
Contents

Preface	xxi
Preface to first edition	xxiii
Notation and abbreviations	xxvii
I Model structure, properties and methods	1
1 Preliminaries: mixtures and Markov chains	3
1.1 Introduction	3
1.2 Independent mixture models	6
1.2.1 Definition and properties	6
1.2.2 Parameter estimation	9
1.2.3 Unbounded likelihood in mixtures	11
1.2.4 Examples of fitted mixture models	12
1.3 Markov chains	14
1.3.1 Definitions and example	14
1.3.2 Stationary distributions	17
1.3.3 Autocorrelation function	18
1.3.4 Estimating transition probabilities	19
1.3.5 Higher-order Markov chains	20
Exercises	23
2 Hidden Markov models: definition and properties	29
2.1 A simple hidden Markov model	29
2.2 The basics	30
2.2.1 Definition and notation	30
2.2.2 Marginal distributions	32
2.2.3 Moments	33
2.3 The likelihood	34
2.3.1 The likelihood of a two-state Bernoulli–HMM	35
2.3.2 The likelihood in general	36
2.3.3 HMMs are not Markov processes	39
2.3.4 The likelihood when data are missing	40

2.3.5	The likelihood when observations are interval-censored	41
	Exercises	41
3	Estimation by direct maximization of the likelihood	47
3.1	Introduction	47
3.2	Scaling the likelihood computation	48
3.3	Maximization of the likelihood subject to constraints	50
3.3.1	Reparametrization to avoid constraints	50
3.3.2	Embedding in a continuous-time Markov chain	52
3.4	Other problems	53
3.4.1	Multiple maxima in the likelihood	53
3.4.2	Starting values for the iterations	53
3.4.3	Unbounded likelihood	53
3.5	Example: earthquakes	54
3.6	Standard errors and confidence intervals	56
3.6.1	Standard errors via the Hessian	56
3.6.2	Bootstrap standard errors and confidence intervals	58
3.7	Example: the parametric bootstrap applied to the three-state model for the earthquakes data	59
	Exercises	60
4	Estimation by the EM algorithm	65
4.1	Forward and backward probabilities	65
4.1.1	Forward probabilities	66
4.1.2	Backward probabilities	67
4.1.3	Properties of forward and backward probabilities	68
4.2	The EM algorithm	69
4.2.1	EM in general	70
4.2.2	EM for HMMs	70
4.2.3	M step for Poisson- and normal-HMMs	72
4.2.4	Starting from a specified state	73
4.2.5	EM for the case in which the Markov chain is stationary	73
4.3	Examples of EM applied to Poisson-HMMs	74
4.3.1	Earthquakes	74
4.3.2	Foetal movement counts	76
4.4	Discussion	77
	Exercises	78
5	Forecasting, decoding and state prediction	81
5.1	Introduction	81

CONTENTS	xiii
5.2 Conditional distributions	82
5.3 Forecast distributions	83
5.4 Decoding	85
5.4.1 State probabilities and local decoding	86
5.4.2 Global decoding	88
5.5 State prediction	92
5.6 HMMs for classification	93
Exercises	94
6 Model selection and checking	97
6.1 Model selection by AIC and BIC	97
6.2 Model checking with pseudo-residuals	101
6.2.1 Introducing pseudo-residuals	101
6.2.2 Ordinary pseudo-residuals	105
6.2.3 Forecast pseudo-residuals	105
6.3 Examples	106
6.3.1 Ordinary pseudo-residuals for the earthquakes	106
6.3.2 Dependent ordinary pseudo-residuals	108
6.4 Discussion	109
Exercises	109
7 Bayesian inference for Poisson–hidden Markov models	111
7.1 Applying the Gibbs sampler to Poisson–HMMs	111
7.1.1 Introduction and outline	111
7.1.2 Generating sample paths of the Markov chain	113
7.1.3 Decomposing the observed counts into regime contributions	114
7.1.4 Updating the parameters	114
7.2 Bayesian estimation of the number of states	114
7.2.1 Use of the integrated likelihood	115
7.2.2 Model selection by parallel sampling	116
7.3 Example: earthquakes	116
7.4 Discussion	119
Exercises	120
8 R packages	123
8.1 The package <code>depmixS4</code>	123
8.1.1 Model formulation and estimation	123
8.1.2 Decoding	124
8.2 The package <code>HiddenMarkov</code>	124
8.2.1 Model formulation and estimation	124
8.2.2 Decoding	126
8.2.3 Residuals	126
8.3 The package <code>msm</code>	126

