NBER WORKING PAPER SERIES

# A NOTE ON OPTIMAL SMOOTHING FOR TIME VARYING COEFFICIENT PROBLEMS

Thomas F. Cooley*
Kent D. Wall**

Working Paper No. 128

COMPUTER RESEARCH CENTER FOR ECONOMICS AND MANAGEMENT SCIENCE
National Bureau of Economic Research, Inc.
575 Technology Square
Cambridge, Massachusetts 02139



March 1976



Preliminary: Not for quotation



*NBER Computer Research Center and Tufts University. Research supported in part by National Science Foundation Grant GJ-1154X3 to the National Bureau of Economic Research, Inc.

**NBER Computer Research Center. Research supported in part by National Science Foundation Grant GJ-1154X3 to the National Bureau of Economic Research, Inc.

## Abstract

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.

TABLE OF CONTENTS

1. INTRODUCTION . . . . . . . . . . . . . . 1

2. THE OPTIMAL SMOOTHING ALGORITHM . . 2

FOOTNOTES. . . . . . . . . . . . . . . . . 6

REFERENCES . . . . . . . . . . . . . . . . 7

## 1. INTRODUCTION

In recent years a significant body of literature has appeared which is addressed to the problem of estimating time varying regression coefficient (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.

Rosenberg [1968, 1973] was the first to develop an algorithm for conditional estimation of the initial coefficients, and present formulae for the terminal period "smoothed" estimates of the coefficients.[1] His results are incomplete, however, in that they do not contain equations for obtaining the smoothed coefficient trajectories with their associated variance covariance matrices. These estimates cannot be obtained from Rosenberg's analysis for two reasons. The first is that he does not provide estimates of the variance-covariance matrix of the initial parameters. The second, and more important, reason is that the form of the smoothing equations he presents are inappropriate because they rely on corrections to the filtered trajectories which in themselves are nonexistent in the absence of proper prior distributions for the initial coefficients.

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.[2] 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 exist in the control theory literature. 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 filter (Fraser [1967]) and the two filter interpretations of the smoother (Fraser and Potter [1968]).

## 2. THE OPTIMAL SMOOTHING ALGORITHM

The time varying coefficient problem is characterized by a regression equation

$$(1) \qquad y_t = x_t'\beta_t + e_t ,$$

and an equation describing the evolution of the coefficients over time

$$(2) \qquad \beta_{t+1} = \Phi\beta_t + u_t .$$

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

(3) $$P_{t/t*} = E[(\beta_t - b_{t/t*})(\beta_t - b_{t/t*})']$$

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 $M_{t/t-1} = P^{-1}_{t/t-1}$ and $M_{t/t} = P^{-1}_{t/t}$. Finally, define the forward information variables

(4) $$f_{t/t-1} = M_{t/t-1}\, b_{t/t-1}$$

$$f_{t/t} = M_{t/t}\, b_{t/t}$$

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

$$M_{1/0} = 0 \text{ and } f_{1/0} = 0.$$

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

Prediction

(5) $$\overline{M}_{t-1} = (\Phi^{-1})'\, M_{t-1/t-1}$$

(6) $$D_{t-1} = [I - \overline{M}_{t-1}(I + \overline{Q}\,\overline{M}_{t-1})^{-1}\overline{Q}]\,(\Phi^{-1})'$$

(7) $$M_{t/t-1} = D_{t-1}\, M_{t-1/t-1}\, \Phi^{-1}$$

(8) $$f_{t/t-1} = D_{t-1}\, f_{t-1/t-1}$$

Correction

(9) $$f_{t/t} = f_{t/t-1} + x_t' \, y_t/\sigma^2$$

(11) $$M_{t/t} = M_{t/t-1} + x_t' \, x_t/\sigma^2$$

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

$$N_{1/0} = 0 \quad r_{1/0} = 0 \; ,$$

and the prediction and correction formulae are:

Prediction

(12) $$D_{\tau-1} = \Phi'[I-N_{\tau-1/\tau-1}(I+QN_{\tau-1/\tau-1})^{-1}Q]$$

(13) $$N_{\tau/\tau-1} = D_{\tau-1} \, N_{\tau-1/\tau-1}\Phi$$

(14) $$r_{\tau/\tau-1} = D_{\tau-1} \, r_{\tau-1/\tau-1}$$

Correction

(15) $$r_{\tau/\tau} = r_{\tau/\tau-1} + X_\tau' \, y_\tau/\sigma^2$$

(16) $$N_{\tau/\tau} = N_{\tau/\tau-1} + X_\tau' \, X_\tau/\sigma^2$$

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, $b_{\tau/T}$, and they may be computed as a weighted combination of $f_{t/t}$ and $r_{t/t}$ (Liebelt [1967]):

(17) $$P_{t/T} = [M_{t/t} + N_{t/t+1}]^{-1}$$

(18) $$b_{t/T} = P_{t/T}[f_{t/t} + r_{t/t+1}]$$

The algorithm outlined above provides the best estimates of the parameter trajectories and their associated variances, without resorting to empirical Bayesian procedures. It also is equally applicable when a proper prior distribution is specified since then one merely sets $M_{1/0}=P^{-1}_{1/0}$ and $f_{1/0}=M_{1/0}\ b_{1/0}$.

FOOTNOTES

1. Smoothed estimates are those which use all of the information in the sample to estimate the coefficient at each point in time.

2. If the transition and covariance parameters of the underlying coefficient process are known, the empirical Bayes approach has computational advantages. This, however, is unlikely to be the case in most applications.

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

$$D_{t-1} = [I - \overline{M}_{t-1}(\overline{Q}^{-1} + \overline{M}_{t-1})^{-1}] (\Phi^{-1})'.$$

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 (12).

# REFERENCES


[1] Cooley, T. F. and E. C. Prescott [1973], "Varying Parameter Regression: A Theory and Some Applications", Annals of Economic and Social Measurement, Vol. 2, No. 4, October.

[2] _________ and _________ [1976], "Estimation in the Presence of Stochastic Parameter Variation", Econometrica, Vol. 44, No. 1, January.

[3] Fraser, D. C. [1967], "A New Technique for the Optimal Smoothing of Data", MIT Instrumentation Lab., Rep. T-474, January.

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

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

[6] Kaminski, P. G., A. E. Bryson and S. F. Schmidt [1974], "Discrete Square Root Filtering: A Survey of Current Techniques", IEEE Trans. Automatic Control, Vol. AC-16, No. 6, December.

[7] Pagan, A. [1974], "A Note on the Extraction of Components from Time Series", Econometrica, Vol. 43, No. 6.

[8] Rosenberg [1968], "Varying Parameter Estimation", unpublished Ph.D. Dissertation, Department of Economics, Harvard University.

[9] Rosenberg, B. M. [1973a], "A Survey of Stochastic Parameter Regression", Annals of Economic and Social Measurement, Vol. 3, No. 4, October.

[10] Rosenberg, B. [1973b], "The Analysis of A Cross Section of Time Series by Stochastically Convergent Parameter Regression", Annals of Economic and Social Measurement, Vol. 3, No. 4, October.

[11] Sarris, A [1973], "A Bayesian Approach to the Estimation of Time Varying Regression Coefficients", Annals of Economic and Social Measurement, Vol. 3, No. 4, October.