Volume Title: Annals of Economic and Social Measurement, Volume 3, number 1

Volume Author/Editor: Sanford V. Berg, editor

Volume Publisher:

Volume URL: http://www.nber.org/books/aesm74-1

Publication Date: 1974

Chapter Title: The Importance of Kalman Filtering Methods for Economic Systems

Chapter Author: Michael Athans

Chapter URL: http://www.nber.org/chapters/c9994

Chapter pages in book: (p. 49 - 64)

Annals of Economic and Social Measurement, 3/1, 1974

THE IMPORTANCE OF KALMAN FILTERING METHODS FOR ECONOMIC SYSTEMS*

BY MICHAEL ATHANS†

The purpose of this paper is to indicate how Kalman filtering techniques are potentially useful in modelling economic systems.

1. INTRODUCTION

The purpose of this paper is to examine problems of parameter estimation for linear and nonlinear econometric models from the point of view of Kalman filtering (see references [1] and [2]). It will be shown that if one makes the assumptions that

(1) all endogenous and exogenous variables can be measured exactly
(2) the structure and order of the econometric model have been fixed
(3) the variance and means of the white noise driving the econometric equations have been estimated
(4) the parameters to be estimated appear linearly in the difference equations

then "on-line" estimates of the parameters of the econometric model can be generated in a straightforward manner through the use of the Kalman filtering algorithm.

The Kalman filter represents one of the major contributions in modern control theory. Since its original development (references [1] and [2]) it has been rederived from several points of view bringing into focus its properties from both a probabilistic and optimization viewpoint (references [3] to [9]). Its importance in stochastic control can be appreciated in view of the numerous applications (references [5], [6], [9] to [16]).

In spite of the recent interest in modern control theory by mathematical economists the potential advantages of Kalman filtering methods have not been fully appreciated by economists and management scientists. One of the reasons is that the straightforward application of Kalman filtering methods involves estimation of state variables, whenever the actual measurements are corrupted by white noise. In most economic applications, the measurements of the endogenous and exogenous variables are assumed exact.

In this paper, we shall indicate that the Kalman filtering algorithm does have potential use for an important class of economic problems, namely those involving the refinement of the parameter estimates (and of their variances) in an econometric model. Right at the start we should like to *emphasize* that the use of the Kalman filtering techniques is viewed *not as a replacement, but rather as a supplement, to*

* This research was conducted at the Decision and Control Sciences Group of the M.I.T. Electronic Systems Laboratory with partial support provided by NSF under grant GK-25781. This paper was presented at the Second Workshop on Stochastic Control and Economic Systems, University of Chicago Business School, Chicago, Illinois, June 7–9, 1973, sponsored by the National Bureau for Economic Research, Inc.

† Room 35-308, Department of Electrical Engineering, M.I.T., Cambridge, Mass., 02139.

traditional econometric methods. We visualize that the Kalman filtering methods should become useful only after an econometrician has constructed the mathematical model of a microeconomic or macroeconomic systems. Thus it may represent a final "tune-up" of the econometric model.

To illustrate these ideas, the paper is organized as follows. In Section 2 we present a summary of the standard discrete time Kalman filter algorithm. In Section 3 we consider the problem of identifying the parameters (constant or stochastically varying) of an econometric model that involves the interrelationships of a single endogenous variable, $y(t)$, to a single exogenous variable, $u(t)$, through the use of the Kalman filter. In Section 4 we consider the problem of parameter estimation of an econometric model with several endogenous and exogenous variables, and show that this class of problems reduces to those analyzed in Section 3. Section 5 contains some additional speculative suggestions on how these techniques can be used to reconcile "different" econometric models.

Before commencing the technical development the author would like to elaborate upon two further points.

1. The author (not being an expert in the art and science of econometrics) does not know if the same technique, under a different name and disguise, is not available in the econometric literature: chances are that it probably is. If indeed this is the case, then at the very least this paper can serve the interchange of experiences in this area between the economic and control community.
2. No simulation results are available up to now to indicate the advantages and disadvantages of using the Kalman filtering algorithm for parameter identification in econometric models; such a project is currently underway in the M.I.T. Electrical Engineering Department, using the econometric model developed by Pindyck (reference [17]) as the first test case. However, no numerical results are available as yet.

2. Summary of the Discrete Kalman Filter

In this section we provide a summary of the discrete time Kalman filter.

2.1. *Mathematical Problem Formulation*

Given a vector system of difference equations (state dynamics)

(2.1)
$$\mathbf{x}(t+1) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) + \mathbf{L}(t)\boldsymbol{\xi}(t)$$

where:

the time index t takes values $t = 0, 1, 2, \ldots$
$\mathbf{x}(t)$ is an n-dimensional vector (the state of (2.1))
$\mathbf{u}(t)$ is an m-dimensional vector (the input of (2.1))
$\boldsymbol{\xi}(t)$ is a p-dimensional vector (the system noise of (2.1))
$\mathbf{A}(t), \mathbf{B}(t), \mathbf{L}(t)$ are known matrices of appropriate dimensions.

Suppose that the following measurement equation holds for $t = 1, 2 \ldots$

(2.2)
$$\mathbf{z}(t) = \mathbf{C}(t)\mathbf{x}(t) + \boldsymbol{\theta}(t)$$

where

$\mathbf{z}(t)$ is an r-dimensional vector of actual measurements
$\boldsymbol{\theta}(t)$ is an r-dimensional vector (the measurement noise)
$\mathbf{C}(t)$ is a known matrix.

