

Václav Šmídl
Anthony Quinn

The Variational Bayes Method in Signal Processing

With 65 Figures

 Springer

Contents

1	Introduction	1
1.1	How to be a Bayesian	1
1.2	The Variational Bayes (VB) Method	2
1.3	A First Example of the VB Method: Scalar Additive Decomposition	3
1.3.1	A First Choice of Prior	3
1.3.2	The Prior Choice Revisited	4
1.4	The VB Method in its Context	6
1.5	VB as a Distributional Approximation	8
1.6	Layout of the Work	10
1.7	Acknowledgement	11
2	Bayesian Theory	13
2.1	Bayesian Benefits	13
2.1.1	Off-line vs. On-line Parametric Inference	14
2.2	Bayesian Parametric Inference: the Off-Line Case	15
2.2.1	The Subjective Philosophy	16
2.2.2	Posterior Inferences and Decisions	16
2.2.3	Prior Elicitation	18
2.2.3.1	Conjugate priors	19
2.3	Bayesian Parametric Inference: the On-line Case	19
2.3.1	Time-invariant Parameterization	20
2.3.2	Time-variant Parameterization	20
2.3.3	Prediction	22
2.4	Summary	22
3	Off-line Distributional Approximations and the Variational Bayes Method	25
3.1	Distributional Approximation	25
3.2	How to Choose a Distributional Approximation	26
3.2.1	Distributional Approximation as an Optimization Problem	26
3.2.2	The Bayesian Approach to Distributional Approximation	27

3.3	The Variational Bayes (VB) Method of Distributional Approximation	28
3.3.1	The VB Theorem	28
3.3.2	The VB Method of Approximation as an Operator	32
3.3.3	The VB Method	33
3.3.4	The VB Method for Scalar Additive Decomposition	37
3.4	VB-related Distributional Approximations	39
3.4.1	Optimization with Minimum-Risk KL Divergence	39
3.4.2	Fixed-form (FF) Approximation	40
3.4.3	Restricted VB (RVB) Approximation	40
3.4.3.1	Adaptation of the VB method for the RVB Approximation	41
3.4.3.2	The Quasi-Bayes (QB) Approximation	42
3.4.4	The Expectation-Maximization (EM) Algorithm	44
3.5	Other Deterministic Distributional Approximations	45
3.5.1	The Certainty Equivalence Approximation	45
3.5.2	The Laplace Approximation	45
3.5.3	The Maximum Entropy (MaxEnt) Approximation	45
3.6	Stochastic Distributional Approximations	46
3.6.1	Distributional Estimation	47
3.7	Example: Scalar Multiplicative Decomposition	48
3.7.1	Classical Modelling	48
3.7.2	The Bayesian Formulation	48
3.7.3	Full Bayesian Solution	49
3.7.4	The Variational Bayes (VB) Approximation	51
3.7.5	Comparison with Other Techniques	54
3.8	Conclusion	56
4	Principal Component Analysis and Matrix Decompositions	57
4.1	Probabilistic Principal Component Analysis (PPCA)	58
4.1.1	Maximum Likelihood (ML) Estimation for the PPCA Model	59
4.1.2	Marginal Likelihood Inference of A	61
4.1.3	Exact Bayesian Analysis	61
4.1.4	The Laplace Approximation	62
4.2	The Variational Bayes (VB) Method for the PPCA Model	62
4.3	Orthogonal Variational PCA (OVPCA)	69
4.3.1	The Orthogonal PPCA Model	70
4.3.2	The VB Method for the Orthogonal PPCA Model	70
4.3.3	Inference of Rank	77
4.3.4	Moments of the Model Parameters	78
4.4	Simulation Studies	79
4.4.1	Convergence to Orthogonal Solutions: VPCA vs. FVPCA	79
4.4.2	Local Minima in FVPCA and OVPCA	82
4.4.3	Comparison of Methods for Inference of Rank	83
4.5	Application: Inference of Rank in a Medical Image Sequence	85
4.6	Conclusion	87

