RM-76-44

USE OF KALMAN FILTERING TECHNIQUES WHEN THE PARAMETERS OF A REGRESSION RELATIONSHIP ARE CHANGING OVER TIME ACCORDING TO A MULTIVARIATE ARIMA PROCESS

Johannes Ledolter

June 1976

Preface

In many areas of applied research at IIASA regression models are entertained to explain the relationship among variables. Assuming that the parameters are not changing over time, least squares methods provide minimum square error (MMSE) estimates. In some cases, however, the assumption of constant parameters is restrictive, and ways of incorporating parameter changes have to be found.

Use of Kalman filtering techniques when the parameters of a regression relationship are changing over time according to a multivariate ARIMA process.

Abstract

It is shown how Kalman filtering methodology can be applied to the estimation of the parameters in a regression model, when the parameters are subject to change over time. A multivariate ARIMA model for the parameters of the regression relationship is entertained and it is shown how this model can be brought into the state variable form. Furthermore it is shown how this procedure specializes to various cases already discussed in the literature.

1. State variable representation of dynamic and stochastic systems - Kalman filtering

A dynamic system with stochastic disturbances may be modelled in a state variable form

$$\begin{aligned} \underset{\sim}{x}_{t+1} &= A\underset{\sim}{x}_t + G\underset{\sim}{u}_t + \underset{\sim}{w}_{t+1} \\ \underset{\sim}{y}_t &= H\underset{\sim}{x}_t + \underset{\sim}{v}_t \end{aligned} \qquad (1.1)$$

where $\underset{\sim}{x}_t$ is a vector state variable which should be considered as an abstract quantity and which does not necessarily have a physical interpretation such as the input vector $\underset{\sim}{u}_t$ and the output vector $\underset{\sim}{y}_t$. $\underset{\sim}{w}_t$ and $\underset{\sim}{v}_t$ are uncrosscorrelated Normal white noise sequences with

$$E\underset{\sim}{w}_t = \underset{\sim}{0} \qquad E\underset{\sim}{w}_t\underset{\sim}{w}_t' = R_1$$

$$E\underset{\sim}{v}_t = \underset{\sim}{0} \qquad E\underset{\sim}{v}_t\underset{\sim}{v}_t' = R_2 \quad .$$

The parameter matrices A, G and H may be either constant or time varying.

Given the dynamic stochastic model with known dynamic and stochastic parameters, Kalman [3,4,5] obtains an estimate for the state vector $\underset{\sim}{x}_t$ given the observations on the input and output variables up to time t. He shows that the conditional distribution of $\underset{\sim}{x}_t$ given observations up to time t is a Normal with mean $\hat{\underset{\sim}{x}}_{t|t}$ and variance $P_{t|t}$ where

$$\hat{\underset{\sim}{x}}_{t|t} = \hat{\underset{\sim}{x}}_{t|t-1} + K_{t|t-1}(\underset{\sim}{y}_t - H\hat{\underset{\sim}{x}}_{t|t-1}) \tag{1.2}$$

$$P_{t|t} = P_{t|t-1} - K_{t|t-1}HP_{t|t-1} \tag{1.3}$$

and

$$\hat{\underset{\sim}{x}}_{t+1|t} = A\hat{\underset{\sim}{x}}_{t|t} + G\underset{\sim}{u}_t \tag{1.4}$$

$$P_{t+1|t} = AP_{t|t}A' + R_1 \tag{1.5}$$

where the Kalman gain is given by

$$K_{t|t-1} = P_{t|t-1}H'(HP_{t|t-1}H' + R_2)^{-1} \quad . \tag{1.6}$$

This set of recursive equations, together with specified initial conditions, provide the estimates and their updating equations for the state variables and their covariance matrix.

An excellent description of the state variable approach to dynamic and stochastic systems is given by MacGregor [6].

2. State variable representation of a regression model when its parameters change according to a multivariate ARIMA model.

