis, Laxenburg, Austria, and Harvard University, Cambridge, Massachusetts, USA.

solutions to the problem are still far from satisfactory and require specification of data that may not be readily available in practice.

The two basic assumptions of EVM in Structural Form are normality and independence (or serial uncorrelatedness) of the explanatory variables. The consequences of relaxing normality were analyzed by Reiersøl [11] who showed that the EVM is identifiable for non-normal distributions. In this paper, we examine the assumption of independence and show that relaxation of this assumption makes the EVM identifiable for normal distributions. This may seem paradoxical at first sight, since one is introducing extra correlation parameters into the problem which may be expected to worsen the identifiability problem. But we show, in this paper, that the introduction of a correlation structure between the independent variables gives extra equations that allow one to identify all the parameters associated with the independent variables without using EVM. This solves the basic identifiability problem of EVM in structural form.

The organization of the paper is as follows. In Section 2, we outline the identifiability problem of EVM. The estimation of the covariance and correlation parameters associated with the independent variable using a first order correlation model is discussed in Section 3. The complete identification and estimation of the EVM model is discussed in Sections 4 and 5. Extensions of the EVM to multiple regression and to randomly time-varying coefficients is discussed in Section 6.

2. Error-In-Variables Model, Structural Form

Consider a simplified linear regression model [8],

$$y_i = \alpha + \beta x_i + u_i \quad , \qquad i = 1,\dots,N \quad , \tag{1}$$

where α and β are constant unknown parameters, x_i and y_i are respectively the independent and the dependent variables and u_i is an error variable, normally distributed, white, zero mean and variance σ_u^2. The variable x_i is measured with error

$$z_i = x_i + v_i \tag{2}$$

where v_i is normally distributed, white, zero mean, and variance σ_v^2. We assume that v_i and u_j are independent for all i, j. Notice that u_i includes both the model error in (1) and the measurement error in y_i. In the structural form of EVM, it is further assumed that x_i, $i = 1,\dots,N$ are independent (of each other and of v_i and u_i) and normally distributed with mean μ and variance σ_x^2.

The unknown parameters in the above model are α, β, σ_v^2, σ_u^2, μ and σ_x^2. It is intuitively clear that the maximum likelihood estimates of these six parameters, if they exist, can be obtained by equating the sample mean and covariance of the pair (y_i, z_i) to their theoretical values, i.e.

$$E(z_i) = \mu \tag{3}$$

$$E(y_i) = \alpha + \beta\mu \tag{4}$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_v^2 \tag{5}$$

$$\sigma_y^2 = \beta^2\sigma_x^2 + \sigma_u^2 \tag{6}$$

and

$$\sigma_{zy} = \beta\sigma_x^2 \ . \tag{7}$$

The five equations (3)-(7) can be solved for only five of the six unknown parameters, thus creating an identifiability problem.* It has been suggested in the literature [1-10] that either σ_v^2 or the ratio σ_u^2/σ_v^2 should be assumed to resolve the identifiability problem.** However, such information is generally not available in practice and it would be desirable to devise alternative techniques. An extensive analysis by Zellner [8] shows that the use of Bayesian techniques incorporating a priori information in a less rigid form than the exact specification of σ_v^2 or σ_u^2/σ_v^2 is possible, but the effect of the prior information remains strong for all sample sizes, as pointed out by Florens et al. [10]. Also a recent robustness study by Brown [32] reveals the extreme sensitivity

---

*In terms of the likelihood function, this implies that no maximum exists in the admissible range of the parameters [8].

**Let $\lambda = \sigma_u^2/\sigma_v^2$ and solve Equations (5) and (6) for σ_x^2. Then using Equation (7), the following quadratic equation for β is obtained:

$$\beta = \sigma_{zy} \frac{\beta^2 - \lambda}{\sigma_y^2 - \lambda\sigma_z^2} \ .$$

of the classical estimator to the assumed value of σ_u^2/σ_v^2. In fact, if the assumed value of σ_u^2/σ_v^2 is in error by more than 25%, the ordinary least-squares estimator, even though biased, turns out to have a lower mean square error compared with the classical EVM estimator.

Since the estimation and identification of the EVM seems so out of proportion with its simplicity, one is inclined to ask the question: Is there something missing in the model? Clearly, any model is an idealization of reality and one should make sure that the simplifying assumptions do not make the model degenerate. In the next section, we examine critically the assumption of independence of x_i's and show that a relaxation of this assumption makes EVM identifiable. In most of the practical applications, some form of correlation either already exists or may be caused to exist between the independent variables, so that the above assumption is useful not only from a mathematical standpoint but also beneficial from a practical standpoint.

3. EVM with Correlated Independent Variables

In this section we analyze a particular correlation structure having a Gauss-Markov or state-vector representation [12]. This structure has been used for Time-Series Analysis and System Identification with great success [13,14,15]. In some applications of EVM, the assumption of this type of correlation structure may not be completely valid and one may use some other structure more suited to the particular application. However, for those applications where the independent variables come from time series (e.g. in forecasting problems) and for illustrative purposes, we consider the following first order Gauss-Markov model for the independent variables x_i. (A more general model will be considered in the next section.)

$$x_{i+1} = \phi x_i + w_i \tag{8}$$

$$z_i = x_i + v_i \quad , \tag{9}$$

where $0 < |\phi| < 1$ and w_i is a sequence of zero mean* Gaussian uncorrelated variables with variance σ_w^2. We have excluded the cases $\phi = 0$ and $|\phi| \geq 1$ since the former leads to EVM with uncorrelated x_i's and the latter leads to a nonstationary sequence. The steady state or stationary covariance of (8) satisfies [13],

$$\sigma_x^2 = \phi^2 \sigma_x^2 + \sigma_w^2$$

or

$$\sigma_x^2 = \sigma_w^2 / (1-\phi^2) \quad . \tag{10}$$

If we choose x_0 to be normally distributed with zero mean and variance σ_x^2, then the sequences (x_i, z_i), $i = 1,2,\ldots$, generated by Equations (8)-(9) are stationary. Now we estimate ϕ, σ_w^2 and σ_v^2 from the sample correlation of the observed sequence z_i, $i = 1,2,\ldots$.

---

*For simplicity, we have assumed $E(x_i) = \mu = 0$. In the general case, one should take $E(w_i) = \mu(1-\phi)$.