Assumptions. We assume that

(a) $\mathbf{x}(0)$ is a Gaussian random vector with known mean $\bar{\mathbf{x}}(0)$ and known covariance matrix $\boldsymbol{\Sigma}(0)$, i.e.,

(2.3) $$E\{\mathbf{x}(0)\} = \bar{\mathbf{x}}(0)$$

(2.4) $$\operatorname{cov}[\mathbf{x}(0);\mathbf{x}(0)] = E\{(\mathbf{x}(0) - \bar{\mathbf{x}}(0))(\mathbf{x}(0) - \bar{\mathbf{x}}(0))'\} = \boldsymbol{\Sigma}(0).$$

(b) $\boldsymbol{\xi}(t)$ is a Gaussian random vector with zero mean for all $t = 0, 1, 2, \ldots,$ and independent in time (discrete white noise), i.e.,

(2.5) $$E\{\boldsymbol{\xi}(t)\} = \mathbf{0}$$

(2.6) $$\operatorname{cov}[\boldsymbol{\xi}(t);\boldsymbol{\xi}(\tau)] = E\{\boldsymbol{\xi}(t)\boldsymbol{\xi}'(\tau)\} = \boldsymbol{\Xi}(t)\delta_{t\tau}$$

where $\delta_{t\tau}$ is the Kroenecker delta

(2.7) $$\delta_{t\tau} = \begin{cases} 1 & \text{if } t = \tau \\ 0 & \text{if } t \neq \tau \end{cases}$$

and $\boldsymbol{\Xi}(t)$ is the known positive semidefinite covariance matrix of $\boldsymbol{\xi}(t)$.

(c) $\boldsymbol{\theta}(t)$ is a Gaussian random vector with zero mean for all $t = 1, 2, \ldots,$ and independent in time (discrete white noise), i.e.,

(2.8) $$E\{\boldsymbol{\theta}(t)\} = \mathbf{0}$$

(2.9) $$\operatorname{cov}[\boldsymbol{\theta}(t);\boldsymbol{\theta}(\tau)] = E\{\boldsymbol{\theta}(t)\boldsymbol{\theta}'(\tau)\} = \boldsymbol{\Theta}(t)\delta_{t\tau}$$

and $\boldsymbol{\Theta}(t)$ is the known positive definite covariance matrix of $\boldsymbol{\theta}(t)$. We assume that $\mathbf{x}(0)$, $\boldsymbol{\xi}(t)$, and $\boldsymbol{\theta}(\tau)$ are mutually independent for all values of t and τ.

(d) $\mathbf{u}(t)$ is a deterministic time sequence.

(e) the matrices $\mathbf{A}(t)$, $\mathbf{B}(t)$, $\mathbf{C}(t)$, $\mathbf{L}(t)$ are all deterministic.

2.2. *Problem Statement*

It is desired to construct a "best" estimate of $\mathbf{x}(t)$ given past values of the input vector

(2.10) $$U(t-1) \triangleq \{\mathbf{u}(0), \mathbf{u}(1), \ldots, \mathbf{u}(t-1)\}$$

and past values of the measurement vector

(2.11) $$Z(t) \triangleq \{\mathbf{z}(1), \mathbf{z}(2), \ldots, \mathbf{z}(t)\}.$$

The best estimate is denoted by $\hat{\mathbf{x}}(t|t)$ and is defined as the *conditional mean* of $\mathbf{x}(t)$ given $U(t-1)$, and $Z(t)$, i.e.,

(2.12) $$\hat{\mathbf{x}}(t|t) = E\{\mathbf{x}(t)|Z(t), U(t-1)\}.$$

2.3. *Solution to the Problem*

The linearity of equations (2.1) and (2.2), together with the Gaussian assumptions on $\mathbf{x}(0)$, $\boldsymbol{\xi}(t)$, $\boldsymbol{\theta}(\tau)$ implies that the conditional probability density

$$p(\mathbf{x}(t)/Z(t), U(t-1))$$

is Gaussian and, hence, it is uniquely characterized by its conditional mean $\hat{\mathbf{x}}(t|t)$ and conditional covariance matrix $\boldsymbol{\Sigma}(t|t)$

(2.13) $$\boldsymbol{\Sigma}(t|t) \triangleq \text{cov}\,[\mathbf{x}(t): \mathbf{x}(t)|Z(t), U(t-1)].$$

The discrete Kalman filter (see references [1] to [11]) yields a powerful sequential algorithm that can be used to generate $\hat{\mathbf{x}}(t|t)$ and $\boldsymbol{\Sigma}(t|t)$.

The calculation of the conditional covariance matrix is done off-line.

Initialization at $t = 0$

(2.14) $$\boldsymbol{\Sigma}(0|0) = \boldsymbol{\Sigma}(0).$$

Covariance prediction equation

(2.15) $$\boldsymbol{\Sigma}(t+1|t) = \mathbf{A}(t)\boldsymbol{\Sigma}(t|t)\mathbf{A}'(t) + \mathbf{L}(t)\boldsymbol{\Xi}(t)\mathbf{L}'(t).$$

Covariance update equation

(2.16) $$\begin{aligned}\boldsymbol{\Sigma}(t+1|t+1) = \boldsymbol{\Sigma}(t+1|t) - \boldsymbol{\Sigma}(t+1|t)\mathbf{C}'(t+1)\\ \times [\mathbf{C}(t+1)\boldsymbol{\Sigma}(t+1|t)\mathbf{C}'(t+1) + \boldsymbol{\Theta}(t+1)]^{-1}\\ \times \mathbf{C}(t+1)\boldsymbol{\Sigma}(t+1|t).\end{aligned}$$

Filter gain calculation

(2.17) $$\mathbf{H}(t+1) = \boldsymbol{\Sigma}(t+1|t+1)\mathbf{C}'(t+1)\boldsymbol{\Theta}^{-1}(t+1).$$

The calculation of the condition mean $\hat{\mathbf{x}}(t|t)$ is done on-line.