Kalman's approach can be used to estimate the parameters in regression models where it is assumed that the parameters follow a general multivariate ARIMA process:

$$\Phi(B)\,\underset{\sim}{\beta}_{t+1} = \Theta(B)\,\underset{\sim}{\alpha}_{t+1}$$

$$y_t = \underset{\sim}{r}_t'\underset{\sim}{\beta}_t + a_t \qquad (2.1)$$

where

y_t is the dependent variable

$\underset{\sim}{r}_t$ is a $(k \times 1)$ vector of predetermined variables

$\underset{\sim}{\beta}_t$ is a $(k \times 1)$ vector of parameters

$\Phi(B) = I - \Phi_1 B - \ldots - \Phi_{p+d} B^{p+d}$

$\Theta(B) = I - \Theta_1 B - \ldots - \Theta_q B^q$

where the $\Phi_i (1 \leq i \leq p+d)$ and the $\Theta_j (1 \leq j \leq q)$ are known $(k \times k)$ matrices. B is the backshift operator $B^m \underset{\sim}{\beta}_t = \underset{\sim}{\beta}_{t-m}$. a_t is a white noise sequence with $Ea_t = 0$ and $Ea_t^2 = \sigma_a^2$. $\underset{\sim}{\alpha}_t$ is a multivariate white noise sequence with $E\underset{\sim}{\alpha}_t = 0$ and $E\underset{\sim}{\alpha}_t\underset{\sim}{\alpha}_t' = \Sigma_\alpha$. It is assumed that the zeros of $\det\{\Phi(B)\}$ lie on or outside the unit circle. The zeros of $\det\{\Theta(B)\}$ are assumed to lie outside the unit circle and $\det\{\Phi(B)\}$ and $\det\{\Theta(B)\}$ do not have common roots. Furthermore a_t and $\underset{\sim}{\alpha}_t$ are uncorrelated, i.e. $Ea_t\underset{\sim}{\alpha}_t = \underset{\sim}{0}$. Multivariate ARIMA processes are generalizations of the univariate

ARIMA processes discussed in great detail by Box and Jenkins [1]. Extensive discussion of the multivariate extension, for example, is given in Hannan [2].

In order to apply the Kalman filtering technique (equations (1.2) - (1.6)) we have to write equations (2.1) in form of state variables. The following theorem gives an equivalent state variable form for system (2.1), thus identifying the matrices H, A, G, R_1 and R_2.

<u>Theorem</u>: The model given in (2.1) has the equivalent state variable representation given below.

For p+d>q:

$$\begin{bmatrix} \underset{\sim}{\beta}^*_{1,t+1} \\ \underset{\sim}{\beta}^*_{2,t+1} \\ \vdots \\ \underset{\sim}{\beta}^*_{p+d-1,t+1} \\ \underset{\sim}{\beta}^*_{p+d,t+1} \end{bmatrix} = \left[\begin{array}{c|c} \Phi_1 & \\ \Phi_2 & \\ \vdots & I_* \\ \Phi_{p+d-1} & \\ \hline \Phi_{p+d} & O_* \end{array}\right] \begin{bmatrix} \underset{\sim}{\beta}^*_{1,t} \\ \underset{\sim}{\beta}^*_{2,t} \\ \vdots \\ \underset{\sim}{\beta}^*_{p+d-1,t} \\ \underset{\sim}{\beta}^*_{p+d,t} \end{bmatrix} + \begin{bmatrix} I \\ -\Theta_1 \\ \vdots \\ -\Theta_q \\ O \\ \vdots \\ O \end{bmatrix} \underset{\sim}{\alpha}_{t+1}$$

$$y_t = [\underset{\sim}{r}'_t \; \underset{\sim}{O}' \ldots \underset{\sim}{O}'] \begin{bmatrix} \underset{\sim}{\beta}^*_{1,t} \\ \underset{\sim}{\beta}^*_{2,t} \\ \vdots \\ \underset{\sim}{\beta}^*_{p+d,t} \end{bmatrix} + a_t$$

where I_* is the $[k(p+d-1) \times k(p+d-1)]$ identity matrix and O_* is a $[k \times (p+d-1)]$ matrix of zeros. O is a $[k \times k]$ matrix of zeros. $\underset{\sim}{O}'$ is a $[1 \times k]$ vector of zeros and I is the $[k \times k]$ identity matrix.