Let

$$c(k) = E(z_i z_{i+k}) \quad , \qquad k = 0,1,2,3,\ldots.$$

A consistent estimator of c(k) is $\hat{c}(k)$ where

$$\hat{c}(k) = \frac{1}{N} \sum_{i=1}^{N-k} z_i z_{i+k} \quad . \tag{11}$$

The correlation sequence c(k) satisfies [13]

$$c(0) = \sigma_z^2 = \sigma_x^2 + \sigma_v^2 \tag{12}$$

$$c(1) = \phi\sigma_x^2 \tag{13}$$

$$c(2) = \phi^2\sigma_x^2 \quad . \tag{14}$$

In general,

$$c(k) = \phi^k\sigma_x^2 \quad , \qquad k = 1,2,3,\ldots. \tag{15}$$

Equations (13) and (14) may be solved for ϕ and σ_x^2:

$$\phi = \frac{c(2)}{c(1)} \tag{16}$$

$$\sigma_x^2 = \frac{c^2(1)}{c(2)} \quad . \tag{17}$$

From equation (12),

$$\sigma_v^2 = c(0) - \frac{c^2(1)}{c(2)} \quad , \tag{18}$$

and from equation (10),

$$\sigma_w^2 = \left(1 - \frac{c^2(2)}{c^2(1)}\right)\frac{c^2(1)}{c(2)} = \frac{1}{c(2)}\left(c^2(1) - c^2(2)\right). \quad (19)$$

It is easily shown that if $\hat{c}(k)$ is used for $c(k)$ in Equations (16)-(19), the corresponding estimates of ϕ, σ_x^2, σ_w^2 are consistent [13]. Using these estimates in Equations (6)-(7), one can obtain consistent estimates of β and σ_u^2. The estimation of α and μ is done from the sample means y_i and z_i using Equations (3) and (4). Thus the EVM with the correlation structure of equation (8) and $\phi \neq 0$ is identifiable. Of course, the accuracy of estimates would depend on ϕ, with smaller values of ϕ tending to give larger standard deviations of the parameter estimates. In the limit as $\phi \to 0$, equations (13)-(15) do not provide any information about σ_x^2 and one has the problem of determining both σ_x^2 and σ_v^2 from equation (12) alone. This gives rise to the identifiability problem of the classical EVM. Theoretically, thus, the EVM is identifiable for nonzero ϕ, however small. Furthermore, if correlation is present, its inclusion in the model would, in general, improve the results.

In the next section, we consider estimation of the EVM with a more general Gauss-Markov correlation structure. From here on, we assume that the independent variable has some correlation. Unless there are strong physical reasons to believe that the independent variable is completely uncorrelated, the above assumption is justified in practice. The procedure to be outlined in the next section may, in fact, be used to test correlatedness.

4. Consistent Estimation and Identification of the Correlated EVM

Since the details of some of the techniques to be described here are also covered elsewhere [13,14,15] we will only sketch these techniques here. The new or special aspects of the EVM will be described in detail.

Consider again the EVM, equations (1)-(2) with scalar x_i. We now generalize the first order correlation structure of equation (8) to an n^{th} order correlation structure using a state-vector model, i.e.

$$s_{i+1} = \Phi s_i + \Gamma w_i \tag{20}$$

$$x_i = h s_i \tag{21}$$

$$z_i = h s_i + v_i \quad , \tag{22}$$

where s_i is $n\times 1$ state vector; $\Phi(n\times n)$, $\Gamma(n\times 1)$ and $h(1\times n)$ are respectively a constant matrix and vectors with unknown parameters. By a basis change, the matrices Φ, Γ and h can always be put into the following canonical form [13,14,15]:

$$\Phi = \begin{bmatrix} 0 & 1 & 0 & - & - & 0 \\ 0 & 0 & 1 & - & - & 0 \\ & & \vdots & & & \\ 0 & 0 & - & - & - & 1 \\ -\phi_1 & -\phi_2 & - & - & - & -\phi_n \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \gamma_1 \\ \cdot \\ \vdots \\ \cdot \\ \gamma_n \end{bmatrix}, \quad h = [1,0,\ldots,0] \quad .$$

Furthermore, w_i can be taken to be of unit variance so that the model (20)-(22) has a total of (2n+1) parameters. Generalizations of the technique discussed in Section 3 to this case are given in Reference [13]. The relevant equations are (23)-(25) below.

$$\begin{bmatrix} \phi_1 \\ \cdot \\ \cdot \\ \phi_n \end{bmatrix} = - \begin{bmatrix} c(1) & \cdots & c(n) \\ \cdot & & \\ \cdot & & \\ c(n) & & c(2n-1) \end{bmatrix}^{-1} \begin{bmatrix} c(n+1) \\ \cdot \\ \cdot \\ c(2n) \end{bmatrix} \qquad (23)$$

where c(k)'s are estimated by Equation (11);

$$\sigma_v^2 = \frac{1}{\phi_1} \sum_{j=0}^{n} \phi_{j+1} c(j) \quad , \qquad \phi_{n+1} = 1 \qquad (24)$$

$$\sigma_x^2 = h\Phi^{-1} \begin{bmatrix} c(1) \\ \cdot \\ \cdot \\ c(n) \end{bmatrix} \quad . \qquad (25)$$

It should be montioned that a model equivalent to Equations (20)-(22) is the following 'Innovation' or 'Kalman Filter' model [13,14,15],

$$\hat{s}_{i+1|i} = \Phi \left[\hat{s}_{i|i-1} + k\nu_i \right] \qquad (26)$$

$$z_i = h\hat{s}_{i|i-1} + \nu_i \quad , \qquad (27)$$

where $\hat{s}_{i+1|i}$ denotes the conditional mean estimate of s_{i+1}

given $\{z_1,\ldots,z_i\}$, and ν_i denotes the sequence of one-step-ahead prediction errors or innovations [13], since from Equation (27)

$$\nu_i = z_i - \hat{z}_{i|i-1} \quad . \tag{28}$$

It is known that [13] ν_i is a zero mean Gaussian white noise sequence with variance $\sigma_\nu^2 = \sigma_v^2/(1-hk)$. The Kalman gain $k(n\times 1)$ is a constant vector of gains related to σ_v^2 and σ_w^2 (cf. Equations (31)-(32)).