Initialization at $t = 0$

(2.18) $$\hat{\mathbf{x}}(0|0) = \bar{\mathbf{x}}(0).$$

Mean prediction equation

(2.19) $$\hat{\mathbf{x}}(t+1|t) = \mathbf{A}(t)\hat{\mathbf{x}}(t|t) + \mathbf{B}(t)\mathbf{u}(t).$$

Residual (innovations) calculation

(2.20) $$\mathbf{r}(t+1) = \mathbf{z}(t+1) - \mathbf{C}(t+1)\hat{\mathbf{x}}(t+1|t).$$

Mean update equation

(2.21) $$\hat{\mathbf{x}}(t+1|t+1) = \hat{\mathbf{x}}(t+1|t) + \mathbf{H}(t+1)\mathbf{r}(t+1).$$

3. Analysis of a Simple Model

In this section we shall consider a simple econometric model involving the interrelationship of a scalar endogenous (output) variable, $y(t)$, and a scalar exogenous (input) variable, $u(t)$. We shall assume that the current value of $y(t)$ depends upon lagged values of itself and of $u(t)$.

To illustrate the usefulness of the Kalman filtering algorithm we shall consider three distinct cases

(a) linear model, with constant parameters

(b) linear model, with time-varying parameters

(c) nonlinear model.

In each case we shall illustrate how to formulate the problem so that the general Kalman filtering algorithm described in the previous section is directly applicable.

3.1. *Linear Model with Constant Parameters*

Let the input (exogenous variable) to the system be denoted by $u(t)$ and its output (endogenous variable) by $y(t)$: we assume that t is a discrete-time index attaining values

$$t = 0, 1, 2, \ldots \tag{3.1}$$

Suppose that the system is described by the following linear time-invariant stochastic difference equation:

$$y(t) = \sum_{i=1}^{n} a_i y(t-i) + \sum_{j=1}^{m} b_j u(t-j) + \theta(t). \tag{3.2}$$

Let us suppose that the parameters

$$\begin{cases} a_1, a_2, \ldots, a_n \\ b_1, b_2, \ldots, b_m \end{cases} \tag{3.3}$$

are *known to be constant*. Also we assume that the "noise" $\theta(t)$ is discrete white noise with *known statistics*, i.e.,

$$E\{\theta(t)\} = 0 \tag{3.4}$$

$$E\{\theta(t)\theta(\tau)\} = \Theta\delta_{t\tau}; \qquad \Theta > 0 \tag{3.5}$$

where $\delta_{t\tau}$ is the Kroenecker delta. Furthermore, we assume that we can measure (and store) *exactly* the values of the output, $y(t)$, and of the input, $u(t)$, at each instant of time.

Next we shall define certain vectors which shall be used to transform the parameter estimation problem into an equivalent filtering problem.

We define the *equivalent state vector*, $\mathbf{x}(t)$, to be the $(n+m)$ dimensional vector of the system parameters a_i and b_j. More precisely,

$$\mathbf{x}'(t) \triangleq [a_1 a_2 \ldots a_n b_1 b_2 \ldots b_m] \tag{3.6}$$

where $\mathbf{x}'(t)$ is a row vector (′ denotes transposition).

Next we define the $(n+m)$ row vector $\mathbf{c}'(t)$ as follows:

$$\mathbf{c}'(t) \triangleq [y(t-1)y(t-2)\ldots y(t-n)u(t-1)u(t-2)\ldots u(t-m)]. \tag{3.7}$$

Using these definitions, the stochastic difference equation (3.2) can be written in the form:

$$\boxed{y(t) = \mathbf{c}'(t)\mathbf{x}(t) + \theta(t)}. \tag{3.8}$$

Note that at time t, the numerical value of the row vector $\mathbf{c}'(t)$ is known from the prior n measurements of the output and the m prior measurements of the input. Although the outputs are random (because of the white noise $\theta(t)$), nonetheless at time t they have been observed, and hence they are no longer random. It therefore follows that at time t, the row vector $\mathbf{c}'(t)$ is a known deterministic quantity.

We can now interpret equation (3.9) as a linear noisy measurement equation on the equivalent state vector $\mathbf{x}(t)$ where

- $y(t)$ represents the actual measurement at time t.
- $\theta(t)$ represents the value of the measurement noise at time t.

Hence, if we could construct a linear difference equation for the equivalent state vector $\mathbf{x}(t)$ we could apply directly the Kalman filtering algorithm.

Under our assumptions, the parameters a_i and b_j are constant. This automatically implies that the equivalent state vector $\mathbf{x}(t)$ satisfies the trivial difference equation

(3.9) $$\boxed{\mathbf{x}(t+1) = \mathbf{x}(t)} \quad \text{for all } t.$$

Hence, equations (3.8) and (3.9) define a simple linear filtering problem as far as the estimation of $\mathbf{x}(t)$ is concerned. In addition to the assumed data we also need the prior statistical information on $\mathbf{x}(0)$, or equivalently the prior means and covariance matrices of the parameters

(3.10) $$E\{a_i\},\ E\{b_j\}$$

(3.11) $$\operatorname{cov}[a_i, a_j],\ \operatorname{cov}[b_i, b_j],\ \operatorname{cov}[a_i, b_j].$$

Knowledge of these quantities will then define the prior mean of the equivalent state vector

(3.12) $$E\{\mathbf{x}(0)\}$$

and its prior covariance matrix

(3.13) $$\operatorname{cov}[\mathbf{x}(0); \mathbf{x}(0)].$$

Direct application of the general Kalman filtering algorithm yields the following relationship between successive estimates $\hat{\mathbf{x}}(t)$, $\hat{\mathbf{x}}(t-1)$ of the equivalent state vector and hence of the parameters a_i, b_j.

