AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

MONASH

WP NO. 12/90

MONASH
UNIVERSITY

AUSTRALIA

GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS
LIBRARY
WITHDRAWN
NOV 01 1990

WHY KALMAN FILTER?

R. D. Snyder

Working Paper No. 12/90

September 1990

DEPARTMENT OF ECONOMETRICS

ISSN 1032-3813

ISBN 0 86746 947 1

WHY KALMAN FILTER?

R. D. Snyder

Working Paper No. 11/90

September 1990

DEPARTMENT OF ECONOMETRICS, FACULTY OF ECONOMICS, COMMERCE & MANAGEMENT

MONASH UNIVERSITY, CLAYTON, VICTORIA 3168, AUSTRALIA.

ABSTRACT

In this paper the Kalman filter and regression approaches for estimating linear state space models are compared. It is argued that the Kalman filter is no more efficient from a computational point of view, is relatively more complex and hence more obtruse, and that as a consequence its central role in the smoothing, estimation and prediction of time series is questionable.

Keywords: Kalman filter, time series analysis, regression analysis, forecasting, exponential smoothing, maximum likelihood estimation.

INTRODUCTION

The Kalman filter (Kalman, 1960) was originally developed to filter noise from time series. Interest in it for forecasting purposes grew when Harrison and Stephens (1971, 1976) demonstrated its close links with common forms of exponential smoothing (Gardner, 1985). Akaike (1974, 1978) also established that the Kalman filter could be used in conjunction with numerical optimization procedures to find maximum likelihood estimates of the parameters in ARMA models (Box and Jenkins, 1976). Together with subsequent developments for ARIMA and structural models (Harvey and Todd, 1983; Harvey and Pierse, 1984; Ansley and Kohn, 1985;) these works established Kalman filtering as a key method for smoothing, estimating and predicting time series.

Despite its central role, the Kalman filter is often bypassed in forecasting courses, presumably because of its relative complexity. Thus, in this paper, an alternative approach based on conventional regression analysis from a parallel but lesser known stream of literature is explored. It transpires that when properly implemented, the regression approach not only involves lower computational loads than numerically stable versions of the Kalman filter, but is inherently simpler and more transparent. As such it is better, amongst other things, for teaching purposes.

STATE SPACE MODELS

The Kalman filter is used to estimate so-called linear state space models, the coefficients of which evolve over time according to shocked

linear first-order recurrence relationships. An important example is the local linear trend model (Harrison, 1967) underpinning trend corrected exponential smoothing (Holt, 1957). A time series, denoted by z_t at time t, has a local mean μ_t and a growth rate δ_t which evolve according to the recurrence relationships:

$$\mu_t = \mu_{t-1} + \delta_{t-1} + \eta_t \tag{1}$$

$$\delta_t = \delta_{t-1} + \zeta_t \tag{2}$$

while z_t itself is given by

$$z_t = \mu_t + \varepsilon_t . \tag{3}$$

The disturbances η_t, ζ_t and ε_t have zero means and variances σ_η^2 σ_ζ^2, σ_ε^2 respectively and satisfy the following independence conditions:

$$E\eta_t\eta_s = 0 ; \quad E\zeta_t\zeta_s = 0 ; \quad E\varepsilon_t\varepsilon_s = 0 \quad t \neq s$$

$$E\eta_t\zeta_s = 0 ; \quad E\eta_t\varepsilon_s = 0 ; \quad E\zeta_t\varepsilon_s = 0 .$$

Duncan and Horn (1972) showed that state space models can be rewritten in the form of a regression

$$y = X\beta + u, \tag{4}$$

y being a random n-vector, X a fixed $n \times k$ design matrix, β a random k-vector vector of coefficients, and u a random n-vector of disturbances with zero mean and a diagonal variance matrix V. The essential idea is that all state space model coefficients such as μ_t and δ_t can be stacked to form the vector β. For example, the result obtained for the local linear model with a sample $\{z_1, z_2, z_4)$ is shown in Figure 1. The equation for z_3 has been withheld because the

corresponding observation is missing from the sample. Since the model is nonstationary and β has a diffuse prior distribution, the conventional linear least squares method can be employed to yield an estimate of it provided values for the variance ratios $q_1 = \sigma_\eta^2/\sigma_\varepsilon^2$ and $q_2 = \sigma_\zeta^2/\sigma_\varepsilon^2$ are specified. The resulting components $\hat{\mu}_1$, $\hat{\mu}_2$, $\hat{\mu}_3$ and $\hat{\mu}_4$ of $\hat{\beta}$, the estimate of β, can be interpreted as smoothed values of the time series and, in particular, $\hat{\mu}_3$ can be viewed as an estimate of the missing value z_3. Future values of the time series can also be predicted with

$$\hat{y}_{4+\ell} = \hat{\mu}_4 + \ell \cdot \hat{\beta}_4 \qquad \ell = 1,2,\ldots$$