For $p+d \leq q$:

$$\begin{bmatrix} \underset{\sim}{\beta}^*_{1,t+1} \\ \vdots \\ \underset{\sim}{\beta}^*_{p+d,t+1} \\ \vdots \\ \underset{\sim}{\beta}^*_{q+1,t+1} \end{bmatrix} = \left[\begin{array}{c|c} \Phi_1 & \\ \Phi_2 & \\ \Phi_{p+d} & I_{**} \\ O & \\ \vdots & \\ \hline O & O_{**} \end{array}\right] \begin{bmatrix} \underset{\sim}{\beta}^*_{1,t} \\ \vdots \\ \underset{\sim}{\beta}^*_{p+d,t} \\ \vdots \\ \underset{\sim}{\beta}^*_{q+1,t} \end{bmatrix} + \begin{bmatrix} I \\ -\Theta_1 \\ \vdots \\ -\Theta_q \end{bmatrix} \underset{\sim}{\alpha}_{t+1}$$

$$y_t = [\underset{\sim}{r}_t' \; \underset{\sim}{O}' \; \dots \; \underset{\sim}{O}'] \begin{bmatrix} \underset{\sim}{\beta}^*_{1,t} \\ \vdots \\ \underset{\sim}{\beta}^*_{q+1,t} \end{bmatrix} + a_t$$

where I_{**} is the $[kq \times kq]$ identity matrix and O_{**} is a $[k \times kq]$ matrix of zeros.

The proof of this theorem follows by simple substitution showing that

$$\Phi(B)\underset{\sim}{\beta}^*_{1t} = \Theta(B)\underset{\sim}{\alpha}_t \quad .$$

3. Several special cases discussed in the literature:

i.) Young [8] considers the case where the parameters follow a first order autoregressive process

$$\begin{cases} \underset{\sim}{\beta}_{t+1} = \Phi\underset{\sim}{\beta}_t + \underset{\sim}{\alpha}_{t+1} \\ y_t = \underset{\sim}{r}_t'\underset{\sim}{\beta}_t + a_t \end{cases} \quad . \tag{3.1}$$

In this case, the updating equations reduce to

$$\hat{\underset{\sim}{\beta}}_t = \Phi\hat{\underset{\sim}{\beta}}_{t-1} + [\underset{\sim}{r}_t'(\Phi P_{t-1}\Phi' + \Sigma_\alpha)\underset{\sim}{r}_t + \sigma_a^2]^{-1}(\Phi P_{t-1}\Phi' + \Sigma_\alpha)\underset{\sim}{r}_t(y_t - \underset{\sim}{r}_t'\Phi\hat{\underset{\sim}{\beta}}_{t-1}) \tag{3.2}$$

and

$$P_t = (\Phi P_{t-1}\Phi' + \Sigma_\alpha)\{I - [\underset{\sim}{r}_t'(\Phi P_{t-1}\Phi' + \Sigma_\alpha)\underset{\sim}{r}_t + \sigma_a^2]^{-1}\underset{\sim}{r}_t\underset{\sim}{r}_t'(\Phi P_{t-1}\Phi' + \Sigma_\alpha)\} \quad . \tag{3.3}$$

ii.) For the special case $\Phi = I$ and $\Sigma_\alpha = 0$ (i.e. $\underset{\sim}{\beta}_{t+1} = \underset{\sim}{\beta}_t$ constant parameters) the updating relations in (3.2) and (3.3) simplify to

$$\hat{\underset{\sim}{\beta}}_t = \hat{\underset{\sim}{\beta}}_{t-1} + (\underset{\sim}{r}_t'P_{t-1}\underset{\sim}{r}_t + \sigma_a^2)^{-1} P_{t-1}\underset{\sim}{r}_t(y_t - \underset{\sim}{r}_t'\hat{\underset{\sim}{\beta}}_{t-1}) \tag{3.4}$$

$$P_t = P_{t-1} - (\underset{\sim}{r}_t'P_{t-1}\underset{\sim}{r}_t + \sigma_a^2)^{-1}P_{t-1}\underset{\sim}{r}_t\underset{\sim}{r}_t'P_{t-1} \quad . \tag{3.5}$$