(3.14) $$\boxed{\hat{\mathbf{x}}(t) = \hat{\mathbf{x}}(t-1) + \mathbf{h}(t)[y(t) - \mathbf{c}'(t)\hat{\mathbf{x}}(t-1)]}$$

with the initial condition

(3.15) $$\hat{\mathbf{x}}(-1) = E\{\mathbf{x}(0)\}$$

and where the $(n+m)$ dimensional column vector $\mathbf{h}(t)$ in equation (3.14) is defined by

(3.16) $$\boxed{\mathbf{h}(t) = \frac{1}{\Theta}\Sigma(t)\mathbf{c}(t)}$$

In equation (3.16), $\Sigma(t)$ is the $(n + m) \times (n + m)$ error covariance matrix which can be evaluated by means of the matrix difference equation

$$(3.17) \quad \boxed{\Sigma(t) = \Sigma(t-1) - \frac{1}{\mathbf{c}'(t)\Sigma(t-1)\mathbf{c}(t) + \Theta} \cdot \Sigma(t-1)\mathbf{c}'(t)\mathbf{c}(t)\Sigma(t-1)}$$

with the initial condition

$$(3.18) \qquad \Sigma(-1) = \text{cov}\,[\mathbf{x}(0) : \mathbf{x}(0)].$$

Let us now discuss the interplay between this algorithm and the traditional econometric model. It should be self evident that in order to apply this algorithm, defined by equations (3.14) to (3.18) one *must* have available:

(a) the statistics (mean and variance) of the white noise $\theta(t)$
(b) the general structure of the system and, in particular, the number, n, of lags in the endogenous variable $y(t)$ and the number of lags, m, of the exogenous variable $u(t)$
(c) an initial guess for the parameters (reflected in the value of the prior mean $E\{\mathbf{x}(0)\}$ in equation (3.15)), and an initial value of the parameter variances and covariances (reflected in the value of the prior covariance matrix cov $[\mathbf{x}(0); \mathbf{x}(0)]$ in equation (3.18)).

It is precisely this information that is available from standard econometric models. *Hence, to use the Kalman filter algorithm one does need a prior econometric model.* The parameter values are then refined through the use of the Kalman filter algorithm, and new values for their variances and covariances obtained. It is then for this reason that we remarked in the Introduction that Kalman filtering techniques should be viewed as supplementary to econometric methods, rather than as a replacement.

3.2. *Linear Model with Time-Varying Parameters*

The above technique can be trivially extended to the case that the parameters of the difference equation are viewed not as constant but rather as being *time-varying and stochastic*. Traditional econometric methods do not appear well suited for the estimation of time-varying parameters simply because one must also identify the variance of the driving white noise $\theta(t)$. One can use the standard econometric techniques to process the economic data to arrive at constant parameter estimates. Using these now as constituting prior estimates, one can relax the constraint of constant parameters, and utilize the Kalman filter to arrive at time-varying parameter estimates.

The technical means by which this is communicated to the mathematics is as follows. Each parameter a_i, b_j now is viewed as being time-varying and, at the simplest level, is supposed to satisfy an equation of the form

$$(3.19) \qquad \begin{cases} a_i(t+1) = a_i(t) + \xi_i(t) \\ \\ b_j(t+1) = b_j(t) + \xi_j(t) \end{cases}$$

where the $\xi_i(t)$ are zero mean white noise sequences. Thus, each parameter can change in time in a stochastic manner. The prior information about the variability of each parameter about its mean value (as determined by the standard econometric model) is used to select (subjectively!!!) variance levels for the white noise levels $\xi_i(t)$.

We should like to stress that the *standard* Kalman filtering algorithm *cannot* be used to estimate the statistics of the white noise sequences $\xi_i(t)$. Rather, one has to postulate that they have zero mean and constant variance (although this stationary assumption is not crucial), and assign *subjective* numerical values for the variances.

The equivalent state vector $\mathbf{x}(t)$ is still modelled by equation (3.6), and the row vector $\mathbf{c}'(t)$ is still defined by equation (3.7). Then the measurement equation (3.8) relating the current measurement $y(t)$ to the current value of the equivalent state vector still holds.

In this case, the stochastic variability of the parameters can be modelled at the simplest level by the vector difference equation

$$\boxed{\mathbf{x}(t+1) = \mathbf{x}(t) + \boldsymbol{\xi}(t)} \tag{3.20}$$

rather than by equation (3.9). One assumes that

$$E\{\boldsymbol{\xi}(t)\} = \mathbf{0} \tag{3.21}$$

$$\operatorname{cov}[\boldsymbol{\xi}(t):\boldsymbol{\xi}(\tau)] = \boldsymbol{\Xi}\delta_{t\tau} \tag{3.22}$$

where $\boldsymbol{\Xi}$ can be selected diagonal, but, at the very least, positive semi-definite.

Under these assumptions, the equivalent state estimate $\hat{\mathbf{x}}(t)$ is still given by equation (3.14), and the vector $\mathbf{h}(t)$ is still given by equation (3.16). The only equation that changes is the covariance equation (3.17). The correct covariance equation for this time varying parameter case is

$$\boxed{\begin{aligned}\boldsymbol{\Sigma}(t) = \boldsymbol{\Sigma}(t-1) + \boldsymbol{\Xi} - &\frac{1}{\mathbf{c}'(t)[\boldsymbol{\Sigma}(t-1)+\boldsymbol{\Xi}]\mathbf{c}(t)+\Theta}\\ &\times[\boldsymbol{\Sigma}(t-1)+\boldsymbol{\Xi}]\mathbf{c}(t)\mathbf{c}'(t)[\boldsymbol{\Sigma}(t-1)+\boldsymbol{\Xi}]\end{aligned}} \tag{3.23}$$

with

$$\boldsymbol{\Sigma}(-1) = \operatorname{cov}[\mathbf{x}(0):\mathbf{x}(0)]. \tag{3.24}$$

Roughly speaking, the covariance matrix $\boldsymbol{\Xi}$ of the white noise $\boldsymbol{\xi}(t)$ increases the parameter uncertainty at each instant of time.