The interesting property of the model (26)-(27), besides the whiteness of the sequence ν_i, is the uncorrelatedness of $\hat{z}_{i|i-1}$ with v_i since $\hat{z}_{i|i-1}$ is a function of $\{z_1,\ldots,z_{i-1}\}$ only. This fact is useful in constructing an instrumental variable [16] for estimating β as follows.

Consider EVM (1)-(2) combined into a single equation

$$y_i = \alpha + \beta z_i - \beta v_i + u_i \quad . \tag{29}$$

Using $\hat{z}_{i|i-1}$ as instrumental variable (IV) [16],

$$E\left(y_i \hat{z}_{i|i-1}\right) = \beta E\left(z_i \hat{z}_{i|i-1}\right)$$

or

$$\beta = \frac{E\left(y_i \hat{z}_{i|i-1}\right)}{E\left(z_i \hat{z}_{i|i-1}\right)} \quad . \tag{30}$$

Equation (3) can be used to obtain a consistent IV estimator for β by replacing the theoretical correlations by their sample values. The sequence $\hat{z}_{i|i-1} = h\hat{s}_{i|i-1}$ is generated using Equations (26)-(27). The matrix Φ is estimated from the correlations of z_i's using equation (23), and k is obtained from the estimates of Γ and σ_v^2, as follows.*

$$k = Mh^T\left(hMh^T + \sigma_v^2\right)^{-1} \quad (31)$$

where

$$M = \Phi\left[M - Mh^T\left(hMh^T + \sigma_v^2\right)^{-1}hM\right]\Phi^T + \Gamma\Gamma^T \quad . \quad (32)$$

Other methods for direct and more efficient estimation of k exist and are described in References [13,14,15]. In practice, however, the Maximum Likelihood method seems to give the best results, and it may be used for the simultaneous estimation of all the parameters, denoted collectively as $\Theta = \left[\beta, \sigma_u, \sigma_v, \phi_1, \ldots, \phi_n, \gamma_1, \ldots, \gamma_n\right]^T$.

In the next section, we describe a Maximum Likelihood (ML) Estimator, keeping in mind that the above correlation procedure is to be used to obtain a consistent estimator $\hat{\Theta}_0$ which will

---

*To maintain the uncorrelatedness of $\hat{z}_{i|i-1}$ with $\{z_i, z_{i+1}, z_{i+2}, \ldots\}$, the estimates of Φ and k used in the Kalman Filter are based on the past data, viz. $\{z_{i-1}, z_{i-2}, \ldots\}$. These estimates are computed on-line by using a recursive form of Equation (23) [13].

be required to start the ML estimation iterative procedure. But first we discuss the problem of determining the order n of the system.

4.1 <u>Order Determination</u>: The state vector model (26)-(27) along with canonical forms for Φ and H may be written in input-output form as [13,14]

$$z_{i+n} + \sum_{j=1}^{n} \phi_j z_{i+j-1} = \nu_{i+n} + \sum_{j=1}^{n} c_j \nu_{i+j-1} \quad . \qquad (33)$$

Equations (26)-(27) and Equation (33) are related by their transfer functions, viz.

$$h(qI - \Phi)^{-1}\Phi k + I = \left(q^n + \sum_{j=1}^{n} \phi_j q^{j-1}\right)^{-1}\left(q^n + \sum_{j=1}^{n} c_j q^{j-1}\right) , \qquad (34)$$

where q is a forward shift operator, i.e.

$$q\hat{s}_{i|i-1} = \hat{s}_{i+1|i} \quad \text{and} \quad qz_i = z_{i+1} \quad . \qquad (35)$$

Equation (33) is an Autoregressive Moving Average (ARMA) model of order (n,n) [17]. Let us successively multiply and take expectations on both sides of Equation (33) by $\hat{z}_{i|i-1}$, $\hat{z}_{i+1|i-1}, \ldots, \hat{z}_{i+n|i-1}$ where $\hat{z}_{i+j|i-1} = E\{z_{i+j}|z_1,\ldots,z_{i-1}\}$ is a function of $(z_1,\ldots,z_{i-1})$ only. Then since $E\{\nu_{i+j}z_k\} = 0$ for $k \leq i-1$ and for $j \geq 0$, we get

$$E\left(z_{i+n}\hat{z}_{i+k|i-1} + \sum_{j=1}^{n} \phi_j \hat{z}_{i+k|i-1} z_{i+j-1}\right) = 0 \quad , \qquad (36)$$

$$k = 0,1,\ldots,n \; .$$

This may be written in matrix form as

$$E \underbrace{\begin{bmatrix} \hat{z}_{i|i-1} z_i, & \cdots & , \hat{z}_{i|i-1} z_{i+n} \\ \hat{z}_{i+1|i-1} z_i, & \ldots, & \hat{z}_{i+1|i-1} z_{i+n} \\ & \vdots & \\ \hat{z}_{i+n|i-1} z_i, & \ldots, & \hat{z}_{i+n|i-1} z_{i+n} \end{bmatrix}}_{C_{\hat{z}\hat{z}}} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \\ 1 \end{bmatrix} = 0 \; . \qquad (37)$$

Using the linearity property of Gaussian conditional expectations, we can write

$$\hat{z}_{i+\ell|i-1} = \sum_{t=1}^{i-1} \alpha_{t,\ell} z_t \; . \qquad (38)$$

For practical purposes, one approximates Equation (38) by

$$\hat{z}^m_{i+\ell|i-1} = \sum_{t=i-m}^{i-1} \alpha^m_{t,\ell} z_t \quad , \qquad (39)$$

where $m < i$ is chosen as the least integer value that essentially achieves the minimum prediction error. From Equation (39)

$$E\left(z_{i+j}\hat{z}^m_{i+\ell|i-1}\right) = \sum_{t=i-m}^{i-1} \alpha^m_{t,\ell} C(i+j-t) \; . \qquad (40)$$

Also from equation (39) and the orthogonality property, $\alpha_{t,k}$ satisfy the equations

$$C(i+\ell-\tau) = \sum_{t=i-m}^{i-1} \alpha_{t,\ell}^{m} \, C(t-\tau) \quad , \quad \tau = i-m,\ldots,i-1 \quad , \quad \ell = 1,\ldots,n \quad . \qquad (41)$$

Equation (37) shows that if the order of the system is n, then the matrix of correlations $C_{\hat{z}\hat{z}}$ has rank (n-1) and the eigenvector corresponding to the zero eigenvalue is $[\phi_1,\ldots,\phi_n,1]$. An estimate of $C_{\hat{z}\hat{z}}$ may be obtained by computing sample correlations $\hat{C}(\ell)$, $\ell = 0,\ldots,m$ (of Equation (11)) and by solving Equation (41) using an efficient recursive algorithm based on the work of Levinson [18], Durbin [19] and Wiggins and Robinson [20]. The algorithm can be made recursive both in the order m of lags and and in the lead variable $j \geq 0$ as shown in Reference [21]. Notice that only the smallest eigenvalue of $C_{\hat{z}\hat{z}}$ needs to be computed for different values of n to decide on the model order.

