Forecasting, structural time series models and the Kalman filter

Forecasting, structural time series models and the Kalman filter

ANDREW C. HARVEY
Department of Statistical and Mathematical Sciences, The London School of Economics and Political Science

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011–4211, USA
10 Stamford Road, Oakleigh, VIC 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1989
Reprinted 1990
First paperback edition 1990
Reprinted 1991, 1992, 1994, 1996, 1999, 2001

British Library Cataloguing in Publication data
Harvey, Andrew C.
Forecasting, structural time series models
and the Kalman filter
1. Time series. Structural models
I. Title
519.5′5

Library of Congress Cataloguing in Publication data
Harvey, Andrew C.
Forecasting, structural time series models and the Kalman filter/
Andrew C. Harvey.
p. cm.
Bibliography.
Includes index.
ISBN 0-521-32196-4
1. Time-series analysis. 2. Kalman filtering. I. Title.
QA280.H38 1990
519.5′5–dc20 89-31417 CIP

ISBN 0 521 32196 4 hardback
ISBN 0 521 40573 4 paperback

Transferred to digital printing 2003

Contents

Figures

Acknowledgement

Table 5.3.1 is reproduced with the permission of the *Journal of the American Statistical Association.*

Tables 5.3.2 and 5.3.3 are reproduced with the permission of the Biometrika Trustees.

Preface

Structural time series models are models which are formulated directly in terms of components of interest. They have a considerable intuitive appeal, particularly for economic and social time series. Furthermore, they provide a clear link with regression models, both in their technical formulation and in the model selection methodology which they employ. The potential of such models is only now beginning to be realised, and it seems to be an appropriate time to write a book which provides a unified view of the area and points the direction towards future research.

The Kalman filter plays a fundamental role in handling structural time series models. This technique was originally developed and exploited in control engineering. It has been increasingly used in areas such as economics, and a good deal of work has been done modifying it for use with small samples. Chapter 3 brings these methods together, and it can be read independently of the material on structural time series models. For those who are primarily interested in carrying out applied work with structural time series models, it should perhaps be stressed that the Kalman filter is simply a statistical algorithm, and it is only necessary to understand what the filter does, rather than how it does it. The same is true of the frequency-domain methods which can be used to construct the likelihood function.

The inclusion of 'forecasting' in the title of the book is perhaps a little rash. It is always very difficult to predict the future on the basis of the past. Indeed it has been likened to driving a car blindfolded while following directions given by a person looking out of the back window. Nevertheless, if this is the best we can do, it is important that it should be done properly, with an appreciation of the potential errors involved. In this way it should at least be possible to negotiate straight stretches of road without a major disaster. Too many forecasting procedures seem to attribute the person in the back seat with supernatural powers when in fact his behaviour is more consistent with that of someone who is mildly inebriated.

Structural time series models are appropriate to many subjects, including economics, sociology, management science and operational research, geography, meteorology and engineering. The emphasis in the book is primarily on economic time series, but examples will be found on such diverse topics as rainfall in Brazil, purse snatching in Chicago, telephone calls from Australia and the effect of the seat belt law in Great Britain. Most of the calculations reported can be carried out on an IBM PC using the program 'STAMP', details of which can be found at the end of chapter 1.

It is assumed that the reader is familiar with statistical theory and with the basic ideas of time series analysis and regression. Certain sections, primarily those dealing with dynamic regression and simultaneous equation systems, presuppose a knowledge of econometrics. Indeed one of the aims of the book is to provide a framework which includes behavioural econometric models as well as the simplest kinds of univariate time series models. On the time series side, it is not necessary to be an expert on what is popularly called Box-Jenkins modelling. In fact this could conceivably be a disadvantage since structural time series modelling starts from a somewhat different point and in doing so challenges some of the underlying assumptions of the Box-Jenkins approach. Nevertheless Box-Jenkins ARIMA models and univariate structural time series models do have a good deal in common in that they are both based on statistical models and are normally handled by classical procedures. As regards mathematics, I assume that the reader is familiar with linear algebra and the calculus but not with more advanced techniques such as measure theory. No attempt is made to provide rigorous proofs on topics such as the asymptotic properties of estimators. The emphasis is on the development of models which can be used in practice, and the way in which such models can be selected.