3.3. *The Nonlinear Model Case*

Up to now we have assumed that the difference equation relating the input sequence $u(t)$ to the output sequence $y(t)$ was linear.

This is not a necessary assumption as far as the parameter estimation problem is concerned. *What is important is to have the parameters to be estimated appear linearly in the difference equation.* To make this point precise, suppose that we have a difference equation of the form

$$y(t) = \sum_{i=1}^{Q} a_i f_i(t) + \theta(t) \tag{3.25}$$

where the a_i are the parameters to be estimated, and where $f_i(t)$ are functions that are allowed to depend *nonlinearly* upon

$$\left.\begin{matrix} y(t-1), y(t-2), \ldots, y(t-n) \\ u(t-1), u(t-2), \ldots, u(t-m) \end{matrix}\right\}.$$

The important thing is that prior measurements of the input and output sequence *uniquely and exactly* define the numerical values of the $f_i(t)$.

If this is the case, we once more construct the equivalent state vector $\mathbf{x}(t)$

$$\mathbf{x}'(t) \triangleq [a_1 a_2 \ldots a_Q] \tag{3.26}$$

and if the a_i are constant we have

$$\mathbf{x}(t+1) = \mathbf{x}(t). \tag{3.27}$$

We define the row vector $\mathbf{c}'(t)$ by

$$\mathbf{c}'(t) = [f_1(t) f_2(t) \ldots f_Q(t)] \tag{3.28}$$

so that equation (3.25) takes the form of

$$y(t) = \mathbf{c}'(t)\mathbf{x}(t) + \theta(t). \tag{3.29}$$

Once more equations (3.27) and (3.29) define a linear filtering problem for $\mathbf{x}(t)$ and the Kalman filter algorithm given by equations (3.14) to (3.18) can be used.

4. Multivariable Problems

Complex microeconomic and macroeconomic systems are in general characterized by several endogenous and exogenous variables. We shall denote the endogenous (output) variables by

$$y_1(t), y_2(t), \ldots, y_N(t) \tag{4.1}$$

and the exogenous (input) variables by

$$u_1(t), u_2(t), \ldots, u_M(t). \tag{4.2}$$

Once more we can consider both linear and nonlinear econometric models. For the sake of exposition we shall only analyze linear models.

In linear econometric models one relates the current value of each endogenous variable $y_i(t)$, $i = 1, 2, \ldots, N$ to

(a) linear combinations of lagged values of (possibly all) endogenous variables $y_j(\tau)$, $\tau = t-1, t-2, \ldots, t-n_i$

(b) linear combinations of lagged values of (possibly all) exogenous variables $u_k(\tau)$, $\tau = t-1, t-2, \ldots, t-m_i$

(c) and in the presence of additive discrete white noise $\theta_i(t)$.

In order to write the structure of the econometric model in a compact form we define the endogenous N-dimensional column vector $\mathbf{y}(t)$ as the one naturally defined by the endogenous variables $y_1(t), y_2(t), \ldots, y_N(t)$. i.e., $\mathbf{y}'(t) = [y_1(t)y_2(t)\ldots y_N(t)]$. Similarly, we define the M-dimensional exogenous vector $\mathbf{u}(t)$ by

$$\mathbf{u}'(t) = [u_1(t)u_2(t)\ldots u_M(t)]. \tag{4.3}$$

Using the above notation a linear econometric model with constant parameters can be written using the vector difference equation

$$\mathbf{y}(t) = \mathbf{A}_1\mathbf{y}(t-1) + \mathbf{A}_2\mathbf{y}(t-2) + \ldots + \mathbf{A}_n\mathbf{y}(t-n) + \mathbf{B}_1\mathbf{u}(t-1) + \mathbf{B}_2\mathbf{u}(t-2) + \ldots + \mathbf{B}_m\mathbf{u}(t-m) + \boldsymbol{\theta}(t) \tag{4.4}$$

where $\boldsymbol{\theta}(t)$ is the N-dimensional vector defined by the scalar white noise sequences

$$\boldsymbol{\theta}'(t) \triangleq [\theta_1(t)\theta_2(t)\ldots\theta_N(t)] \tag{4.5}$$

and

$$n = \max_i n_i, \qquad m = \max_i m_i \tag{4.6}$$

and where the constant parameters appear as elements of the matrices

$$\begin{cases} \mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n \\ \mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m \end{cases} \tag{4.7}$$

(Note that because of the variability of the lags appearing in each endogenous equation many of these matrices may have several zero elements.)

The available econometric methods provide at the very least

- mean values for all parameters
- their variances (and perhaps covariances)
- the covariance matrix of $\boldsymbol{\theta}(t)$

We now introduce the following notation. Let

$\mathbf{a}_j^{i\prime}$ denote the j-th row vector of the matrix $\mathbf{A}_i$
$\mathbf{b}_j^{i\prime}$ denote the j-th row vector of the matrix $\mathbf{B}_i$.