One recognizes the recursive updating formulae in (3.4) and (3.5) as the recursive updating algorithm for the least squares estimate $\hat{\underset{\sim}{\beta}}_t$ and its covariance matrix P_t given by Plackett [7].

iii.) For the case $k = 1$, $\phi = 1$, $E\alpha^2 = \sigma_\alpha^2$ and $Ea^2 = \sigma_a^2$ the recursive algorithm is given by

$$\hat{\beta}_t = \hat{\beta}_{t-1} + \frac{(P_{t-1} + \sigma_\alpha^2) r_t}{r_t^2 (P_{t-1} + \sigma_\alpha^2) + \sigma_a^2} (y_t - \hat{\beta}_{t-1} r_t) \qquad (3.6)$$

$$P_t = \frac{\sigma_a^2 (P_{t-1} + \sigma_\alpha^2)}{r_t^2 (P_{t-1} + \sigma_\alpha^2) + \sigma_a^2} \quad . \qquad (3.7)$$

We notice that in all these updating formulae the estimate of the parameter at time t is a linear combination of the parameter estimate at time t-1 and the one step ahead forecast error at time t. In (3.4) the parameter is updated by giving equal weights to all the observations. In (3.6) the introduction of σ_α^2 is similar in effect to an exponential data weighting function (Young [8]).

iv.) Box and Jenkins [1] consider a random walk model with added noise ($k = 1$, $\phi = 1$, $r_t = 1$ for all t)

$$\begin{cases} \beta_{t+1} = \beta_t + \alpha_{t+1} \\ y_t = \beta_t + a_t \end{cases} \quad . \qquad (3.8)$$

They show that this model is equivalent to the integrated first order moving average model

$$y_t - y_{t-1} = a_t - \theta a_{t-1} \qquad (3.9)$$

with

$$\theta = \frac{\sigma_a^2}{\sigma_\alpha^2} > 0 \quad .$$

β_{t+1} is the one step ahead forecast $\hat{y}_t(1)$ and is given by an exponential weighted sum of previous observations.

$$\hat{y}_t(1) = (1-\theta) \sum_{j \geq 0} \theta^j y_{t-j} \quad , \tag{3.9}$$

and the forecasts are updated by

$$\hat{y}_t(1) = \hat{y}_{t-1}(1) + (1-\theta)(y_t - \hat{y}_{t-1}(1)) \quad .$$

References

[1] Box, G.E.P. and Jenkins, G.M., Time Series Analysis, Forecasting and Control, Holden-Day; San Francisco, 1970.

[2] Hannan, E.J., Multiple Time Series, John Wiley, New York, 1970.

[3] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," Jour. Basic Eng., 82, 35, 1960.

[4] Kalman, R.E., "Contributions to the Theory of Optimal Control," Bol. de Soc. Mat. Mexicana, 102, 1960.

[5] Kalman, R.E., "The Mathematical Description of Linear Dynamical Systems," SIAM Jour. on Control, 1, 152, 1963.

[6] MacGregor, J.F., "Topics in Control: 1. State Variable Approach to Time Series Representation and Forecasting," Technical Report No. 306, University of Wisconsin, Madison, 1972.

[7] Plackett, R.L., "Some Theorems in Least Squares," Biometrika, 37, 149, 1950.

[8] Young, P.C., "Applying Parameter Estimation to Dynamic Systems--Part I and II," Control Engineering, 119-125, 118-124, 1969.