
Contents

Preface	xvii
Notation and abbreviations	xxi
PART ONE 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	10
1.2.4 Examples of fitted mixture models	11
1.3 Markov chains	15
1.3.1 Definitions and example	16
1.3.2 Stationary distributions	18
1.3.3 Reversibility	19
1.3.4 Autocorrelation function	19
1.3.5 Estimating transition probabilities	20
1.3.6 Higher-order Markov chains	22
Exercises	24
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	34
2.3 The likelihood	35
2.3.1 The likelihood of a two-state Bernoulli–HMM	35
2.3.2 The likelihood in general	37
2.3.3 The likelihood when data are missing at random	39

2.3.4	The likelihood when observations are interval-censored	40
	Exercises	41
3	Estimation by direct maximization of the likelihood	45
3.1	Introduction	45
3.2	Scaling the likelihood computation	46
3.3	Maximization subject to constraints	47
3.3.1	Reparametrization to avoid constraints	47
3.3.2	Embedding in a continuous-time Markov chain	49
3.4	Other problems	49
3.4.1	Multiple maxima in the likelihood	49
3.4.2	Starting values for the iterations	50
3.4.3	Unbounded likelihood	50
3.5	Example: earthquakes	50
3.6	Standard errors and confidence intervals	53
3.6.1	Standard errors via the Hessian	53
3.6.2	Bootstrap standard errors and confidence intervals	55
3.7	Example: parametric bootstrap	55
	Exercises	57
4	Estimation by the EM algorithm	59
4.1	Forward and backward probabilities	59
4.1.1	Forward probabilities	60
4.1.2	Backward probabilities	61
4.1.3	Properties of forward and backward probabilities	62
4.2	The EM algorithm	63
4.2.1	EM in general	63
4.2.2	EM for HMMs	64
4.2.3	M step for Poisson- and normal-HMMs	66
4.2.4	Starting from a specified state	67
4.2.5	EM for the case in which the Markov chain is stationary	67
4.3	Examples of EM applied to Poisson-HMMs	68
4.3.1	Earthquakes	68
4.3.2	Foetal movement counts	70
4.4	Discussion	72
	Exercises	73
5	Forecasting, decoding and state prediction	75
5.1	Conditional distributions	76

CONTENTS	xi
5.2 Forecast distributions	77
5.3 Decoding	80
5.3.1 State probabilities and local decoding	80
5.3.2 Global decoding	82
5.4 State prediction	86
Exercises	87
6 Model selection and checking	89
6.1 Model selection by AIC and BIC	89
6.2 Model checking with pseudo-residuals	92
6.2.1 Introducing pseudo-residuals	93
6.2.2 Ordinary pseudo-residuals	96
6.2.3 Forecast pseudo-residuals	97
6.3 Examples	98
6.3.1 Ordinary pseudo-residuals for the earthquakes	98
6.3.2 Dependent ordinary pseudo-residuals	98
6.4 Discussion	100
Exercises	101
7 Bayesian inference for Poisson–HMMs	103
7.1 Applying the Gibbs sampler to Poisson–HMMs	103
7.1.1 Generating sample paths of the Markov chain	105
7.1.2 Decomposing observed counts	106
7.1.3 Updating the parameters	106
7.2 Bayesian estimation of the number of states	106
7.2.1 Use of the integrated likelihood	107
7.2.2 Model selection by parallel sampling	108
7.3 Example: earthquakes	108
7.4 Discussion	110
Exercises	112
8 Extensions of the basic hidden Markov model	115
8.1 Introduction	115
8.2 HMMs with general univariate state-dependent distribution	116
8.3 HMMs based on a second-order Markov chain	118
8.4 HMMs for multivariate series	119
8.4.1 Series of multinomial-like observations	119
8.4.2 A model for categorical series	121
8.4.3 Other multivariate models	122
8.5 Series that depend on covariates	125
8.5.1 Covariates in the state-dependent distributions	125
8.5.2 Covariates in the transition probabilities	126

