

Time Series Analysis by State Space Methods

Second Edition

J. Durbin

London School of Economics
and Political Science
and University College London

S. J. Koopman

Vrije Universiteit Amsterdam

OXFORD
UNIVERSITY PRESS

Contents

Introduction	1
1.1 Basic ideas of state space analysis	1
1.2 Linear models	1
1.3 Non-Gaussian and nonlinear models	3
1.4 Prior knowledge	4
1.5 Notation	4
1.6 Other books on state space methods	5
1.7 Website for the book	6

PART I THE LINEAR STATE SPACE MODEL

2. Local level model	9
2.1 Introduction	9
2.2 Filtering	11
2.2.1 The Kalman filter	11
2.2.2 Regression lemma	13
2.2.3 Bayesian treatment	14
2.2.4 Minimum variance linear unbiased treatment	15
2.2.5 Illustration	16
2.3 Forecast errors	17
2.3.1 Cholesky decomposition	17
2.3.2 Error recursions	18
2.4 State smoothing	19
2.4.1 Smoothed state	19
2.4.2 Smoothed state variance	21
2.4.3 Illustration	23
2.5 Disturbance smoothing	23
2.5.1 Smoothed observation disturbances	23
2.5.2 Smoothed state disturbances	24
2.5.3 Illustration	25
2.5.4 Cholesky decomposition and smoothing	25
2.6 Simulation	26
2.6.1 Illustration	27
2.7 Missing observations	28
2.7.1 Illustration	30

2.8	Forecasting	30
2.8.1	Illustration	31
2.9	Initialisation	32
2.10	Parameter estimation	34
2.10.1	Loglikelihood evaluation	34
2.10.2	Concentration of loglikelihood	36
2.10.3	Illustration	37
2.11	Steady state	37
2.12	Diagnostic checking	38
2.12.1	Diagnostic tests for forecast errors	38
2.12.2	Detection of outliers and structural breaks	39
2.12.3	Illustration	40
2.13	Exercises	41
3.	Linear state space models	43
3.1	Introduction	43
3.2	Univariate structural time series models	44
3.2.1	Trend component	44
3.2.2	Seasonal component	45
3.2.3	Basic structural time series model	46
3.2.4	Cycle component	48
3.2.5	Explanatory variables and intervention effects	49
3.2.6	STAMP	51
3.3	Multivariate structural time series models	51
3.3.1	Homogeneous models	51
3.3.2	Common levels	52
3.3.3	Latent risk model	52
3.4	ARMA models and ARIMA models	53
3.5	Exponential smoothing	57
3.6	Regression models	60
3.6.1	Regression with time-varying coefficients	60
3.6.2	Regression with ARMA errors	60
3.7	Dynamic factor models	61
3.8	State space models in continuous time	62
3.8.1	Local level model	63
3.8.2	Local linear trend model	64
3.9	Spline smoothing	66
3.9.1	Spline smoothing in discrete time	66
3.9.2	Spline smoothing in continuous time	68
3.10	Further comments on state space analysis	69
3.10.1	State space versus Box-Jenkins approaches	69
3.10.2	Benchmarking	71
3.10.3	Simultaneous modelling of series from different sources	73
3.11	Exercises	74

4.	Filtering, smoothing and forecasting	76
4.1	Introduction	76
4.2	Basic results in multivariate regression theory	77
4.3	Filtering	82
4.3.1	Derivation of the Kalman filter	82
4.3.2	Kalman filter recursion	85
4.3.3	Kalman filter for models with mean adjustments	85
4.3.4	Steady state	86
4.3.5	State estimation errors and forecast errors	86
4.4	State smoothing	87
4.4.1	Introduction	87
4.4.2	Smoothed state vector	88
4.4.3	Smoothed state variance matrix	90
4.4.4	State smoothing recursion	91
4.4.5	Updating smoothed estimates	91
4.4.6	Fixed-point and fixed-lag smoothers	92
4.5	Disturbance smoothing	93
4.5.1	Smoothed disturbances	93
4.5.2	Smoothed disturbance variance matrices	95
4.5.3	Disturbance smoothing recursion	96
4.6	Other state smoothing algorithms	96
4.6.1	Classical state smoothing	96
4.6.2	Fast state smoothing	97
4.6.3	The Whittle relation between smoothed estimates	98
4.6.4	Two filter formula for smoothing	98
4.7	Covariance matrices of smoothed estimators	100
4.8	Weight functions	104
4.8.1	Introduction	104
4.8.2	Filtering weights	105
4.8.3	Smoothing weights	106
4.9	Simulation smoothing	107
4.9.1	Simulation smoothing by mean corrections	107
4.9.2	Simulation smoothing for the state vector	108
4.9.3	de Jong-Shephard method for simulation of disturbances	109
4.10	Missing observations	110
4.11	Forecasting	112
4.12	Dimensionality of observational vector	113
4.13	Matrix formulations of basic results	114
4.13.1	State space model in matrix form	114
4.13.2	Matrix expression for densities	115
4.13.3	Filtering in matrix form: Cholesky decomposition	116
4.13.4	Smoothing in matrix form	118
4.13.5	Matrix expressions for signal	119
4.13.6	Simulation smoothing	120
4.14	Exercises	121