Given the optimal properties of this approach it can be established that these results are identical to those normally obtained with the Kalman filter.

$$\begin{bmatrix} 0 \\ 0 \\ z_1 \\ 0 \\ 0 \\ z_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ z_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -1 & & & & & & & & & \\ & 1 & & -1 & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & 1 & 1 & -1 & & & & & & & \\ & & & 1 & & -1 & & & & & & \\ & & & & 1 & & & & & & & \\ & & & & 1 & 1 & -1 & & & & & \\ & & & & & 1 & & -1 & & & & \\ & & & & & & 1 & 1 & -1 & & & \\ & & & & & & & 1 & & -1 & & \\ & & & & & & & & 1 & & & \end{bmatrix} \begin{bmatrix} \mu_0 \\ \delta_0 \\ \mu_1 \\ \delta_1 \\ \mu_2 \\ \delta_2 \\ \mu_3 \\ \delta_3 \\ \mu_4 \\ \delta_4 \\ \\ \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \zeta_1 \\ \varepsilon_1 \\ \eta_2 \\ \zeta_2 \\ \varepsilon_2 \\ \eta_3 \\ \zeta_3 \\ \eta_4 \\ \zeta_4 \\ \varepsilon_4 \end{bmatrix}$$

Figure 1. Regression Formulation of Local Trend Model

MAXIMUM LIKELIHOOD ESTIMATION

In most applications parameters such as the variance ratios q_1 and q_2 are themselves unknown and maximum likelihood estimates of them must be obtained. However, the likelihood function is undefined for non-stationary models because β has a diffuse prior distribution. An alternative strategy is to use the marginal likelihood function (Kalbfleish and Sprott, 1970) without the nuisance coefficients β. It can be established that this marginal likelihood function can be written in terms of the one-step ahead prediction errors e_i as follows:

$$\ell = (2\pi\sigma_\varepsilon^2)^{-(n-k)/2}\left(\prod_{i=k+1}^{n} h_i\sigma_\varepsilon^2\right)^{-1/2} \exp\left\{-\sum_{i=k+1}^{n} e_i^2/(2h_i\sigma_\varepsilon^2\right\} \qquad (5)$$

where the h_i are scaled mean squared errors of the e_i. This result is sometimes referred to as the 'prediction error decomposition of the likelihood function' (Schweppe, 1965).

Procedures to find the maximum likelihood estimates require the evaluation of (5) for many trial values of the parameters. The Kalman filter is conventionally employed to generate the associated one-step ahead prediction errors e_i and their scaled mean squared errors h_i for each trial. However, it can be shown that these same crucial quantities can be obtained as a by-product of the orthogonalization procedure based on fast Givens transformations (Gentleman 1973; Stirling, 1981) when it is used to compute the least squares estimate $\hat{\beta}$ in a regression model. As such, the Kalman filter can be replaced by the regression approach in maximum likelihood procedures.

CONCLUDING REMARKS

It has been briefly demonstrated how a regression approach can be used instead of the more complex Kalman filter in smoothing, estimation and prediction problems in time series analysis. Although the resulting regressions are often quite large, a study by Page and Saunders (1977) indicates that efficient least squares routines which automatically exploit sparsity have computational loads and storage requirements that are comparable to those associated with numerically stable versions of the Kalman filter. Although V^{-1} may not exist in special but important cases where some of the variances equal zero, the associated relationships reduce to linear restrictions and these can be handled automatically by Stirlings (1981) least squares method. And although the approach in this paper was concerned with a nonstationary model, it is also readily adapted to cope with stationary cases where β has an informative prior distribution.

Interestingly, the regression approach works better in non-stationary cases such as the local trend model. The Kalman filter, when initiated with a diffuse prior distribution can be numerically unstable and special, more complex algorithms (Ansley and Kohn, 1985; Snyder, 1988) with relatively high computational loads are required to guard against such a possibility. It transpires that during initialization, the regression approach involves lower than normal computational loads so that its performance definitely is superior in these situations. Overall, then, the regression approach is not only inherently simpler and more transparent, but the evidence suggests that it can involve lower computational loads. As such it leads us to ask: "Why Kalman Filter?"

References

Akaike, H., 1974, "Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes", *Annals of the Institute of Statistical Mathematics*, 26, 363-387.

Akaike, H., 1978, "Covariance matrix computation of the state variable of a stationary Gaussian process", *Annals of the Institute of Statistical Mathematics*, 30, 499-504.