Many of my students at the LSE have helped in the development and implementation of the ideas presented in this book. I would particularly like to mention Phillipa Todd, Carla Inclan, Luiz Hotta, Lorenzo Figliuoli, Javier Fernández, Neil Shephard, Mariane Streibel, Ester Ruiz, Pablo Marshall and Christiano Fernandes. Simon Peters, who was my research assistant for three years, also played a major role in this respect. Furthermore it was he who developed the software for the STAMP program, and he must take a great deal of the credit for actually putting the methods of structural time series modelling into practice. I have also benefitted considerably from discussions with my colleagues at the LSE, and I am particularly grateful for the advice I have received from Jim Durbin and Peter Robinson. My conversations with Jim Durbin have been especially valuable in providing me with insight into the development of the subject of time series and in convincing me that the structural approach to time series modelling is indeed the best way to proceed. Outside the LSE, I have gained enormously from contacts with Piet de Jong and Jim Stock. In fact many of the ideas on multivariate models and continuous time have come from joint work carried out with Jim Stock. In addition, the section on time-varying parameters owes a good deal to discussions with Bill Brainard.

I am grateful to Alicia Kacperek, Christine Wills, Elaine Hartwell and, most notably, Lavinia Harvey for typing assistance and to the Economic and Social Research Council for financial support. Parts of the book were written during visits to IMPA, the mathematics institute in Rio de Janeiro, and to the Australian National University in Canberra. Leigh Roberts, Jim Durbin, Mark Watson, Andy Tremayne and Johannes Ledolter read through various parts of the book and made valuable comments. I am very grateful to them, but they are of course, absolved from blame for any errors and from any responsibility for the views expressed in the book.

Notation and conventions

The following conventions are adopted in the text.

(a) Matrices and vectors are printed in bold type. Vectors are denoted by lower-case letters, with a prime indicating a row vector.

(b) Greek letters are used to denote parameters or states. The nearest corresponding Latin letter denotes the linear estimator. A tilde (˜) over a Greek letter indicates a maximum likelihood estimator of a parameter or the minimum mean square estimator of a state.

(c) A single time subscript on an estimate or estimator indicates the use of information up to, and including, that time subscript. The conditional notation, for example $t|\tau$, for a subscript denotes an estimate or estimator of a quantity at time t based on information available at time τ, where τ may be less than or greater than t.

Writing $\hat{y}_{t+j|t}$ denotes an estimate of y_{t+j} made on the basis of information available at time t. A tilde rather than a hat (ˆ) indicates that the estimate is, in some sense, optimal.

(d) The notation 'log' always denotes the natural logarithm.

Abbreviations

a.c.f.	autocovariance, or autocorrelation, function
a.c.g.f.	autocovariance generating function
AIC	Akaike information criterion
AN	asymptotically normal
AR	autoregressive
ARCH	autoregressive conditional heteroscedasticity
ARE	algebraic Riccati equation
ARIMA	autoregressive integrated moving average
ASE	asymptotic standard error
BIC	Bayes information criterion
BSM	basic structural model
CSS	conditional sum of squares
CUSUM	cumulative sum
DLS	discounted least squares
D-W	Durbin-Watson
ESS	extrapolative sum of squares
EWMA	exponentially weighted moving average
FD	frequency domain
FIML	full information maximum likelihood
GLIM	general linear model
GLS	generalised least squares
IV	instrumental variable
KF	Kalman filter
LBI	locally best invariant
LIML	limited information maximum likelihood
LM	Lagrange multiplier
LR	likelihood ratio
MA	moving average
MD	mean deviation
ML	maximum likelihood
MMSE	minimum mean square estimator or estimate
MMSLE	minimum mean square linear estimator
MPI	most powerful invariant

MSE	mean square error
$NID(\mu, \sigma^2)$	normally and independently distributed with mean μ and variance σ^2
OLS	ordinary least squares
p.d.f.	probability density function
p.e.v.	prediction error variance
p.(s.)d.	positive (semi-) definite
RMSE	root mean square error
SEM	simultaneous equation model
s.g.f.	spectral generating function
SSE	sum of squared errors
SSF	state space form
TD	time domain
UCARIMA	unobserved components ARIMA
VAR	vector autoregression
VARMA	vector autoregressive moving average
WLS	weighted least squares