8.3.1	Model formulation and estimation	126
8.3.2	Decoding	128
8.4	The package R2OpenBUGS	128
8.5	Discussion	129
II	Extensions	131
9	HMMs with general state-dependent distribution	133
9.1	Introduction	133
9.2	General univariate state-dependent distribution	133
9.2.1	HMMs for unbounded counts	133
9.2.2	HMMs for binary data	134
9.2.3	HMMs for bounded counts	134
9.2.4	HMMs for continuous-valued series	135
9.2.5	HMMs for proportions	135
9.2.6	HMMs for circular-valued series	136
9.3	Multinomial and categorical HMMs	136
9.3.1	Multinomial-HMM	136
9.3.2	HMMs for categorical data	137
9.3.3	HMMs for compositional data	138
9.4	General multivariate state-dependent distribution	138
9.4.1	Longitudinal conditional independence	138
9.4.2	Contemporaneous conditional independence	140
9.4.3	Further remarks on multivariate HMMs	141
	Exercises	142
10	Covariates and other extra dependencies	145
10.1	Introduction	145
10.2	HMMs with covariates	145
10.2.1	Covariates in the state-dependent distributions	146
10.2.2	Covariates in the transition probabilities	147
10.3	HMMs based on a second-order Markov chain	148
10.4	HMMs with other additional dependencies	150
	Exercises	152
11	Continuous-valued state processes	155
11.1	Introduction	155
11.2	Models with continuous-valued state process	156
11.2.1	Numerical integration of the likelihood	157
11.2.2	Evaluation of the approximate likelihood via forward recursion	158
11.2.3	Parameter estimation and related issues	160
11.3	Fitting an SSM to the earthquake data	160
11.4	Discussion	162

12 Hidden semi-Markov models and their representation as HMMs	165
12.1 Introduction	165
12.2 Semi-Markov processes, hidden semi-Markov models and approximating HMMs	165
12.3 Examples of HSMMs represented as HMMs	167
12.3.1 A simple two-state Poisson–HSMM	167
12.3.2 Example of HSMM with three states	169
12.3.3 A two-state HSMM with general dwell-time distribution in one state	171
12.4 General HSMM	173
12.5 R code	176
12.6 Some examples of dwell-time distributions	178
12.6.1 Geometric distribution	178
12.6.2 Shifted Poisson distribution	178
12.6.3 Shifted negative binomial distribution	179
12.6.4 Shifted binomial distribution	180
12.6.5 A distribution with unstructured start and geometric tail	180
12.7 Fitting HSMMs via the HMM representation	181
12.8 Example: earthquakes	182
12.9 Discussion	184
Exercises	184
13 HMMs for longitudinal data	187
13.1 Introduction	187
13.2 Models that assume some parameters to be constant across component series	189
13.3 Models with random effects	190
13.3.1 HMMs with continuous-valued random effects	191
13.3.2 HMMs with discrete-valued random effects	193
13.4 Discussion	195
Exercises	196
III Applications	197
14 Introduction to applications	199
15 Epileptic seizures	201
15.1 Introduction	201
15.2 Models fitted	201
15.3 Model checking by pseudo-residuals	204
Exercises	206

16 Daily rainfall occurrence	207
16.1 Introduction	207
16.2 Models fitted	207
17 Eruptions of the Old Faithful geyser	213
17.1 Introduction	213
17.2 The data	213
17.3 The binary time series of short and long eruptions	214
17.3.1 Markov chain models	214
17.3.2 Hidden Markov models	216
17.3.3 Comparison of models	219
17.3.4 Forecast distributions	219
17.4 Univariate normal-HMMs for durations and waiting times	220
17.5 Bivariate normal-HMM for durations and waiting times	223
Exercises	224
18 HMMs for animal movement	227
18.1 Introduction	227
18.2 Directional data	228
18.2.1 Directional means	228
18.2.2 The von Mises distribution	228
18.3 HMMs for movement data	229
18.3.1 Movement data	229
18.3.2 HMMs as multi-state random walks	230
18.4 A basic HMM for <i>Drosophila</i> movement	232
18.5 HMMs and HSMMs for bison movement	235
18.6 Mixed HMMs for woodpecker movement	238
Exercises	242
19 Wind direction at Koeberg	245
19.1 Introduction	245
19.2 Wind direction classified into 16 categories	245
19.2.1 Three HMMs for hourly averages of wind direction	245
19.2.2 Model comparisons and other possible models	248
19.3 Wind direction as a circular variable	251
19.3.1 Daily at hour 24: von Mises-HMMs	251
19.3.2 Modelling hourly change of direction	253
19.3.3 Transition probabilities varying with lagged speed	253
19.3.4 Concentration parameter varying with lagged speed	254
Exercises	257