Another procedure proposed by Akaike [22] is to use the method of canonical correlations between the sets of variables $\xi = \left\{\hat{z}_{i|i-1},\ldots,z_{i+n \; i-1}\right\}$ and $\eta = \{z_{i-1},\ldots,z_{i-m}\}$ for m sufficiently large. In this method, correlations between all normalized linear combinations of ξ and η viz $A\xi$ and $B\eta$, with $||A\xi|| = ||B\eta|| = 1$ are checked, and the combinations with the least correlation are tested for uncorrelatedness. In essence, a Singular Value Decomposition [23] of the matrix $E\{\xi\eta^T\}$ is performed and the lowest characteristic value is checked for significance. The test has been found useful in practical problems, but seems to involve more computation than the method proposed above.

Remark: 1. Akaike [22] has shown that the state vector of the system may be defined as

$$\hat{s}_{i|i-1} = \begin{bmatrix} \hat{z}_{i|i-1} \\ \vdots \\ \hat{z}_{i+n-1|i-1} \end{bmatrix} .$$

Then $\hat{s}_{i|i-1}$ represents all the information from the past needed to predict the future outputs of the system. Thus for an nth order, $\hat{z}_{i+n|i-1}$ will be linearly dependent on $\hat{s}_{i|i-1}$, which also follows from Equation (37).

2. In deriving Equation (37) from Equation (33), $\left(\hat{z}_{i+k|i-1},\ k = 0,\ldots,n\right)$ were used as instrumental variables. If one uses, instead, lagged values of z, viz. $(z_{i-1}, z_{i-2}, \ldots)$ as instruments, Modified Yule-Walker Equations are obtained [13]. The advantage of using $\hat{z}_{i+k|i-1}$'s as instruments is an improvement in efficiency of estimating $(\phi_1,\ldots,\phi_n)$ since the resulting equations have a structure similar to the Maximum Likelihood estimator discussed below. It is important for order determination that the estimates of $(\phi_1,\ldots,\phi_n)$ be as efficient as possible within the constraints of the computation burden.

5. Maximum Likelihood Estimation of the Correlated EVM

The EVM described by Equations (1) and (20)-(22) may be written in state-vector form as

$$s_{i+1} = \Phi s_i + \Gamma w_i \tag{20}$$

$$y_i = \beta h s_i + u_i \tag{23*}$$

$$z_i = h s_i + v_i \tag{22}$$

$$i = 1,\ldots,N \quad .$$

Equations (22) and (23) may be combined into a vector equation

$$m_i = H s_i + n_i \quad , \tag{24}$$

where $m_i = \begin{pmatrix} y_i \\ z_i \end{pmatrix}$ is 2 × 1 vector of measurements, $n_i = \begin{pmatrix} u_i \\ v_i \end{pmatrix}$ is 2 × 1 vector of noises with covariance matrix

$$R = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix} \quad , \tag{25}$$

and H is 2 × n matrix defined as

$$H = \begin{bmatrix} \beta \\ 1 \end{bmatrix} h \quad . \tag{26}$$

Let $\theta((2n + 3) \times 1)$ be the vector of all unknown parameters in the above model, i.e.

$$\theta = [\beta,\sigma_u,\sigma_v,\phi_1,\ldots,\phi_n,\gamma_1,\ldots,\gamma_n]^T \quad . \tag{27}$$

---

*It is assumed at $\alpha = 0$ or equivalently the mean value of y_i has been subtracted already.

We have shown in Section 4 that θ is identifiable as long as $\theta_1,\ldots,\theta_n$ are not identically zero and the roots of θ lie inside the unit circle. Thus the maximum likelihood estimate of θ would be expected to be consistent. In fact, stronger results can be proved, viz. that under the above conditions, the MLE is asympotically unbiased, efficient, normal and strongly consistent [24,25]. We describe here only the procedure for obtaining MLE of θ. For further computational details the reader may refer to [26,27].

The log-likelihood function is

$$\begin{aligned} L(\theta) &= \log p(y_1,\ldots,y_N,\ z_1,\ldots,z_N|\theta) \\ &= \log p(m_1,\ldots,m_N|\theta) \\ &= \sum_{j=1}^{N} \log p(m_j|m_1,\ldots,m_{j-1},\theta) \quad . \end{aligned} \tag{28}$$

The conditional density $p(m_j|m_1,\ldots,m_{j-1},\theta)$ is normal with mean and covariance denoted respectively by* $\hat{\hat{m}}_{j|j-1}$ and $\hat{\hat{P}}_{j|j-1}$. As is well known, these quantities can be computed recursively using a Kalman Filter [12] of the following form:

$$\hat{\hat{s}}_{j+1|j} = \Phi\hat{\hat{s}}_{j|j} \quad . \tag{29}$$

$$\hat{\hat{s}}_{j|j} = \hat{\hat{s}}_{j|j-1} + K_j(m_j - H\hat{\hat{s}}_{j|j-1}) \tag{30}$$

$$P_{j+1|j} = \Phi P_{j|j}\,\Phi^T + \Gamma\Gamma^T \tag{31}$$

---

*Double hats are used for estimates conditional on the joint set of measurements $\{m_1,\ldots,m_{j-1}\}$. Thus $\hat{\hat{s}}_{j|j-1} = E\{s_j|m_1,\ldots,m_{j-1},\theta\} = E\{\hat{s}_{j|j-1}|y_1,\ldots,y_{j-1},\theta\}$.

$$K_j = P_{j|j-1} H^T (H P_{j|j-1} H^T + R)^{-1} \qquad (32)$$

$$P_{j|j} = (I - K_j H) P_{j|j-1} \quad . \qquad (33)$$

The initial conditions are specified from a priori knowledge as $\hat{s}_{0|0} = s_0$ and $P_{0|0} = P_0$. If no a priori information is available, one may use the Information Form of the Kalman Filter [28] that propagates $P^{-1}_{j|j-1}$ and $P^{-1}_{j|j}$ starting from an initial value of zero. Another method often used in practice is to set P_0 to a very large value which essentially eliminates the dependence of the Kalman filter on initial values.

