Volume Title: Annals of Economic and Social Measurement, Volume 6, number 4

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm77-4

Publication Date: October 1977

Chapter Title: A Note on Optimal Smoothing for Time Varying Coefficient Problems

Chapter Author: Thomas F. Cooley, Barr Rosenberg

Chapter URL: http://www.nber.org/chapters/c10528

Chapter pages in book: (p. 453 - 456)

Annals of Economic and Social Measurement, 6/4, 1977

A NOTE ON OPTIMAL SMOOTHING FOR TIME VARYING COEFFICIENT PROBLEMS

BY THOMAS F. COOLEY, BARR ROSENBERG, AND KENT D. WALL

An algorithm is presented which provides a complete solution to the optimal estimation problem for time-varying parameters when no proper prior distribution is specified. The key ideas involve a combination of the information-form Kalman filter with the two-filter interpretation of the optimal smoother. The algorithm produces efficient estimates of the parameter trajectories over the entire sample and is equally applicable when a proper prior distribution has been specified.

1. INTRODUCTION

In recent years, a significant body of literature has appeared that is addressed to the problem of estimating time varying regression coefficients (Pagan [1974], Rosenberg [1973a, 1973b], Sarris [1973], Cooley and Prescott [1973, 1976]). The most fashionable approach to these problems has been to apply Kalman filtering theory to the estimation of the coefficient trajectories. The application of filtering theory, however, requires either a priori knowledge or the estimation of an initial coefficient vector.

One approach to the starting problem is to use an empirical Bayesian technique like that suggested by Kaminsky et al. [1974] or Garbade [1975]. This approach involves selecting a subsample of the observations and using them to compute a prior distribution.[1] This, in turn, is used to initialize the filter for the remaining observations.

The components of a theoretically complete solution to the initialization and smoothing problems have existed in the control theory literature for some time.[2] Fraser [1967], Rosenberg [1967, 1968], and Fraser and Potter [1968] developed solutions to the smoothing problem which do not require proper prior distributions for the parameter vector. The purpose of this note is to integrate these components and provide an algorithm for optimal smoothing which does not require dichotomization of the sample. The approach is a simple one based on a combination of the information form of the filter and the two-filter interpretations of the smoother.

[1] If the transition and covariance parameters of the underlying coefficient process are known, the empirical Bayes approach has computational advantages. However, this is unlikely to be the case in most applications.

[2] Smoothed estimates are those that use all of the information in the sample to estimate the coefficient at each point in time.

2. The Oilmai Smoothing Algorithm

The time varying coefficient problem is characterized by a regression equation:

$$y_t = x_t'\beta_t + e_t, \tag{1}$$

$$\beta_{t+1} = \Phi\beta_t + u_t. \tag{2}$$

The variables y and x represent the observables of the system, Φ is a $k \times k$ matrix which governs the transitions of the k component parameter vector β. The disturbances e_t and u_t are independently and identically distributed random variables with zero means and covariance matrices σ^2 and Q, respectively. The problem is to obtain estimates of the β_t based on the observations $[y_1, \ldots, y_T]$. Let $b_{t|t^*}$ be an estimate of β_t based on observations $[y_1, \ldots, y_{t^*}]$ and let $P_{t|t^*}$ be the variance-covariance matrix of the estimated coefficients.

$$P_{t|t^*} = E[(\beta_t - b_{t|t^*})(\beta_t - b_{t|t^*})']. \tag{3}$$

Information Filters

We now let $P_{t|t-1}$ and $P_{t|t}$ be the variance-covariance matrices before and after making an observation (the "predicted" and "corrected" matrices) and define the corresponding "information" matrices $H_{t|t-1} = P\dagger_{t|t-1}$ and $H_{t|t} = P\dagger_{t|t}$, where "$\dagger$" denotes the pseudoinverse. Finally, define the forward information variables:

$$\begin{aligned} f_{t|t-1} &= H_{t|t-1}b_{t|t-1} \\ f_{t|t} &= H_{t|t}b_{t|t}, \end{aligned} \tag{4}$$

which play the role of the estimates in the information form of the filter. Estimation proceeds by assuming a diffuse prior distribution for the β_t's, expressed by initializing the problem with

$$H_{1|0} = 0 \quad \text{and} \quad f_{1|0} = 0.$$

Letting $\bar{Q} = \Phi^{-1}Q$, the prediction and correction formulas are:[3]

Prediction

$$\bar{H}_{t+1} = (\Phi^{-1})'H_{t|t}\Phi^{-1} \tag{5}$$

[3]Equation (6) has appeared in the literature in slightly different form:

$$K_{t+1} = [I - \bar{H}_{t+1}(\bar{Q}^{-1} + \bar{H}_{t+1})^{-1}](\Phi^{-1})'.$$

This has led some people to the conclusion that the information form of the filter cannot accomodate singular Q matrices. A similar observation applies to equation (11).