CONTENTS	xvii
20 Models for financial series	259
20.1 Financial series I: A multivariate normal–HMM for returns on four shares	259
20.2 Financial series II: Discrete state-space stochastic volatility models	262
20.2.1 Stochastic volatility models without leverage	263
20.2.2 Application: FTSE 100 returns	265
20.2.3 Stochastic volatility models with leverage	265
20.2.4 Application: TOPIX returns	268
20.2.5 Non-standard stochastic volatility models	270
20.2.6 A model with a mixture AR(1) volatility process	271
20.2.7 Application: S&P 500 returns	272
Exercises	273
21 Births at Edendale Hospital	275
21.1 Introduction	275
21.2 Models for the proportion Caesarean	275
21.3 Models for the total number of deliveries	282
21.4 Conclusion	285
22 Homicides and suicides in Cape Town, 1986–1991	287
22.1 Introduction	287
22.2 Firearm homicides as a proportion of all homicides, suicides and legal intervention homicides	287
22.3 The number of firearm homicides	289
22.4 Firearm homicides as a proportion of all homicides, and firearm suicides as a proportion of all suicides	291
22.5 Proportion in each of the five categories	295
23 A model for animal behaviour which incorporates feedback	297
23.1 Introduction	297
23.2 The model	298
23.3 Likelihood evaluation	300
23.3.1 The likelihood as a multiple sum	301
23.3.2 Recursive evaluation	301
23.4 Parameter estimation by maximum likelihood	302
23.5 Model checking	302
23.6 Inferring the underlying state	303
23.7 Models for a heterogeneous group of subjects	304
23.7.1 Models assuming some parameters to be constant across subjects	304
23.7.2 Mixed models	305

23.7.3	Inclusion of covariates	306
23.8	Other modifications or extensions	306
23.8.1	Increasing the number of states	306
23.8.2	Changing the nature of the state-dependent distribution	306
23.9	Application to caterpillar feeding behaviour	307
23.9.1	Data description and preliminary analysis	307
23.9.2	Parameter estimates and model checking	307
23.9.3	Runlength distributions	311
23.9.4	Joint models for seven subjects	313
23.10	Discussion	314
24	Estimating the survival rates of Soay sheep from mark-recapture-recovery data	317
24.1	Introduction	317
24.2	MRR data without use of covariates	318
24.3	MRR data involving individual-specific time-varying continuous-valued covariates	321
24.4	Application to Soay sheep data	324
24.5	Conclusion	328
A	Examples of R code	331
A.1	The functions	331
A.1.1	Transforming natural parameters to working	332
A.1.2	Transforming working parameters to natural	332
A.1.3	Computing minus the log-likelihood from the working parameters	332
A.1.4	Computing the MLEs, given starting values for the natural parameters	333
A.1.5	Generating a sample	333
A.1.6	Global decoding by the Viterbi algorithm	334
A.1.7	Computing log(forward probabilities)	334
A.1.8	Computing log(backward probabilities)	334
A.1.9	Conditional probabilities	335
A.1.10	Pseudo-residuals	336
A.1.11	State probabilities	336
A.1.12	State prediction	336
A.1.13	Local decoding	337
A.1.14	Forecast probabilities	337
A.2	Examples of code using the above functions	338
A.2.1	Fitting Poisson-HMMs to the earthquakes series	338
A.2.2	Forecast probabilities	339

CONTENTS	xix
B Some proofs	341
B.1 A factorization needed for the forward probabilities	341
B.2 Two results needed for the backward probabilities	342
B.3 Conditional independence of \mathbf{X}_1^t and \mathbf{X}_{t+1}^T	343
References	345
Author index	359
Subject index	365