8.6	Models with additional dependencies	128
	Exercises	129
PART TWO Applications		133
9	Epileptic seizures	135
9.1	Introduction	135
9.2	Models fitted	135
9.3	Model checking by pseudo-residuals	138
	Exercises	140
10	Eruptions of the Old Faithful geyser	141
10.1	Introduction	141
10.2	Binary time series of short and long eruptions	141
10.2.1	Markov chain models	142
10.2.2	Hidden Markov models	144
10.2.3	Comparison of models	147
10.2.4	Forecast distributions	148
10.3	Normal–HMMs for durations and waiting times	149
10.4	Bivariate model for durations and waiting times	152
	Exercises	153
11	<i>Drosophila</i> speed and change of direction	155
11.1	Introduction	155
11.2	Von Mises distributions	156
11.3	Von Mises–HMMs for the two subjects	157
11.4	Circular autocorrelation functions	158
11.5	Bivariate model	161
	Exercises	165
12	Wind direction at Koeberg	167
12.1	Introduction	167
12.2	Wind direction classified into 16 categories	167
12.2.1	Three HMMs for hourly averages of wind direction	167
12.2.2	Model comparisons and other possible models	170
12.2.3	Conclusion	173
12.3	Wind direction as a circular variable	174
12.3.1	Daily at hour 24: von Mises–HMMs	174
12.3.2	Modelling hourly change of direction	176
12.3.3	Transition probabilities varying with lagged speed	176

CONTENTS	xiii
12.3.4 Concentration parameter varying with lagged speed	177
Exercises	180
13 Models for financial series	181
13.1 Thinly traded shares	181
13.1.1 Univariate models	181
13.1.2 Multivariate models	183
13.1.3 Discussion	185
13.2 Multivariate HMM for returns on four shares	186
13.3 Stochastic volatility models	190
13.3.1 Stochastic volatility models without leverage	190
13.3.2 Application: FTSE 100 returns	192
13.3.3 Stochastic volatility models with leverage	193
13.3.4 Application: TOPIX returns	195
13.3.5 Discussion	197
14 Births at Edendale Hospital	199
14.1 Introduction	199
14.2 Models for the proportion Caesarean	199
14.3 Models for the total number of deliveries	205
14.4 Conclusion	208
15 Homicides and suicides in Cape Town	209
15.1 Introduction	209
15.2 Firearm homicides as a proportion of all homicides, suicides and legal intervention homicides	209
15.3 The number of firearm homicides	211
15.4 Firearm homicide and suicide proportions	213
15.5 Proportion in each of the five categories	217
16 Animal behaviour model with feedback	219
16.1 Introduction	219
16.2 The model	220
16.3 Likelihood evaluation	222
16.3.1 The likelihood as a multiple sum	223
16.3.2 Recursive evaluation	223
16.4 Parameter estimation by maximum likelihood	224
16.5 Model checking	224
16.6 Inferring the underlying state	225
16.7 Models for a heterogeneous group of subjects	226
16.7.1 Models assuming some parameters to be constant across subjects	226

16.7.2	Mixed models	227
16.7.3	Inclusion of covariates	227
16.8	Other modifications or extensions	228
16.8.1	Increasing the number of states	228
16.8.2	Changing the nature of the state-dependent distribution	228
16.9	Application to caterpillar feeding behaviour	229
16.9.1	Data description and preliminary analysis	229
16.9.2	Parameter estimates and model checking	229
16.9.3	Runlength distributions	233
16.9.4	Joint models for seven subjects	235
16.10	Discussion	236
A	Examples of R code	239
A.1	Stationary Poisson–HMM, numerical maximization	239
A.1.1	Transform natural parameters to working	240
A.1.2	Transform working parameters to natural	240
A.1.3	Log-likelihood of a stationary Poisson–HMM	240
A.1.4	ML estimation of a stationary Poisson–HMM	241
A.2	More on Poisson–HMMs, including EM	242
A.2.1	Generate a realization of a Poisson–HMM	242
A.2.2	Forward and backward probabilities	242
A.2.3	EM estimation of a Poisson–HMM	243
A.2.4	Viterbi algorithm	244
A.2.5	Conditional state probabilities	244
A.2.6	Local decoding	245
A.2.7	State prediction	245
A.2.8	Forecast distributions	246
A.2.9	Conditional distribution of one observation given the rest	246
A.2.10	Ordinary pseudo-residuals	247
A.3	Bivariate normal state-dependent distributions	248
A.3.1	Transform natural parameters to working	248
A.3.2	Transform working parameters to natural	249
A.3.3	Discrete log-likelihood	249
A.3.4	MLEs of the parameters	250
A.4	Categorical HMM, constrained optimization	250
A.4.1	Log-likelihood	251
A.4.2	MLEs of the parameters	252
B	Some proofs	253
B.1	Factorization needed for forward probabilities	253
B.2	Two results for backward probabilities	255

CONTENTS	xv
B.3 Conditional independence of \mathbf{X}_1^t and \mathbf{X}_{t+1}^T	256
References	257
Author index	267
Subject index	271