Using this notation we can write the general econometric relation (4.4) as follows. For $i = 1, 2, \ldots, N$

$$y_i(t) = \sum_{j=1}^{n} \mathbf{a}_j^{i\prime}\mathbf{y}(t-j) + \sum_{k=1}^{m} \mathbf{b}_k^{i\prime}\mathbf{u}(t-k) + \theta_i(t). \tag{4.8}$$

For each i, we can now define an equivalent state vector $\mathbf{x}_i(t)$ which contains all unknown parameters in the i-th endogenous equation (4.8).

$$\mathbf{x}_i'(t) = [\mathbf{a}_1^{i\prime}, \mathbf{a}_2^{i\prime}, \ldots, \mathbf{a}_1^{i\prime}, \mathbf{b}_1^{i\prime}, \mathbf{b}_2^{i\prime}, \ldots, \mathbf{b}_m^{i\prime}]. \tag{4.9}$$

We also define the row vector $\mathbf{c}'(t)$ (independent of i) by combining all appropriate past measurements of all endogenous and exogenous variables

$$\mathbf{c}'(t) \triangleq [\mathbf{y}'(t-1)\mathbf{y}'(t-2)\ldots\mathbf{y}'(t-n)\mathbf{u}'(t-1)\mathbf{u}'(t-2)\ldots\mathbf{u}'(t-m)]. \tag{4.10}$$

Using this notation, we write equation (4.8) in the form

$$\boxed{y_i(t) = \mathbf{c}'(t)\mathbf{x}_i(t) + \theta_i(t)}: \qquad i = 1, 2, \ldots, n \tag{4.11}$$

which represents a linear noisy measurement relationship in the equivalent state vector $\mathbf{x}_i(t)$. Once more, the assumption that the parameters are constant leads to the trivial difference equation

$$\boxed{\mathbf{x}_i(t+1) = \mathbf{x}_i(t)} \tag{4.12}$$

for the parameters associated with the i-th equation (4.8).

Hence, exactly the same techniques as those described in the previous section 3.1 can be used to determine the appropriate parameter estimates

$$\hat{\mathbf{x}}_i(t): \qquad i = 1, 2, \ldots, N \tag{4.13}$$

and the corresponding error covariance matrices

$$\mathbf{\Sigma}_i(t): \qquad i = 1, 2, \ldots, N \tag{4.13}$$

namely:

$$\hat{\mathbf{x}}_i(t) = \hat{\mathbf{x}}_i(t-1) + \mathbf{h}_i(t)[y_i(t) - \mathbf{c}'(t)\hat{\mathbf{x}}_i(t-1)]; \qquad \hat{\mathbf{x}}_i(-1) = E\{\mathbf{x}_i(0)\} \tag{4.15}$$

$$\mathbf{h}_i(t) = \frac{1}{\theta_i}\mathbf{\Sigma}_i(t)\mathbf{c}(t) \tag{4.16}$$

$$\mathbf{\Sigma}_i(t) = \mathbf{\Sigma}_i(t-1) - \frac{1}{\mathbf{c}'(t)\mathbf{\Sigma}_i(t-1)\mathbf{c}(t) + \theta_i} \cdot \mathbf{\Sigma}_i(t-1)\mathbf{c}(t)\mathbf{c}'(t)\mathbf{\Sigma}_i(t-1) \tag{4.17}$$

$$\mathbf{\Sigma}_i(-1) = \operatorname{cov}[\mathbf{x}_i(0) : \mathbf{x}_i(0)]. \tag{4.18}$$

Once more the standard econometric model is necessary to specify

1. the prior estimated white noise variances θ_i
2. the prior parameter estimates $E\{\mathbf{x}_i(0)\}$
3. the prior parameter covariance matrices $\operatorname{cov}[\mathbf{x}_i(0) : \mathbf{x}_i(0)]$.

The estimate generation of all parameters in the econometric model, then reduces to running N independent Kalman filters.

It is self evident that the remarks made in Section 3 regarding time-varying parameters and/or nonlinear models are also applicable to this multivariable case.

5. Reconciliation of Different Econometric Models

In this section we shall examine an extension of the previous results, which may be of *potential* usefulness in reconciling different econometric models. As we have remarked, different assumptions and statistical techniques enter in the construction of econometric models. Hence, it is not surprising that the predictions and forecasts of econometric models are variable. Reference [18] contains several papers dealing with this question, where attention was focused upon comparisons of quarterly econometric models of the limited states. See also reference [19].

It is, of course, extremely difficult to draw any precise conclusions about the "goodness" of any given econometric model, since they are usually different from the aggregation viewpoint, and they differ in the time series techniques utilized to estimate the parameters. Thus, this is not the type of problem that we shall address in this paper.

What we shall examine is the problem of estimating the parameters of a single econometric model starting from N econometric models with "identical" structure, but different parameter values.

In order to make precise the concepts and calculations involved, we shall outline the method and objectives in a mathematical framework. Then we shall explain what one may have to do when one deals with "real" econometric models. The ideas will be illustrated by considering linear time-invariant econometric models involving the interrelationship of a single endogenous variable, $y(t)$, and a single exogenous variable, $u(t)$, of the type described in Section 3. The techniques are also applicable to multivariable models of the type discussed in Section 4, in view of the decomposition properties.

Suppose that the *true* econometric model is of the form

$$y(t) = \sum_{i=1}^{n} a_i y(t-i) + \sum_{j=1}^{m} b_j u(t-j) + \theta(t) \tag{5.1}$$

where a_i and b_j are the true unknown parameters, and $\theta(t)$ is the actual white noise.

