
Contents

Part I Probabilistic Modeling

1 A Leisurely Look at Statistical Inference

<i>Dirk Husmeier</i>	3
1.1 Preliminaries	3
1.2 The Classical or Frequentist Approach	5
1.3 The Bayesian Approach	10
1.4 Comparison	12
References	15

2 Introduction to Learning Bayesian Networks from Data

<i>Dirk Husmeier</i>	17
2.1 Introduction to Bayesian Networks	17
2.1.1 The Structure of a Bayesian Network	17
2.1.2 The Parameters of a Bayesian Network	25
2.2 Learning Bayesian Networks from Complete Data	25
2.2.1 The Basic Learning Paradigm	25
2.2.2 Markov Chain Monte Carlo (MCMC)	28
2.2.3 Equivalence Classes	35
2.2.4 Causality	38
2.3 Learning Bayesian Networks from Incomplete Data	41
2.3.1 Introduction	41
2.3.2 Evidence Approximation and Bayesian Information Criterion	41
2.3.3 The EM Algorithm	43
2.3.4 Hidden Markov Models	44
2.3.5 Application of the EM Algorithm to HMMs	49
2.3.6 Applying the EM Algorithm to More Complex Bayesian Networks with Hidden States	52
2.3.7 Reversible Jump MCMC	54
2.4 Summary	55

References	55
------------------	----

3 A Casual View of Multi-Layer Perceptrons as Probability Models

<i>Richard Dybowski</i>	59
3.1 A Brief History	59
3.1.1 The McCulloch-Pitts Neuron	59
3.1.2 The Single-Layer Perceptron	60
3.1.3 Enter the Multi-Layer Perceptron	62
3.1.4 A Statistical Perspective	63
3.2 Regression	63
3.2.1 Maximum Likelihood Estimation	65
3.3 From Regression to Probabilistic Classification	65
3.3.1 Multi-Layer Perceptrons	67
3.4 Training a Multi-Layer Perceptron	69
3.4.1 The Error Back-Propagation Algorithm	70
3.4.2 Alternative Training Strategies	73
3.5 Some Practical Considerations	73
3.5.1 Over-Fitting	74
3.5.2 Local Minima	75
3.5.3 Number of Hidden Nodes	77
3.5.4 Preprocessing Techniques	77
3.5.5 Training Sets	78
3.6 Further Reading	78
References	79

Part II Bioinformatics

4 Introduction to Statistical Phylogenetics

<i>Dirk Husmeier</i>	83
4.1 Motivation and Background on Phylogenetic Trees	84
4.2 Distance and Clustering Methods	90
4.2.1 Evolutionary Distances	90
4.2.2 A Naive Clustering Algorithm: UPGMA	93
4.2.3 An Improved Clustering Algorithm: Neighbour Joining ..	96
4.2.4 Shortcomings of Distance and Clustering Methods	98
4.3 Parsimony	100
4.3.1 Introduction	100
4.3.2 Objection to Parsimony	104
4.4 Likelihood Methods	104
4.4.1 A Mathematical Model of Nucleotide Substitution	104
4.4.2 Details of the Mathematical Model of Nucleotide Substitution	106
4.4.3 Likelihood of a Phylogenetic Tree	111
4.4.4 A Comparison with Parsimony	118

4.4.5	Maximum Likelihood	120
4.4.6	Bootstrapping	127
4.4.7	Bayesian Inference	130
4.4.8	Gaps	135
4.4.9	Rate Heterogeneity	136
4.4.10	Protein and RNA Sequences	138
4.4.11	A Non-homogeneous and Non-stationary Markov Model of Nucleotide Substitution	139
4.5	Summary	141
	References	142

5 Detecting Recombination in DNA Sequence Alignments

	<i>Dirk Husmeier, Frank Wright</i>	147
5.1	Introduction	147
5.2	Recombination in Bacteria and Viruses	148
5.3	Phylogenetic Networks	148
5.4	Maximum Chi-squared	152
5.5	PLATO	156
5.6	TOPAL	159
5.7	Probabilistic Divergence Method (PDM)	162
5.8	Empirical Comparison I	167
5.9	RECPARS	170
5.10	Combining Phylogenetic Trees with HMMs	171
5.10.1	Introduction	171
5.10.2	Maximum Likelihood	175
5.10.3	Bayesian Approach	176
5.10.4	Shortcomings of the HMM Approach	180
5.11	Empirical Comparison II	181
5.11.1	Simulated Recombination	181
5.11.2	Gene Conversion in Maize	184
5.11.3	Recombination in <i>Neisseria</i>	184
5.12	Conclusion	187
5.13	Software	188
	References	188

6 RNA-Based Phylogenetic Methods

	<i>Magnus Rattray, Paul G. Higgs</i>	191
6.1	Introduction	191
6.2	RNA Structure	193
6.3	Substitution Processes in RNA Helices	196
6.4	An Application: Mammalian Phylogeny	201
6.5	Conclusion	207
	References	208