5.	Initialisation of filter and smoother	123
5.1	Introduction	123
5.2	The exact initial Kalman filter	126
5.2.1	The basic recursions	126
5.2.2	Transition to the usual Kalman filter	129
5.2.3	A convenient representation	130
5.3	Exact initial state smoothing	130
5.3.1	Smoothed mean of state vector	130
5.3.2	Smoothed variance of state vector	132
5.4	Exact initial disturbance smoothing	134
5.5	Exact initial simulation smoothing	135
5.5.1	Modifications for diffuse initial conditions	135
5.5.2	Exact initial simulation smoothing	136
5.6	Examples of initial conditions for some models	136
5.6.1	Structural time series models	136
5.6.2	Stationary ARMA models	137
5.6.3	Nonstationary ARIMA models	138
5.6.4	Regression model with ARMA errors	140
5.6.5	Spline smoothing	141
5.7	Augmented Kalman filter and smoother	141
5.7.1	Introduction	141
5.7.2	Augmented Kalman filter	141
5.7.3	Filtering based on the augmented Kalman filter	142
5.7.4	Illustration: the local linear trend model	144
5.7.5	Comparisons of computational efficiency	145
5.7.6	Smoothing based on the augmented Kalman filter	146
6.	Further computational aspects	147
6.1	Introduction	147
6.2	Regression estimation	147
6.2.1	Introduction	147
6.2.2	Inclusion of coefficient vector in state vector	148
6.2.3	Regression estimation by augmentation	148
6.2.4	Least squares and recursive residuals	150
6.3	Square root filter and smoother	150
6.3.1	Introduction	150
6.3.2	Square root form of variance updating	151
6.3.3	Givens rotations	152
6.3.4	Square root smoothing	153
6.3.5	Square root filtering and initialisation	154
6.3.6	Illustration: local linear trend model	154
6.4	Univariate treatment of multivariate series	155
6.4.1	Introduction	155
6.4.2	Details of univariate treatment	155

6.4.3	Correlation between observation equations	158
6.4.4	Computational efficiency	158
6.4.5	Illustration: vector splines	159
6.5	Collapsing large observation vectors	161
6.5.1	Introduction	161
6.5.2	Collapse by transformation	162
6.5.3	A generalisation of collapsing by transformation	163
6.5.4	Computational efficiency	164
6.6	Filtering and smoothing under linear restrictions	164
6.7	Computer packages for state space methods	165
6.7.1	Introduction	165
6.7.2	<i>SsfPack</i>	165
6.7.3	The basic <i>SsfPack</i> functions	166
6.7.4	The extended <i>SsfPack</i> functions	166
6.7.5	Illustration: spline smoothing	167
	Maximum likelihood estimation of parameters	170
7.1	Introduction	170
7.2	Likelihood evaluation	170
7.2.1	Loglikelihood when initial conditions are known	170
7.2.2	Diffuse loglikelihood	171
7.2.3	Diffuse loglikelihood via augmented Kalman filter	173
7.2.4	Likelihood when elements of initial state vector are fixed but unknown	174
7.2.5	Likelihood when a univariate treatment of multivariate series is employed	174
7.2.6	Likelihood when the model contains regression effects	175
7.2.7	Likelihood when large observation vector is collapsed	176
7.3	Parameter estimation	177
7.3.1	Introduction	177
7.3.2	Numerical maximisation algorithms	177
7.3.3	The score vector	179
7.3.4	The EM algorithm	182
7.3.5	Estimation when dealing with diffuse initial conditions	184
7.3.6	Large sample distribution of estimates	185
7.3.7	Effect of errors in parameter estimation	186
7.4	Goodness of fit	187
7.5	Diagnostic checking	188
	Illustrations of the use of the linear model	190
8.1	Introduction	190
8.2	Structural time series models	190
8.3	Bivariate structural time series analysis	195
8.4	Box-Jenkins analysis	198
8.5	Spline smoothing	200
8.6	Dynamic factor analysis	202

PART II NON-GAUSSIAN AND NONLINEAR STATE SPACE MODELS

9.	Special cases of nonlinear and non-Gaussian models	209
9.1	Introduction	209
9.2	Models with a linear Gaussian signal	209
9.3	Exponential family models	211
9.3.1	Poisson density	211
9.3.2	Binary density	212
9.3.3	Binomial density	212
9.3.4	Negative binomial density	213
9.3.5	Multinomial density	213
9.3.6	Multivariate extensions	214
9.4	Heavy-tailed distributions	215
9.4.1	t-distribution	215
9.4.2	Mixture of normals	215
9.4.3	General error distribution	215
9.5	Stochastic volatility models	216
9.5.1	Multiple volatility factors	217
9.5.2	Regression and fixed effects	217
9.5.3	Heavy-tailed disturbances	218
9.5.4	Additive noise	218
9.5.5	Leverage effects	219
9.5.6	Stochastic volatility in mean	220
9.5.7	Multivariate SV models	220
9.5.8	Generalised autoregressive conditional heteroscedasticity	221
9.6	Other financial models	222
9.6.1	Durations: exponential distribution	222
9.6.2	Trade frequencies: Poisson distribution	223
9.6.3	Credit risk models	223
9.7	Nonlinear models	224
10.	Approximate filtering and smoothing	226
10.1	Introduction	226
10.2	The extended Kalman filter	226
10.2.1	A multiplicative trend-cycle decomposition	228
10.2.2	Power growth model	229
10.3	The unscented Kalman filter	230
10.3.1	The unscented transformation	230
10.3.2	Derivation of the unscented Kalman filter	232
10.3.3	Further developments of the unscented transform	233
10.3.4	Comparisons between EKF and UKF	236
10.4	Nonlinear smoothing	237
10.4.1	Extended smoothing	237
10.4.2	Unscented smoothing	237