Let us now suppose that different econometricians have constructed N different econometric models that in their view represent a "good" approximation to the true system (5.1). We shall let $k = 1, 2, \ldots, N$ index these distinct econometric models. Thus, each of the models constructed is described by the equation

$$y(t) = \sum_{i=1}^{n} a_i^k y(t-i) + \sum_{j=1}^{m} b_j^k u(t-j) + \theta_k(t) \tag{5.2}$$

where, for the sake of simplicity, we assume that

$$E\{\theta_k(t)\} = 0 \qquad \text{for all } k \tag{5.3}$$

$$E\{\theta_k(t)\theta_k(\tau)\} = \Theta_k \delta_{t\tau}. \tag{5.4}$$

In addition to the estimates a_i^k, b_j^k of the parameters, each econometric model also provides us with the variances and possibly covariances, of the parameters; these would also differ from model to model.

Now imagine that a central agency, e.g., the Federal Reserve Board, would like to make some econometric predictions, and that it has these N different models at its disposal. Roughly speaking, it may be interested in "combining" all these models into a single one; this may correspond to obtaining new estimates for the parameters of the model (5.1), on the basis of information already available by the N econometric models. The central agency may have some prior probabilities

$$P_1(0), P_2(0), \ldots, P_N(0) \tag{5.5}$$

that reflect their subjective prior confidence on the accuracy of each econometric model. One simple way then would be to construct new parameter estimates by weighing the different coefficients, i.e.,

$$a_i = \sum_{k=1}^{N} P_k(0)a_i^k. \tag{5.6}$$

However, this technique may strongly be influenced by the prior biases of the agency as expressed by the prior probabilities $P_k(0)$.

A more sophisticated way is to utilize the Kalman filtering algorithm described in Section 3, together with a hypothesis testing algorithm, with each econometric model corresponding to a distinct hypothesis. We describe below how this can be accomplished.

Following the techniques of Section 3 each econometric model is described by its equivalent state vector $\mathbf{x}_k(t)$ satisfying the trivial difference equation

$$\mathbf{x}_k(t+1) = \mathbf{x}_k(t) \tag{5.7}$$

and the equivalent measurement equation

$$y(t) = \mathbf{c}'(t)\mathbf{x}_k(t) + \theta_k(t) \tag{5.8}$$

where $\mathbf{c}'(t)$ is common to all models and given by equation (3.7). Each model is characterized by different values of

$$E\{\mathbf{x}_k(0)\} \tag{5.9}$$

$$\operatorname{cov}[\mathbf{x}_k(0):\mathbf{x}_k(0)] \tag{5.10}$$

$$\Theta_k \tag{5.11}$$

where (5.9) implies that the numerical parameter estimates of the parameters of each model are different, equation (5.10) reflects the fact that the variances of the parameter estimates are also different, and (5.11) implies that the estimated variance of the driving white noise also differs from model to model.

The problem faced by the central agency is to construct an estimate $\hat{\mathbf{x}}(t)$ of the parameters associated with the overall model (5.1). This estimate $\hat{\mathbf{x}}(t)$ can be constructed by the following algorithm

$$\hat{\mathbf{x}}(t) = \sum_{k=1}^{N} \hat{\mathbf{x}}_k(t)P_k(t) \tag{5.12}$$

where

$\hat{\mathbf{x}}_k(t)$ is the estimate generated by the Kalman filter matched to the k-th econometric model [see equations (3.14) to (3.18)].

$P_k(t)$ is the posterior probability, given measurements up to time t, that the k-th econometric model is the correct one.

The optimality of the estimate (5.12) is most clearly understood if one makes the following assumptions. *All random variables are Gaussian.* Then $\hat{\mathbf{x}}(t)$ is the true conditional mean given the measurements up to time t, $\hat{\mathbf{x}}_k(t)$ is the conditional mean under the hypothesis that the k-th model is the correct one, and $P_k(t)$ is the conditional probability that the hypothesis of the k-th model being the correct one is valid (see reference [20]).

Since we have already given the algorithm that generates the individual model estimates $\hat{x}_k(t)$ using the Kalman filtering algorithm of Section 3, we shall now describes how one generates the probabilities $P_k(t)$. Use of Bayes rule yields the equation

$$P_k(t) = \frac{p(Y(t)|H_k)P_k(0)}{\sum_{j=1}^{N} p(Y(t)|H_j)P_j(0)}. \tag{5.13}$$

In equation (5.13) the $P_k(0)$ are the prior probabilities reflecting the initial confidence of the central agency upon the k-th econometric model. The term $p(Y(t)|H_k)$ is the probability density function of observing the entire set of measurements

$$Y(t) = \{y(t), y(t-1), \ldots, u(t-1), u(t-2), \ldots\} \tag{5.14}$$

under the assumption that the k-th model is the correct one.

Let $p(Y(t)|H_k)$ denote the conditional density function, so that $P(Y(t)|H_k)$ can be obtained by the integration of $p(Y(t)|H_k)$. We write

$$Y(t) = \{y(t), Y(t-1)\} \tag{5.15}$$

and, hence, by Bayes rule

$$p(Y(t)|H_k) = p(y(t), Y(t-1)|H_k) = p(y(t)|Y(t-1), H_k)p(Y(t-1)|H_k). \tag{5.16}$$

Equation (5.16) provides a recursive relationship between the conditional density functions $p(Y(t)|H_k)$ and $p(Y(t-1)|H_k)$; hence, given $p(Y(t-1)|H_k)$, the only additional quantity to be evaluated is

$$p(y(t)|Y(t-1), H_k) \tag{5.17}$$

so as to compute $p(Y(t)|H_k)$.