The log-likelihood function (28) may now be written as

$$L(\theta) = -\sum_{j=1}^{N} \tfrac{1}{2}\{(m_j - \hat{m}_{j|j-1})^T (HP_{j|j-1}H^T + R)^{-1} (m_j - \hat{m}_{j|j-1})$$

$$+ \log|HP_{j|j-1}H^T + R|\} + \text{constants} \quad . \qquad (34)$$

We now maximize $L(\theta)$ with respect to $\theta \varepsilon \Theta$ subject to the constraints of Eqs.(29)-(33). Since this is a nonlinear programming problem, a commonly used method is modified Gauss-Newton, the details of which are given in [27]. The basic iteration is

$$\theta^{j+1} = \theta^j + \rho M^{\#}(\theta^j)\left(\frac{\partial L}{\partial \theta^j}\right) \quad , \qquad (35)$$

where θ^j is the value of θ during the j^{th} iteration, $0 < \rho < 1$ is a step-size parameter, $M(\theta^j)$ is an estimate of the Fisher Information Matrix at θ^j defined as

$$M(\theta^j) = E\left\{\frac{\partial^2 L}{\partial\theta^2}\right\}\Bigg|_{\theta^j} = E\left\{\left(\frac{\partial L}{\partial\theta}\right)\left(\frac{\partial L}{\partial\theta}\right)^T\right\}_{\theta^j} ,$$

and $M^{\#}$ is a modified inverse of M obtained by setting eigenvalues of M below a certain threshold (such as 10^{-8} times the largest eigenvalue) to larger positive values. Most of the computation in this algorithm is involved in calculating the partial derivatives $\frac{\partial \hat{\hat{s}}}{\partial\theta}_{j|j-1}$ and $\frac{\partial P}{\partial\theta}_{j|j-1}$ from a set of linear recursive equations. As indicated in Ref. [26,27], simplifications to these computations are possible in practice. Notice that $M^{-1}(\hat{\theta})$ evaluated at the MLE $\hat{\theta}$ is the Cramer-Rao lower bound. For large samples, it gives a good estimate of the covariance of the ML estimates.

Remarks: 1. Since the log-likelihood function may be multimodal, it is important to have a good starting estimate $\theta_0 \varepsilon \Theta$ of the parameters. The Innovation Correlation - Instrumental Variable technique described in Section 4 is recommended for this purpose. In the small sample case, even the order n may have to be rechecked using MLE along with an F-test or Information criterion [22].

2. It is also possible to develop a maximum likelihood estimator using the alternative model (26)-(27). This formulation leads to some simplifications and is also useful in the Multiple Regression case to be discussed in the next section, since in that case, a multivariate version of (26)-(27) is identified directly. The new set of equations is

$$\hat{s}_{i+1|i} = \Phi[\hat{s}_{i|i-1} + k(\underbrace{z_i - h\,\hat{s}_{i|i-1}}_{\nu_i})] \tag{20'}$$

$$y_i = \beta h\,\hat{s}_{i|i-1} + \varepsilon_i \quad , \tag{21'}$$

where $$\varepsilon_i = \beta(\nu_i - v_i) + u_i \tag{22'}$$

and $$\sigma_\varepsilon^2 = \beta^2\sigma_\nu^2\, h\, k + \sigma_u^2 \quad . \tag{23'}$$

Defining

$$\theta' = [\beta,\sigma_u,\sigma_\nu,\phi_1,\ldots,\phi_n,\ k_1,\ldots,k_n] \quad , \tag{24'}$$

the log-likelihood function $L(\theta')$ may be written in terms of the Kalman Filter,*

$$\hat{\hat{s}}_{i+1|i} = \Phi[\hat{\hat{s}}_{i|i-1} + (I - kh)\ g_i(y_i - \beta h\,\hat{\hat{s}}_{i|i-1})$$

$$+ k(z_i - h\,\hat{\hat{s}}_{i|i-1})] \tag{25'}$$

$$P_{i+1|i} = \Phi(I - kh)(I - \beta g_i h)P_{i|i-1}(I - kh)^T\,\Phi^T \tag{26'}$$

$$g_i = \beta P_{i|i-1}h^T(\beta^2\, h\, P_{i|i-1}\, h^T + \beta^2\sigma_\nu^2\, hk + \sigma_u^2)^{-1} \quad . \tag{27'}$$

$L(\theta')$ has the same form as Eq.(34) except that σ_v^2 in R is replaced by $(I - hk)\,\sigma_\nu^2$ and Eqs. (25')-(27') are used to evaluate $P_{j|j-1}$ and $\hat{\hat{s}}_{j|j-1}$ in terms of θ' parameters. Notice that no matrix inversion is required in Eqs. (26')-(27').

---

*The Kalman Filter for Equations (20')-(21') is derived by regarding Equation (20') as an equation containing no process noise, viz.

$$\hat{s}_{i+1|i} = \Phi(I - kh)\hat{s}_{i|i-1} + kz_i \quad ,$$

where z_i is a known sequence.

6. Extensions

In this section, we consider two extensions:

(i) vector β and vector x case (Multiple Regression EVM), and

(ii) randomly time-varying β.