(6) $K_{t-1} = [I - \bar{H}_{t-1}(I + \bar{Q}\bar{H}_{t-1})^{-1}\bar{Q}](\Phi^{-1})' = (\Phi' + H_{t-1|t-1}\bar{Q})^{-1}$

(7) $$H_{t|t-1} = K_{t-1}H_{t-1|t-1}\Phi^{-1}$$

(8) $$f_{t|t-1} = K_{t-1}f_{t-1|t-1}$$

Correction

(9) $$f_{t|t} = f_{t|t-1} + x_t y_t/\sigma^2$$

(10) $$H_{t|t} = H_{t|t-1} + x_t x_t'/\sigma^2.$$

In addition to the forward filtered estimates, we require the backward or reverse time filtered estimates which evolve in $\tau = T - t$. Denote by $G_{t|t+1}$ and $G_{t|t}$, the predicted and corrected reverse time information matrices, and let the corresponding filtered variables be $r_{t|t+1}$ and $r_{t|t}$. The reverse time filter is initialized with:

$$G_{T|T+1} = 0 \quad \text{and} \quad r_{T|T+1} = 0.$$

and the prediction and correction formulas are:

Prediction

(11) $J_{t+1} = \Phi'[I - G_{t+1|t+1}(I + QG_{t+1|t+1})^{-1}Q] = \Phi'(I + G_{t+1|t+1}Q)^{-1}$

(12) $$G_{t|t+1} = J_{t+1}G_{t+1|t+1}\Phi$$

(13) $$r_{t|t+1} = J_{t+1}r_{t+1|t+1}.$$

Correction

(14) $$r_{t|t} = r_{t|t+1} + x_t y_t/\sigma^2$$

(15) $$G_{t|t} = G_{t|t+1} + x_t x_t'/\sigma^2.$$

The Optimal Smoother

It is clear that in econometric applications interest should focus on the most efficient estimates of the parameters which use all of the information available. These are the smoothed estimates, $b_{t|T}$, and they may be computed as a weighted combination of $f_{t|t}$ and $r_{t|t}$ (Liebelt [1967]):

(16) $$P_{t|T} = [H_{t|t} + G_{t|t+1}]^{-1}$$

(17) $$b_{t|T} = P_{t|T}[f_{t|t} + r_{t|t+1}].$$

The algorithm outlined above provides the best estimates of the parameter trajectories and their associated variances, without resorting to ad hoc procedures. It also is equally applicable when a proper prior distribution is specified, since then one merely sets $H_{1|0} = P_{1|0}^{-1}$ and $f_{1|0} = H_{1|0}b_{1|0}$.

NBER and University of California, Santa Barbara;
University of California, Berkeley;
NBER and Systems Control, Inc.

Submitted July 1976
Revised April 1977

REFERENCES

Cooley, T. F., and Prescott, E. C. 1973. "Varying Parameter Regression: A Theory and Some Applications." *Annals of Economic and Social Measurement* 2, no. 4 (October).

———. 1976. "Estimation in the Presence of Stochastic Parameter Variation." *Econometrica* 44, no. 1 (January).

Fraser, D. C. 1967. "A New Technique for the Optimal Smoothing of Data." MIT Instrumentation Laboratory Report T-474, January.

Fraser, D. C., and Potter, J. E. 1967. "The Optimum Linear Smoother as a Combination of Two Optimum Linear Filters." *IEEE Transactions on Automatic Control*, vol. AC14, no. 4.

Garbade, K. 1975. "The Initialization Problem in Variable Parameter Regression." Working Paper, New York University.

Kaminski, P. G., Bryson, A. E., and Schmidt, S. F. 1974. "Discrete Square Root Filtering: A Survey of Current Techniques." *IEEE Transactions on Automatic Control*, vol AC16, no. 6 (December).

Liebelt, P. 1967. *An Introduction to Optimal Estimation.* Reading: Mass.: Addison-Wesley.

Pagan, A. 1974. "A Note on the Extraction of Components from Time Series." *Econometrica* 43, no. 6.

Rosenberg, B. M. 1967. "Regression in the Presence of Stochastically Varying Parameters." Unpublished manuscript. School of Business Administration, University of California, Berkeley. Presented at the Econometric Society, Washington, D.C. (December).

———. 1968. "Varhing Parameter Estimation." Unpublished Ph.D. dissertation, Department of Economics, Harvard University.

———. 1973a. "A Survey of Stochastic Parameter Regression." *Annals of Economic and Social Measurement* 3, no. 4 (October).

———. 1973b. "The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression." *Annals of Economic and Social Measurement* 3, no. 4 (October).

Sarris, A. 1973. "A Bayesian Approach to the Estimation of Time Varying Regression Coefficients." *Annals of Economic and Social Measurement* 3, no. 4 (October).