Under the Gaussian assumptions, the density (5.17) turns out to be Gaussian, and hence uniquely characterized by its mean and covariance matrix. The conditional mean is

$$E\{y(t)|Y(t-1), H_k\} = \mathbf{c}'(t)\hat{\mathbf{x}}_k(t-1) \tag{5.18}$$

where $\hat{\mathbf{x}}_k(t-1)$ is the estimate generated by the Kalman filter matched to the k-th econometric model. The conditional covariance matrix is given by

$$\text{cov}\,[y(t); y(t)|Y(t-1), H_k] = \mathbf{c}'(t)\boldsymbol{\Sigma}_k(t-1)\mathbf{c}'(t) + \Theta_k \tag{5.19}$$

where the covariance matrices $\boldsymbol{\Sigma}_k(t-1)$ are also generated by the Kalman filters (see references [20] to [22]).

In summary, the use of the N Kalman filters can be used to generate:

1. The probability that each econometric model is the correct one.
2. A weighted probabilistic average for the parameters of the econometric model to be used by the central agency, based upon the individual parameter estimates of each econometric model.

Although the theory and algorithms are relatively straightforward, much research needs to be done before one can use this method to "reconcile" and "combine" *actual* econometric models. because existing econometric models can be drastically different in their structural and aggregation properties. The method described does not seem to extend itself naturally if one wants to reconcile. say, a linear and a nonlinear econometric model.

However. it may be possible to compare different linear econometric models. with different dependence on lagged endogenous and exogenous variables. In this case, the central model *must* contain *all* lagged variables appearing in *all* models. If in a particular econometric model, a lagged variable does not appear. then it should be introduced with a zero coefficient. To be specific. suppose that in the k-th econometric model the dependence on $y(t-3)$ is absent. We can introduce this dependence by adding the term $\beta_k y(t-3)$. i.e.,

$$y(t) = \text{original terms} + \beta_k y(t-3). \tag{5.20}$$

The fact that the prior model did not include this term is communicated by setting

$$E\{\beta_k\} = 0. \tag{5.21}$$

However. one must *subjectively* use a *non-zero* initial variance $E\{\beta_k^2\}$ whose size reflects the feelings of the central agency on the importance of this term. Then the k-th Kalman filter will generate a non-zero estimate of β_k.

If this procedure is adopted, then each individual econometric model can indeed be made to have the same structure, and the techniques described in this section can be used. However, no numerical simulations are available at this time so as to make a definitive judgement even on the potential usefulness of this technique from an econometric viewpoint.

Massachusetts Institute of Technology

REFERENCES

[1] R. E. Kalman. "A New Approach to Linear Filtering and Prediction Problems." *Trans. ASME Journal of Basic Engineering*. Vol. 82. pp. 34–45. March 1960.
[2] R. E. Kalman and R. S. Bucy. "New Results in Linear Filtering and Prediction Theory." *Trans. ASME. Journal of Basic Engineering*, Vol. 83. pp. 95–107. December 1971.
[3] R. S. Bucy and P. D. Joseph. *Filtering for Stochastic Processes with Applications to Guidance*. Wiley. Inc.. New York. 1970.
[4] A. H. Jazwinskii. *Stochastic Processes and Filtering Theory*. Academic Press. Inc.. New York. 1970.
[5] R. C. K. Lee. *Optimal Estimation. Identification, and Control*. Cambridge. Mass.. M.I.T. Press. 1964.
[6] J. S. Meditch. *Stochastic Optimal Linear Estimation and Control*. McGraw-Hill Book Co.. Inc.. New York. 1969.
[7] N. Nahi. *Estimation Theory and Applications*, Wiley. Inc.. New York. 1969.
[8] A. P. Sage and J. L. Melsa. *Estimation Theory with Applications to Communications and Control*. McGraw-Hill Book Co.. Inc.. New York. 1970.
[9] *IEEE Transactions on Automatic Control*, Special Issue on the Linear Quadratic Gaussian Problem. Vol. AC-16. No. 6. December 1971.
[10] M. Athans. "The Discrete Time Linear-Quadratic-Gaussian Stochastic Control Problem." *Annals of Economic and Social Measurement*. Vol. 1. No. 4. pp. 449–491. 1972.
[11] K. J. Astrom, *Introduction to Stochastic Control Theory*, Academic Press. Inc.. New York, 1970.
[12] A. E. Bryson and Y. C. Ho. *Applied Optimal Control*. Blaisdell, Waltham. Mass.. 1969.
[13] H. J. Kushner. *Introduction to Stochastic Control*. Holt. Rinehart and Winston. Inc.. New York. 1971.

[14] M. Aoki. *Optimization of Stochastic Systems*. Academic Press, Inc., New York, 1967.
[15] D. Sworder. *Optimal Adaptive Control Systems*. Academic Press, Inc., New York, 1966.
[16] C. T. Leondes and J. O. Pearson. "Kalman Filtering of Systems with Parameter Uncertainties: A Survey." *Intern. Journal of Control*. Vol. 17. No. 4. pp. 785-801. 1973.
[17] R. S. Pindyck. *Optimal Planning for Economic Stabilization*. North Holland Publishing Company, 1973.
[18] B. G. Hickman (ed.). "Econometric Models of Cyclical Behavior." Vol. I and II. *Studies in Income and Wealth*. National Bureau of Economic Research. Columbia Univ. Press. 1972.
[19] V. Zarnowitz. "An Appraisal of Short-Term Economic Forecasts." National Bureau of Economic Research. Occasional Paper. No. 104. New York.
[20] D. T. Magill. "Optimal Adaptive Estimation of Sampled Stochastic Processes." *IEEE Trans. Automatic Control*. Vol. AC-10. pp. 434-439. 1965.
[21] D. G. Lainiotis. "Optimal Adaptive Estimation: Structure and Parameter Adaptation." *IEEE Trans. Automatic Control*. Vol. AC-16. pp. 160-170. April 1971.
[22] D. Willner. "Observation and Control of Partially Unknown Systems." Ph.D. Dissertation, Department of Electrical Engineering. M.I.T., May 1973.