6.1 Multiple Regression EVM

Let $\underline{x}$ denote a p × 1 vector of independent variables and $\underline{\beta}$ be the corresponding vector of regression coefficients.*
The EVM is

$$y_i = \underline{x}_i^T \underline{\beta} + u_i \tag{36}$$

$$\underline{z}_i = \underline{x}_i + \underline{v}_i \quad , \qquad i = 1,\dots,N \quad . \tag{37}$$

We now develop a state-vector model for the series $\{\underline{x}_i\}$, of the same form as Eqs. (20)-(22) except that $\underline{h}$ is a matrix (p × n). The identification of this model is more complicated, but follows the same basic principles as outlined in Sections 4 and 5. The essential differences lie in choosing a canonical form which in the multi-output case depends on p integers $\{n_1,\dots,n_p\}$ such that $\sum_{i=1}^{p} n_i = n$. The state vector $\hat{s}_{i|i-1}$ is defined as

$$\underset{(n\times 1)}{\hat{s}_{i|i-1}} = \begin{bmatrix} \hat{z}_{i|i-1}(1) \\ \vdots \\ \hat{z}_{i+n_1|i-1}(1) \\ \hat{z}_{i|i-1}(2) \\ \vdots \\ \hat{z}_{i+n_2|i-1}(2) \\ \vdots \\ \hat{z}_{i+n_p|i-1}(p) \end{bmatrix} \tag{38}$$

---

*A bar under a scalar variable denotes a vector and a bar under a vector denotes a matrix. A bar under a matrix denotes another matrix of different dimensions.

where $\hat{z}_{i+j|i-1}(k)$ denotes the (j+1)-step-ahead predicted estimate of the kth component of z_{i+j}. The integers $n_1, n_2, \ldots, n_p$ are determined by examining corelations between the above variables in the order $\hat{z}_{i|i-1}(1)$, $\hat{z}_{i|i-1}(2), \ldots,$ $\hat{z}_{i|i-1}(p)$, $\hat{z}_{i+1|i-1}(1)$, $\hat{z}_{i+1|i-1}(2), \ldots, \hat{z}_{i+n_j+1|i-1}(j)$, where j refers to the output variable with the highest value n_j. Thus, n_1 is determined when $z_{i+n_1|i-1}(1)$ becomes linearly correlated to its antecedents. The procedure is quite straightforward and is well described in Ref. [22]. The procedure described in Section 4 using Eqs. (36)-(41) is also easily extended to the vector case using recursive algorithms of Wiggins-Robinson [20,21]. We now summarize the complete procedure adding a few more practical details.

1. Compute the sample correlation matrices $\hat{\underline{c}}(k)$, of $\{\underline{z}_1, \ldots, \underline{z}_n\}$ after subtracting the mean, for k up to $m \approx N/_{10}$.*
2. Determine a state-vector model for $\underline{x}_i$'s using either the Canonical Correlation Procedure of Akaike [22] or the procedure of Section 4 extended to the vector case [21]. During this step, the order n, output numbers $n_1, \ldots, n_p$ and matrices $(\Phi, \underline{k}, \underline{h})$ in canonical form are determined:

$$\hat{s}_{i+1|i} = \Phi[\hat{s}_{i|i-1} + \underline{k}\ \underline{\nu}_i] \tag{39}$$

$$\underline{\nu}_i = \underline{z}_i - \underline{h}\ \hat{s}_{i|i-1} \ . \tag{40}$$

3. This step should be performed if, due to small sample size, the procedure of step 2 is expected to yield inefficient estimates that may also affect the correct

---

*$N/_{10}$ is an empirical number beyond which the accuracy of correlations is found to degrade seriously.

determination of $(n_1, \ldots, n_p)$ [22]. During this step, obtain maximum likelihood estimates of parameters in Φ, $\underline{k}$, $\underline{h}$ and $\sum_{\underline{\nu}\underline{\nu}}$ (covariance matrix of $\underline{\nu}$'s), denoted collectively by vector ψ, by maximizing the log-likelihood function,

$$L(\psi) = -\tfrac{1}{2} \sum_{i=1}^{N} \underline{\nu}_i^T \sum_{\underline{\nu}\underline{\nu}}^{-1} \underline{\nu}_i - \frac{N}{2} \ln\left|\sum_{\underline{\nu}\underline{\nu}}\right| \qquad (41)$$

4. Use $\hat{\underline{x}}_{i|i-1} = h\hat{s}_{i|i-1}$ as instrumental variables with Eqs. (36) and (37) to obtain a consistent estimate of $\underline{\beta}$:

$$\hat{\underline{\beta}} = \left[\sum_{i=1}^{N} (\hat{\underline{x}}_{i|i-1} \underline{z}_i^T)\right]^{-1} \left(\sum_{i=1}^{N} \hat{\underline{x}}_{i|i-1} y_i\right) . \qquad (42)$$

From the sum-of-squares of the residuals $(y_i - z_i^T \hat{\underline{\beta}})$, obtain an estimate of σ_u^2 using estimated values of $\underline{\beta}$ and $\sum_{\underline{v}\underline{v}} = (I - \underline{h}\underline{k}) \sum_{\underline{\nu}\underline{\nu}}$, covariance of measurement noise $\underline{v}_i$.

5. Steps 1-4 give the model structure and a consistent estimate of all the unknown parameters, $\theta \in \{\beta, \sigma_u, \sum_{\underline{\nu}\underline{\nu}}, \Phi, \underline{k}, \underline{h}\}$. We now perform final maximum likelihood estimation by maximizing with respect to $\theta \in \Theta$,

$$L(\theta) = -\tfrac{1}{2} \sum_{i=1} \left\{ [\zeta_i^T, \underline{r}_i^T] (H\, P_{i|i-1} \underline{H}^T + \sum_{\underline{v}\underline{v}})^{-1} \begin{bmatrix} \zeta_i \\ r_i \end{bmatrix} \right.$$

$$\left. + \ln\left|\underline{H}\, P_{i|i-1} \underline{H}^T + \sum_{\underline{v}\underline{v}}\right| \right\} ,$$

subject to the constraints

$$\hat{\hat{s}}_{i+1|i} = \Phi[\hat{\hat{s}}_{i|i-1} + (I - \underline{k}\underline{h})g_i \zeta_i + \underline{k}\ \underline{r}_i] \tag{43}$$

$$\zeta_i = y_i - \underline{\beta}^T\underline{h}\ \hat{\hat{s}}_{i|i-1} \tag{44}$$

$$\underline{r}_i = \underline{z}_i - \underline{h}\ \hat{\hat{s}}_{i|i-1} \tag{45}$$

$$P_{i+1|i} = \Phi(I - \underline{k}\underline{h})(I - \underline{\beta}^T g_i\underline{h})P_{i|i-1}(I - \underline{k}\underline{h})^T\Phi^T \tag{46}$$

$$g_i = P_{i|i-1}\underline{h}^T\underline{\beta}(\underline{\beta}^T\underline{h}\ P_{i|i-1}\underline{h}^T\underline{\beta} + \sigma_u^2 + \underline{\beta}^T\underline{h}\underline{k}\ \Sigma_{\underline{v}\underline{v}}\ \underline{\beta})^{-1} \tag{47}$$

$$\Sigma_{\underline{v}\underline{v}} = (I - \underline{h}\underline{k})\ \Sigma_{\underline{v}\underline{v}} \tag{48}$$

$$H = \begin{bmatrix} \underline{\beta}^T \\ I \end{bmatrix} \underline{h}\ . \tag{49}$$