7 Statistical Methods in Microarray Gene Expression Data Analysis

<i>Claus-Dieter Mayer, Chris A. Glasbey</i>	211
7.1 Introduction	211
7.1.1 Gene Expression in a Nutshell	211
7.1.2 Microarray Technologies	212
7.2 Image Analysis	214
7.2.1 Image Enhancement	215
7.2.2 Gridding	216
7.2.3 Estimators of Intensities	216
7.3 Transformation	218
7.4 Normalization	222
7.4.1 Explorative Analysis and Flagging of Data Points	222
7.4.2 Linear Models and Experimental Design	225
7.4.3 Non-linear Methods	227
7.4.4 Normalization of One-channel Data	228
7.5 Differential Expression	228
7.5.1 One-slide Approaches	228
7.5.2 Using Replicated Experiments	229
7.5.3 Multiple Testing	232
7.6 Further Reading	234
References	235

8 Inferring Genetic Regulatory Networks from Microarray Experiments with Bayesian Networks

<i>Dirk Husmeier</i>	239
8.1 Introduction	240
8.2 A Brief Revision of Bayesian Networks	241
8.3 Learning Local Structures and Subnetworks	244
8.4 Application to the Yeast Cell Cycle	247
8.4.1 Biological Findings	248
8.5 Shortcomings of Static Bayesian Networks	251
8.6 Dynamic Bayesian Networks	252
8.7 Accuracy of Inference	252
8.8 Evaluation on Synthetic Data	253
8.9 Evaluation on Realistic Data	257
8.10 Discussion	263
References	265

9 Modeling Genetic Regulatory Networks using Gene Expression Profiling and State-Space Models

<i>Claudia Rangel, John Angus, Zoubin Ghahramani, David L. Wild</i>	269
9.1 Introduction	269
9.2 State-Space Models (Linear Dynamical Systems)	272
9.2.1 State-Space Model with Inputs	272

9.2.2	EM Applied to SSM with Inputs	274
9.2.3	Kalman Smoothing	275
9.3	The SSM Model for Gene Expression	277
9.3.1	Structural Properties of the Model	277
9.3.2	Identifiability and Stability Issues	278
9.4	Model Selection by Bootstrapping	281
9.4.1	Objectives	281
9.4.2	The Bootstrap Procedure	281
9.5	Experiments with Simulated Data	283
9.5.1	Model Definition	283
9.5.2	Reconstructing the Original Network	283
9.5.3	Results	283
9.6	Results from Experimental Data	288
9.7	Conclusions	289
	References	291

Part III Medical Informatics

10 An Anthology of Probabilistic Models for Medical Informatics

	<i>Richard Dybowski, Stephen Roberts</i>	297
10.1	Probabilities in Medicine	297
10.2	Desiderata for Probability Models	297
10.3	Bayesian Statistics	298
10.3.1	Parameter Averaging and Model Averaging	299
10.3.2	Computations	300
10.4	Logistic Regression	301
10.5	Bayesian Logistic Regression	302
10.5.1	Gibbs Sampling and GLIB	304
10.5.2	Hierarchical Models	306
10.6	Neural Networks	307
10.6.1	Multi-Layer Perceptrons	307
10.6.2	Radial-Basis-Function Neural Networks	308
10.6.3	“Probabilistic Neural Networks”	309
10.6.4	Missing Data	310
10.7	Bayesian Neural Techniques	311
10.7.1	Moderated Output	311
10.7.2	Hyperparameters	312
10.7.3	Committees	313
10.7.4	Full Bayesian Models	314
10.8	The Naïve Bayes Model	316
10.9	Bayesian Networks	317
10.9.1	Probabilistic Inference over BNs	318
10.9.2	Sigmoidal Belief Networks	321

10.9.3	Construction of BNs: Probabilities	321
10.9.4	Construction of BNs: Structures	322
10.9.5	Missing Data	322
10.10	Class-Probability Trees	323
10.10.1	Missing Data	324
10.10.2	Bayesian Tree Induction	325
10.11	Probabilistic Models for Detection	326
10.11.1	Data Conditioning	327
10.11.2	Detection, Segmentation and Decisions	330
10.11.3	Cluster Analysis	331
10.11.4	Hidden Markov Models	335
10.11.5	Novelty Detection	338
	References	338

11 Bayesian Analysis of Population

Pharmacokinetic/Pharmacodynamic Models

	<i>David J. Lunn</i>	351
11.1	Introduction	351
11.2	Deterministic Models	352
11.2.1	Pharmacokinetics	352
11.2.2	Pharmacodynamics	359
11.3	Stochastic Model	360
11.3.1	Structure	360
11.3.2	Priors	363
11.3.3	Parameterization Issues	364
11.3.4	Analysis	365
11.3.5	Prediction	366
11.4	Implementation	367
11.4.1	PKBugs	367
11.4.2	WinBUGS Differential Interface	368
	References	369