Ansley, C.F. and R. Kohn, 1985, "Estimation, filtering, and smoothing in state space models with incompletely specified initial conditions", *The Annals of Statistics*, 13, 1286-1316.

Box G.E.P. and G.W. Jenkins, 1976, *Time Series Analysis: Forecasting and Control*, (Holden-Day, San Francisco).

Duncan, D.B. and S.D. Horn, 1972, "Linear dynamic recursive estimation from the view point of regression analysis", *Journal of the American Statistical Association*, 67, 815-821.

Gardner, E.S., 1985, "Exponential smoothing: the state of the art", *Journal of Forecasting*, 4, 1-28.

Gentleman, W.M., 1973, "Least squares computations by Givens transformations with square roots," *Inst. Math. Applics.*, 12, 329-336.

Harrison, P.J., "Exponential smoothing and short term sales forecasting", *Management Science*, 13, 821-842.

Harrison, P.J. and C.F. Stephens, 1971, "A Bayesian approach to short-term forecasting", *Operational Research Quarterly*, 22, 341-362.

Harrison, P.J. and C.F. Stephens, 1976, "Bayesian forecasting", *Journal of the Royal Statistical Society, B*, 38, 205-228.

Harvey, A.C. and R.G. Pierse, 1984, "Estimating missing observations in economic time series", *Journal of the American Statistical Association*, 79, 125-131.

Harvey, A.C. and P.H.J. Todd, 1983, "Forecasting Economic Time Series with Structural and Box-Jenkins Models: A Case Study," *Journal of Business and Economic Statistics*, 1, 299-316.

Holt, C.E., 1957, "Forecasting Trends and Seasonals by Exponentially Weighted Averages", *ONR Memorandum No. 52*, Carnegie Institute of Technology.

Kalbfleish, J.D. and D.A. Sprott, 1970, "Application of Likelihood Methods to Models Involving Large Numbers of Parameters," *Journal of the Royal Statistical Society B*, 32, 175-192.

Kalman, R.E., 1960, "A new approach to linear filtering and prediction problems", *Trans ASME Series D, Journal of Basic Engineering*, 82, 35-45.

Page, C.C. and M.A. Saunders, 1977, "Linear least squares estimation of discrete linear dynamic systems using orthogonal transformations", *SIAM Journal of Numerical Analysis*, 15, 180-193.

Schweppe, F.C. (1965), "Evaluation of the likelihood function for Gaussian signals", *IEEE Transactions on Information Theory*, 11, 61-70.

Snyder, R.D., 1988, "Computational aspects of Kalman filtering with a diffuse prior distribution", *Journal of Statistical Computation and Simulation*, 29, 77-86.

Stirling, W.D., 1981, "Least squares subject to linear constraints", *Applied Statistics*, 30, 204-212.

MONASH UNIVERSITY

DEPARTMENT OF ECONOMETRICS

WORKING PAPERS

11/89 R.D. Snyder. "A Computerized System for Forecasting Spare Parts Sales: A Case Study".

12/89 Maxwell L. King and Ping X. Wu. "Small-Disturbance Asymptotics and the Durbin-Watson and Related Tests in the Dynamic Regression Model".

13/89 W.T.M. Dunsmuir and R.D. Snyder. "ABC Analysis in Inventory Control - The Issue of Stability".

1990

1/90 P. Burridge. "The Functional Central Limit Theorem: An Introductory Exposition with Application to Testing for Unit Roots in Economic Time Series".

2/90 Maxwell L. King and Ping X. Wu. "Locally Optimal One-Sided Tests for Multiparameter Hypotheses."

3/90 Grant H. Hillier. "On Multiple Diagnostic Procedures for the Linear Model."

4/90 Jean-Marie Dufour and Maxwell L. King. "Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR(1) Errors."

5/90 Keith R. McLaren. "A Reappraisal of the Neoclassical Approach to Modelling Business Investment."

6/90 Francis Vella. "Non-Wage Benefits in a Simultaneous Model of Wages and Hours: Labor Supply Functions of Young Females."

7/90 Francis Vella. "A Simple Estimator for Simultaneous Models with Censored Endogenous Regressors."

8/90 Nicola J. Crichton and Timothy R.L. Fry. "An Analysis of the Effect of an Offender's Employment Status on the Type of Sentence Chosen by the Magistrate."

9/90 Paramsothy Silvapulle and Maxwell L. King. "Testing Moving Average Against Autoregressive Disturbances in the Linear Regression Model."

10/90 Asraul Hoque and Brett A. Inder. "Structural Unemployment in Australia."

11/90 Maxwell L. King, Chandra Shah and Kees Jan van Garderen. "Tutoring in Economic Statistics: The Monash experience."