6.2 EVM with Randomly Time-Varying Coefficients

In this section we propose an approximate technique based on Extended Kalman Filtering [28,29] for estimation of EVM with time-varying coefficients. For simplicity, consider the simple EVM (1)-(2) with $\alpha = 0$, and β a function of i, which now explicitly refers to time. (This connection with time is not essential, but helps motivation, since such models generally arise in forecasting applications where i is a time variable.)

$$y_i = \beta_i\ x_i + u_i \tag{50}$$

$$z_i = x_i + v_i\ . \tag{2}$$

One way of modeling random changes in β_i that has been used successfully in practice with ordinary regression models [30] is

$$\beta_{i+1} = \delta\ \beta_i + \Delta_i \tag{51}$$

where $0 \leq \delta < 1$ is a constant unknown parameter and Δ_i is a Gaussian white noise sequence* with unknown variance σ_Δ^2.

Let us assume, again for simplicity, that x_i's obey a first order model (cf Eq. (8)):

$$x_{i+1} = \phi\ x_i + w_i\ . \tag{8}$$

Regarding (x_i, β_i) as the state vector, the above four equations constitute a linear state-vector model with <u>nonlinear</u> measurements since the product of states, $\beta_i x_i$, appears in Eq. (50). Initial estimation of ϕ and σ_v^2 can still be carried out in the same fashion as before, but to estimate β_i's, we use an Extended Kalman Filter of the following form [28],

$$\begin{bmatrix} \hat{\hat{x}}_{i+1|i} \\ \hat{\hat{\beta}}_{i+1|i} \end{bmatrix} = \underbrace{\begin{bmatrix} \phi & 0 \\ 0 & \delta \end{bmatrix}}_{D} \begin{bmatrix} \hat{\hat{x}}_{i|i-1} \\ \hat{\hat{\beta}}_{i|i-1} \end{bmatrix} + K_i \begin{bmatrix} z_i - \hat{\hat{x}}_{i|i-1} \\ y_i - \hat{\hat{\beta}}_{i|i-1}\hat{x}_{i|i-1} \end{bmatrix} \tag{52}$$

$$K_i = S_{i|i-1}\ A_i^T (A_i S_{i|i-1}\ A_i^T + R)^{-1} \tag{53}$$

$$S_{i+1|i} = D(I - K_i A_i) S_{i|i-1} D^T + E\ , \tag{54}$$

---

*As shown in Zellner [8], the assumption of Gaussian prior distribution of β_i may lead to inadmissible values. In such cases, we assume that β_i is a transformed variable with Gaussian density.

where

$$A_i = \begin{bmatrix} 1 & 0 \\ \hat{\hat{\beta}}_{i|i-1} & \hat{\hat{x}}_{i|i-1} \end{bmatrix} \quad (55)$$

$$R = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix} \quad (56)$$

$$D = \begin{bmatrix} \phi & 0 \\ 0 & \delta \end{bmatrix} \quad (57)$$

$$E = \begin{bmatrix} \sigma_w^2 & 0 \\ 0 & \sigma_\Delta^2 \end{bmatrix} . \quad (58)$$

The Extended Kalman Filter is not an optimal nonlinear filter for estimation of β's and x's. Other filters such as the Iterated Sequential - Extended Kalman Filter [29], which involves a little more computation, may give better results. The choice of the filter is dependent qualitatively on the amount of noise and the extent of nonlinearity in the equations.

Approximate Maximum Likelihood estimation of the unknown parameters $\theta = (\sigma_u, \sigma_v, \delta, \sigma_\Delta, \phi, \sigma_w)^T$ may be performed by maximizing the following approximate log-likelihood function subject to Eqs. (52)-(58).

$$L(\theta) = -\tfrac{1}{2} \sum_{i=1}^{N} e_i^T (A_i s_{i|i-1} A_2^T + R)^{-1} e_i + \ln|A_i s_{i|i-1} A_i^T + R| , \tag{59}$$

where

$$e_i = \begin{bmatrix} z_i - \hat{\hat{x}}_{i|i-1} \\ y_i - \hat{\hat{\beta}}_{i|i-1} \hat{\hat{x}}_{i|i-1} \end{bmatrix} . \tag{60}$$

$L(\theta)$ is approximate log-likelihood since e_i is not exactly Gaussian and white. Further details of this procedure may be found in Ref. [31].

Remarks: 1. The above method is extended easily to Multiple Regression EVM with a general correlation model for the independent variables.

2. In many forecasting applications where regression is used, the values of the independent variables also have to be predicted for the future. If the independent variables are regarded as serially uncorrelated, this cannot be done. Our procedure directly gives forecasts of both the independent and dependent variables via the equations

$$\hat{\hat{s}}_{i+j|i} = \Phi^{j-1} \hat{\hat{s}}_{i+1|i} \tag{61}$$

$$\begin{bmatrix} \hat{\hat{y}}_{i+j|i} \\ \hat{\hat{x}}_{i+j|i} \end{bmatrix} = \underline{H}\, \hat{\hat{s}}_{i+j|i} . \tag{62}$$

6. Conclusions

The EVM in structural form is completely identifiable as long as some serial correlation is present in the independent variables. Both least squares and maximum likelihood techniques have been given to identify and estimate the serial correlations and the EVM parameters. Construction of Bayesian techniques is also straightforward and will be discussed elsewhere. The following summarize what are believed to be the original contributions of the paper:

1. The assumption of no serial correlation of independent variables is a cause of the identifiability problem. The assumption is generally not justified in practical applications such as forecasting where regression models are commonly used.