12 Assessing the Effectiveness of Bayesian Feature Selection

	<i>Ian T. Nabney, David J. Evans, Yann Brulé, Caroline Gordon</i>	371
12.1	Introduction	371
12.2	Bayesian Feature Selection	372
12.2.1	Bayesian Techniques for Neural Networks	372
12.2.2	Automatic Relevance Determination	374
12.3	ARD in Arrhythmia Classification	375
12.3.1	Clinical Context	375
12.3.2	Benchmarking Classification Models	376
12.3.3	Variable Selection	379
12.3.4	Conclusions	380
12.4	ARD in Lupus Diagnosis	381
12.4.1	Clinical Context	381

12.4.2	Linear Methods for Variable Selection	383
12.4.3	Prognosis with Non-linear Models	383
12.4.4	Bayesian Variable Selection	385
12.4.5	Conclusions	386
12.5	Conclusions	387
	References	388

13 Bayes Consistent Classification of EEG Data by Approximate Marginalization

	<i>Peter Sykacek, Iead Rezek, and Stephen Roberts</i>	391
13.1	Introduction	391
13.2	Bayesian Lattice Filter	393
13.3	Spatial Fusion	396
13.4	Spatio-temporal Fusion	400
	13.4.1 A Simple DAG Structure	401
	13.4.2 A Likelihood Function for Sequence Models	402
	13.4.3 An Augmented DAG for MCMC Sampling	403
	13.4.4 Specifying Priors	404
	13.4.5 MCMC Updates of Coefficients and Latent Variables	405
	13.4.6 Gibbs Updates for Hidden States and Class Labels	407
	13.4.7 Approximate Updates of the Latent Feature Space	408
	13.4.8 Algorithms	409
13.5	Experiments	411
	13.5.1 Data	412
	13.5.2 Classification Results	413
13.6	Conclusion	415
	References	416

14 Ensemble Hidden Markov Models with Extended Observation Densities for Biosignal Analysis

	<i>Iead Rezek, Stephen Roberts</i>	419
14.1	Introduction	419
14.2	Principles of Variational Learning	421
14.3	Variational Learning of Hidden Markov Models	423
	14.3.1 Learning the HMM Hidden State Sequence	425
	14.3.2 Learning HMM Parameters	426
	14.3.3 HMM Observation Models	427
	14.3.4 Estimation	431
14.4	Experiments	435
	14.4.1 Sleep EEG with Arousal	435
	14.4.2 Whole-Night Sleep EEG	435
	14.4.3 Periodic Respiration	436
	14.4.4 Heartbeat Intervals	437
	14.4.5 Segmentation of Cognitive Tasks	439
14.5	Conclusion	440

A	Model Free Update Equations	442
B	Derivation of the Baum-Welch Recursions.....	443
C	Complete KL Divergences	445
	C.1 Negative Entropy	446
	C.2 KL Divergences	446
	C.3 Gaussian Observation HMM	447
	C.4 Poisson Observation HMM.....	448
	C.5 Linear Observation Model HMM	448
	References	449

15 A Probabilistic Network for Fusion of Data and Knowledge in Clinical Microbiology

Steen Andreassen, Leonard Leibovici, Mical Paul, Anders D. Nielsen, Alina Zalounina, Leif E. Kristensen, Karsten Falborg, Brian

	<i>Kristensen, Uwe Frank, Henrik C. Schönheyder.....</i>	451
15.1	Introduction	451
15.2	Institution of Antibiotic Therapy	453
15.3	Calculation of Probabilities for Severity of Sepsis, Site of Infection, and Pathogens.....	454
	15.3.1 Patient Example (Part 1)	454
	15.3.2 Fusion of Data and Knowledge for Calculation of Probabilities for Sepsis and Pathogens.....	456
15.4	Calculation of Coverage and Treatment Advice	461
	15.4.1 Patient Example (Part 2)	461
	15.4.2 Fusion of Data and Knowledge for Calculation of Coverage and Treatment Advice	466
15.5	Calibration Databases	467
15.6	Clinical Testing of Decision-support Systems	468
15.7	Test Results	468
15.8	Discussion	469
	References	470

16 Software for Probability Models in Medical Informatics

Richard Dybowski

16.1	Introduction	473
16.2	Open-source Software	474
16.3	Logistic Regression Models.....	474
	16.3.1 S-Plus and R.....	475
	16.3.2 BUGS.....	476
16.4	Neural Networks.....	477
	16.4.1 Netlab	477
	16.4.2 The Stuttgart Neural Network Simulator	478
16.5	Bayesian Networks	478
	16.5.1 Hugin and Netica	481
	16.5.2 The Bayes Net Toolbox	481

16.5.3	The OpenBayes Initiative	483
16.5.4	The Probabilistic Networks Library	483
16.5.5	The gR Project	484
16.5.6	The VIBES Project	484
16.6	Class-probability trees	484
16.7	Hidden Markov Models	485
16.7.1	Hidden Markov Model Toolbox for Matlab	486
References	487
A Appendix: Conventions and Notation		
	491
Index	495

Probabilistic Modeling in Bioinformatics and Medical
Informatics

Husmeier, D.; Dybowski, R.; Roberts, S. (Eds.)

2005, XX, 508 p., Hardcover

ISBN: 978-1-85233-778-0