10.5	Approximation via data transformation	238
10.5.1	Partly multiplicative decompositions	239
10.5.2	Stochastic volatility model	239
10.6	Approximation via mode estimation	240
10.6.1	Mode estimation for the linear Gaussian model	240
10.6.2	Mode estimation for model with linear Gaussian signal	241
10.6.3	Mode estimation by linearisation	243
10.6.4	Mode estimation for exponential family models	245
10.6.5	Mode estimation for stochastic volatility model	245
10.7	Further advances in mode estimation	247
10.7.1	Linearisation based on the state vector	247
10.7.2	Linearisation for linear state equations	248
10.7.3	Linearisation for nonlinear models	250
10.7.4	Linearisation for multiplicative models	251
10.7.5	An optimal property for the mode	252
10.8	Treatments for heavy-tailed distributions	254
10.8.1	Mode estimation for models with heavy-tailed densities	254
10.8.2	Mode estimation for state errors with \wedge -distribution	255
10.8.3	A simulation treatment for \wedge -distribution model	255
10.8.4	A simulation treatment for mixture of normals model	258
11.	Importance sampling for smoothing	260
11.1	Introduction	260
11.2	Basic ideas of importance sampling	261
11.3	Choice of an importance density	263
11.4	Implementation details of importance sampling	264
11.4.1	Introduction	264
11.4.2	Practical implementation of importance sampling	264
11.4.3	Antithetic variables	265
11.4.4	Diffuse initialisation	266
11.5	Estimating functions of the state vector	268
11.5.1	Estimating mean functions	268
11.5.2	Estimating variance functions	268
11.5.3	Estimating conditional densities	269
11.5.4	Estimating conditional distribution functions	270
11.5.5	Forecasting and estimating with missing observations	270
11.6	Estimating loglikelihood and parameters	271
11.6.1	Estimation of likelihood	271
11.6.2	Maximisation of loglikelihood	272
11.6.3	Variance matrix of maximum likelihood estimate	273
11.6.4	Effect of errors in parameter estimation	273
11.6.5	Mean square error matrix due to simulation	273
11.7	Importance sampling weights and diagnostics	275

12. Particle filtering	276
12.1 Introduction	276
12.2 Filtering by importance sampling	276
12.3 Sequential importance sampling	278
12.3.1 Introduction	278
12.3.2 Recursions for particle filtering	279
12.3.3 Degeneracy and resampling	280
12.3.4 Algorithm for sequential importance sampling	282
12.4 The bootstrap particle filter	283
12.4.1 Introduction	283
12.4.2 The bootstrap filter	283
12.4.3 Algorithm for bootstrap filter	283
12.4.4 Illustration: local level model for Nile data	284
12.5 The auxiliary particle filter	287
12.5.1 Algorithm for auxiliary filter	287
12.5.2 Illustration: local level model for Nile data	288
12.6 Other implementations of particle filtering	288
12.6.1 Importance density from extended or unscented filter	288
12.6.2 The local regression filter	291
12.6.3 The mode equalisation filter	294
12.7 Rao-Blackwellisation	296
12.7.1 Introduction	296
12.7.2 The Rao-Blackwellisation technique	297
13. Bayesian estimation of parameters	299
13.1 Introduction	299
13.2 Posterior analysis for linear Gaussian model	300
13.2.1 Posterior analysis based on importance sampling	300
13.2.2 Non-informative priors	302
13.3 Posterior analysis for a nonlinear non-Gaussian model	303
13.3.1 Posterior analysis of functions of the state vector	303
13.3.2 Computational aspects of Bayesian analysis	305
13.3.3 Posterior analysis of parameter vector	307
13.4 Markov chain Monte Carlo methods	309
14. Non-Gaussian and nonlinear illustrations	312
14.1 Introduction	312
14.2 Nonlinear decomposition: UK visits abroad	312
14.3 Poisson density: van drivers killed in Great Britain	314
14.4 Heavy-tailed density: outlier in gas consumption	317

14.5 Volatility: pound/dollar daily exchange rates	319
14.5.1 Data transformation analysis	319
14.5.2 Estimation via importance sampling	321
14.5.3 Particle filtering illustration	322
14.6 Binary density: Oxford-Cambridge boat race	324
References	326
Author Index	340
Subject Index	343