

Contents

Preface	xi
Acknowledgments	xii
1 Introduction	1
1.1 Laser Example	2
1.2 State Space Models	4
1.2.1 Tasks	6
1.3 Discrete HMMs	7
1.3.1 Example: Quantized Lorenz Time Series	11
1.3.2 Example: Hidden States as Parts of Speech	13
1.3.3 Remarks	15
2 Basic Algorithms	19
2.1 The Forward Algorithm	20
2.1.1 Vector Notation	23
2.1.2 General Filtering	24
2.2 The Backward Algorithm	25
2.3 The Viterbi Algorithm	27
2.3.1 General Decoding	29
2.3.2 MAP Sequence of States or Sequence of MAP States?	29
2.4 The Baum–Welch Algorithm	31
2.4.1 Weights and Reestimation	32
2.4.2 Remarks	37
2.5 The EM Algorithm	38
2.5.1 Monotonicity	42
2.5.2 Convergence	43
3 Variants and Generalizations	47
3.1 Gaussian Observations	48
3.1.1 Independent Scalar Observations	48
3.1.2 Singularities of the Likelihood Function and Regularization	50
3.1.3 The EM Algorithm for MAP Estimation	51
3.1.4 Vector Autoregressive Observations	52
3.2 Related Models	55

4	Continuous States and Observations and Kalman Filtering	59
4.1	Algorithms with Integrals	60
4.1.1	Forward Algorithm	60
4.1.2	Backward Algorithm	61
4.2	Linear Gaussian Systems	62
4.2.1	Kalman Filter: The Forward Algorithm	64
4.2.2	The Backward Algorithm	65
4.2.3	Smoothing	65
4.3	Algorithm Derivations and Details	66
4.3.1	Forward Kalman Filter	66
4.3.2	Backward Recursion	69
4.3.3	Smoothing	71
4.3.4	Inverse Covariance Form	71
4.3.5	Extended Kalman Filter	71
5	Performance Bounds and a Toy Problem	73
5.1	Fidelity Criteria and Entropy	78
5.1.1	Definitions	78
5.2	Stretching and Entropy	82
5.2.1	Maps of the Unit Circle	82
5.3	Lyapunov Exponents and Pesin's Formula	84
5.3.1	A Theoretical Bound on Model Likelihood	85
5.4	Benettin's Procedure for Calculating Lyapunov Exponents Numerically	87
5.5	A Practical Performance Bound	89
5.6	Approaching the Bound	94
6	Obstructive Sleep Apnea	97
6.1	The Challenge and the Data	97
6.1.1	The Data	99
6.2	First Classification Algorithms and Two Useful Features	101
6.2.1	Using Information from Experts to Train	101
6.2.2	Nonlinear Dynamics	102
6.2.3	The Excellent Eye of Dr. McNames	103
6.3	Decoding Sequences of Classifications	104
6.4	Assembling the Pieces	106
6.4.1	Extracting a Low Pass Filtered Heart Rate	107
6.4.2	Extracting Respiration Information	108
6.4.3	Classifying Records	110
6.4.4	Model Topology and Training Data	112
6.4.5	Tunable Parameters	112
6.4.6	Results	114
6.5	Classification Versus Estimation	116
A	Formulas for Matrices and Gaussians	117
B	Notes on Software	121

Contents ix

Bibliography	125
Books and Collections	125
Review Articles	126
Dissertations and Theses	127
Research Articles	127
Web Sites	129
 Index	 131