2. Consistent estimates of the correlation parameters are obtained by analyzing the series of independent variables alone. These estimates are then used along with a new instrumental variable technique to obtain consistent estimates of the EVM parameters.

3. In Section 4.1, a computationally efficient technique is given for model order determination.

4. In Sections 5 and 6, a maximum likelihood technique using the observations one at a time and incorporating new information into an 'Innovation' model is described (cf. Eqs. (20')-(27') and (42)-(49)).

5. An EVM with randomly time-varying coefficients is estimated using Extended Kalman Filtering and Approximate Maximum Likelihood Estimation. The technique is applicable to nonlinear systems, as well.

REFERENCES

[1] Wald, A., "The filtering of straight lines if both variables are subject to error." Ann. Math. Stat., Vol. 11, pp. 284-300 (1940).

[2] Neyman, J. and Scott, E., "Consistent estimates based on partially consistent observations." Econometrica, Vol. 16, pp. 1-32 (1948).

[3] Koopmans, T.C. and Reiersøl, O., "The identification of structural characteristics." Ann. Math. Stat., Vol. 21, pp. 165-181 (1950).

[4] Neyman, J., "Existence of a consistent estimate of the directional paramenter in a linear structural relation between two variables." Ann. Math. Stat., Vol. 22, pp. 497-512 (1951).

[5] Kiefer, J. and Wolfowitz, J., "Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters." Ann. Math. Stat., Vol. 27, pp. 887-906 (1956).

[6] Madansky, A., "The fitting of straight lines when both variables are subject to error." J. Amer. Stat. Assoc., Vol. 54, pp. 173-205 (1959).

[7] Solari, M.E., "The maximum likelihood solution of the problem of estimating a linear functional relationship." J. Roy. Statist. Soc., Ser. B, Vol. 31, pp. 372-375 (1969).

[8] Zellner, A., An Introduction to Bayesian Inference in Econometrics. Wiley, New York (1971).

[9] Moran, P.A.P., "Estimating structural and functional relationships." J. Multivariate Analysis, Vol. 1, pp. 232-255 (1971).

[10] Florens, J.P., Mouchart, M., and Richard, J.F., "Bayesian Inference in Error-in-Variables Models." J. of Multivariate Analysis, Vol. 4, pp. 419-452 (1974).

[11] Reiersøl, O., "Identifiablity of a linear relation between variables which are subject to error." Econometrica, Vol. 18, pp. 375-389 (1950).

[12] Bryson, A.E., and Ho, Y.C., Applied Optimal Control. Xerox Publishing Company, New York (1969).

[13] Mehra, R.K., "On Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering." IEEE Trans. Aut. Cont., Vol AC-16, No. 1 (Feb., 1971).

[14] Mehra, R.K., "Identification in Control and Econometrics, Similarities and Differences." Annals of Economic and Social Measurement (Jan., 1974).

[15] Mehra, R.K., "On the Identification of Variances and Adaptive Kalman Filtering." IEEE Trans. Aut. Cont. Vol. AC-15, No. 2 (April, 1970).

[16] Johnston, J., Econometric Methods. McGraw-Hill, New York (1963).

[17] Box, G.E.P., and Jenkins, G.M., Time Series Analysis Forecasting and Control. Holden Day, San Francisco (1970).

[18] Levinson, N., "The Wiener RMS Criterion in Filter Design and Prediction," Appendix B of N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series, Wiley, New York (1949).

[19] Durbin, J., "The fitting of Time Series Models." Rev. Inst. Int. Statis., Vol. 28, pp. 233-244 (1960).

[20] Wiggins, R.A., and Robinson, E.A., "Recursive solution to the multichannel filtering problem." J. Geophysics Res., Vol. 70, pp. 1885-1891 (1965).

[21] Krishnaprasad, P.S., and Mehra, R.K., "Fast recursive algorithms for parameter estimation." Technical Report, Division of Engineering and Applied Physics, Harvard University, Cambridge, Massachusetts (Dec., 1974).

[22] Akaike, H., "Canonical correlation analysis of time series and the use of an information criterion," in System Identification, Advances and Case Studies," R.K. Mehra and D.G. Lainiotis, Eds. Marcel-Dekker, New York (1975).

[23] Golub, G.H., "Matrix decompositions and statistical calculations," in Statistical Computation, R.C. Milton and J.A. Nelda, Eds. Academic Press, New York (1969).

[24] Astrom, K.J., Bohlin, T., and Wenmark, S., "Automatic construction of linear stochastic dynamic models for stationary industrial processes with random disturbances using operating records." Rep. TP 18.150, IBM Nordic Laboratories, Lindingo, Sweden (1965).

[25] Ljung, L., "On the consistency of prediction error identification methods," in Systems Identification, Advances and Case Studies., R.K. Mehra and D.G. Lainiotis, Eds., Marcel-Dekker, New York (1975).

[26] Mehra, R.K., "Identification of stochastic linear dynamic systems using Kalman filter representation." AIAA Journal (Jan., 1971).

[27] Gupta, N.K., and Mehra, R.K., "Computational aspects of maximum likelihood estimation and reduction in sensitivity function calculations." IEEE Trans. Aut. Cont. (Dec., 1974).

[28] Schweppe, F.C., Uncertain Dynamic Systems. Prentice Hall, New York (1974).

[29] Mehra, R.K., "A comparison of several nonlinear filters for re-entry vehicle tracking." IEEE Trans. Aut. Cont. (August, 1971).

[30] Mehra, R.K., and Krishnaprasad, P.S., "A unified approach to structural estimation of distributed lags and stochastic differential equations." Third NBER Conf. on Stochastic Control, Washington, D.C. (1974).

[31] Mehra, R.K., and Tyler, J.S., "Case studies in aircraft parameter identification." Third IFAC Conference on Identification, The Hague, Netherlands (June, 1973).

[32] Brown, M.L., "Robust Line Estimation with Errors in Both Variables." Working Paper No. 83, Computer Research Centre for Economics And Management Science, National Bureau of Economic Research, Inc., Cambridge, Mass. (May, 1975).