5	Functional Analysis of Medical Image Sequences	89
5.1	A Physical Model for Medical Image Sequences	90
5.1.1	Classical Inference of the Physiological Model	92
5.2	The FAMIS Observation Model	92
5.2.1	Bayesian Inference of FAMIS and Related Models	94
5.3	The VB Method for the FAMIS Model	94
5.4	The VB Method for FAMIS: Alternative Priors	99
5.5	Analysis of Clinical Data Using the FAMIS Model	102
5.6	Conclusion	107
6	On-line Inference of Time-Invariant Parameters	109
6.1	Recursive Inference	110
6.2	Bayesian Recursive Inference	110
6.2.1	The Dynamic Exponential Family (DEF)	112
6.2.2	Example: The AutoRegressive (AR) Model	114
6.2.3	Recursive Inference of non-DEF models	117
6.3	The VB Approximation in On-Line Scenarios	118
6.3.1	Scenario I: VB-Marginalization for Conjugate Updates	118
6.3.2	Scenario II: The VB Method in One-Step Approximation	121
6.3.3	Scenario III: Achieving Conjugacy in non-DEF Models via the VB Approximation	123
6.3.4	The VB Method in the On-Line Scenarios	126
6.4	Related Distributional Approximations	127
6.4.1	The Quasi-Bayes (QB) Approximation in On-Line Scenarios	128
6.4.2	Global Approximation via the Geometric Approach	128
6.4.3	One-step Fixed-Form (FF) Approximation	129
6.5	On-line Inference of a Mixture of AutoRegressive (AR) Models	130
6.5.1	The VB Method for AR Mixtures	130
6.5.2	Related Distributional Approximations for AR Mixtures	133
6.5.2.1	The Quasi-Bayes (QB) Approximation	133
6.5.2.2	One-step Fixed-Form (FF) Approximation	135
6.5.3	Simulation Study: On-line Inference of a Static Mixture	135
6.5.3.1	Inference of a Many-Component Mixture	136
6.5.3.2	Inference of a Two-Component Mixture	136
6.5.4	Data-Intensive Applications of Dynamic Mixtures	139
6.5.4.1	Urban Vehicular Traffic Prediction	141
6.6	Conclusion	143
7	On-line Inference of Time-Variant Parameters	145
7.1	Exact Bayesian Filtering	145
7.2	The VB-Approximation in Bayesian Filtering	147
7.2.1	The VB method for Bayesian Filtering	149
7.3	Other Approximation Techniques for Bayesian Filtering	150
7.3.1	Restricted VB (RVB) Approximation	150
7.3.2	Particle Filtering	152

7.3.3	Stabilized Forgetting	153
7.3.3.1	The Choice of the Forgetting Factor	154
7.4	The VB-Approximation in Kalman Filtering	155
7.4.1	The VB method	156
7.4.2	Loss of Moment Information in the VB Approximation	158
7.5	VB-Filtering for the Hidden Markov Model (HMM)	158
7.5.1	Exact Bayesian filtering for known T	159
7.5.2	The VB Method for the HMM Model with Known T	160
7.5.3	The VB Method for the HMM Model with Unknown T	162
7.5.4	Other Approximate Inference Techniques	164
7.5.4.1	Particle Filtering	164
7.5.4.2	Certainty Equivalence Approach	165
7.5.5	Simulation Study: Inference of Soft Bits	166
7.6	The VB-Approximation for an Unknown Forgetting Factor	168
7.6.1	Inference of a Univariate AR Model with Time-Variant Parameters	169
7.6.2	Simulation Study: Non-stationary AR Model Inference via Unknown Forgetting	173
7.6.2.1	Inference of an AR Process with Switching Parameters	173
7.6.2.2	Initialization of Inference for a Stationary AR Process	174
7.7	Conclusion	176
8	The Mixture-based Extension of the AR Model (MEAR)	179
8.1	The Extended AR (EAR) Model	179
8.1.1	Bayesian Inference of the EAR Model	181
8.1.2	Computational Issues	182
8.2	The EAR Model with Unknown Transformation: the MEAR Model	182
8.3	The VB Method for the MEAR Model	183
8.4	Related Distributional Approximations for MEAR	186
8.4.1	The Quasi-Bayes (QB) Approximation	186
8.4.2	The Viterbi-Like (VL) Approximation	187
8.5	Computational Issues	188
8.6	The MEAR Model with Time-Variant Parameters	191
8.7	Application: Inference of an AR Model Robust to Outliers	192
8.7.1	Design of the Filter-bank	192
8.7.2	Simulation Study	193
8.8	Application: Inference of an AR Model Robust to Burst Noise	196
8.8.1	Design of the Filter-Bank	196
8.8.2	Simulation Study	197
8.8.3	Application in Speech Reconstruction	201
8.9	Conclusion	201

9	Concluding Remarks	205
9.1	The VB Method	205
9.2	Contributions of the Work	206
9.3	Current Issues	206
9.4	Future Prospects for the VB Method	207
	Required Probability Distributions	209
A.1	Multivariate Normal distribution	209
A.2	Matrix Normal distribution	209
A.3	Normal-inverse-Wishart ($\mathcal{NIW}_{A,\Omega}$) Distribution	210
A.4	Truncated Normal Distribution	211
A.5	Gamma Distribution	212
A.6	Von Mises-Fisher Matrix distribution	212
	A.6.1 Definition	213
	A.6.2 First Moment	213
	A.6.3 Second Moment and Uncertainty Bounds	214
A.7	Multinomial Distribution	215
A.8	Dirichlet Distribution	215
A.9	Truncated Exponential Distribution	216
	References	217
	Index	225