
Contents

1	Introduction	1
1.1	Motivation	1
1.2	State of the art	3
1.2.1	Operator supports	3
1.2.2	Mainstream multivariate techniques	4
1.2.3	Probabilistic dynamic optimized decision-making	6
1.3	Developed advising and its role in computer support	6
1.4	Presentation style, readership and layout	7
1.5	Acknowledgements	10
2	Underlying theory	11
2.1	General conventions	11
2.2	Basic notions and notations	12
2.3	Decision making under uncertainty	16
2.3.1	Complete ordering of decision rules	17
2.3.2	Calculus with pdfs	21
2.3.3	Basic decision-making lemma	24
2.4	Dynamic design	25
2.4.1	Dynamic programming	25
2.4.2	Fully probabilistic design	28
2.4.3	Asymptotic of the design	31
2.5	Learning	33
2.5.1	Bayesian filtration	33
2.5.2	Bayesian estimation	36
2.5.3	Asymptotic of estimation	38
3	Approximate and feasible learning	43
3.1	Estimation with forgetting	44
3.2	Exponential family	47
3.3	Structure estimation in the nested exponential family	49
3.4	Equivalence approach	50

3.4.1	Recursively feasible representation	51
3.4.2	Approximation as a point estimation	53
3.4.3	Specification of $\mathcal{E}[f_t(\Psi) g_t]$	54
4	Approximate design	57
4.1	Adaptive systems	57
4.2	Suboptimal design	58
4.2.1	Strategies examining reduced space	58
4.2.2	Strategies simplifying models	60
4.3	Decomposition of decision-making	62
4.3.1	Offline phase	63
4.3.2	Online phase	65
5	Problem formulation	67
5.1	Design principle and design conditions	67
5.1.1	Systems and data spaces	68
5.1.2	Basic scenario and design principle	69
5.1.3	Reduction of surplus data of the operator	71
5.1.4	Construction of a true user's ideal pdf	72
5.1.5	Extension of a true user's ideal pdf to the surplus p-data	73
5.2	Learning conditions	74
5.3	Mixtures as approximate models and predictors	79
5.4	Design of advisory systems	82
5.4.1	Types of advisory systems	82
5.4.2	Advices as actions of the p-system	83
5.4.3	Unguided and guided models for respective designs	84
5.4.4	Academic design	85
5.4.5	Industrial design	87
5.4.6	Simultaneous academic and industrial design	88
5.5	Interactions with the operator	89
5.5.1	Assigning priorities	89
5.5.2	Stimulating the operator	91
5.6	Design summary	92
5.6.1	Influence of advices on the o-system	92
5.6.2	Overall scenario and design subtasks	93
6	Solution and principles of its approximation: learning part	95
6.1	Common tools	96
6.1.1	Prediction and model selection	97
6.1.2	Likelihood on variants	98
6.1.3	Branch-and-bound techniques	100
6.2	Data preprocessing	102
6.2.1	Data transformation	103
6.2.2	Outlier removal	104

6.2.3	Filtering	106
6.2.4	Filters generating factors in an exponential family	110
6.2.5	Statistics for the exponential family	112
6.2.6	Prediction in EF with statistics gained by filtering	113
6.3	Use of prior knowledge at the factor level	114
6.3.1	Internally consistent fictitious data blocks	114
6.3.2	Translation of input-output characteristics into data ...	115
6.3.3	Merging of knowledge pieces	117
6.4	Construction of the prior estimate	120
6.4.1	Iterative construction of the prior pdf	120
6.4.2	Common bounding mapping	122
6.4.3	Flattening mapping	123
6.4.4	Geometric mean as branching mapping	134
6.4.5	Random branching of statistics	136
6.4.6	Prior-posterior branching	136
6.4.7	Branching by forgetting	138
6.4.8	Branching by factor splitting	140
6.4.9	Techniques applicable to static mixtures	148
6.5	Approximate parameter estimation	154
6.5.1	Quasi-Bayes estimation	155
6.5.2	EM estimation	159
6.5.3	Batch quasi-Bayes estimation	162
6.6	Structure estimation	164
6.6.1	Estimation of factor structure	165
6.6.2	Structure estimation in factor splitting	166
6.6.3	Estimation of component structure	166
6.6.4	Merging and cancelling of components	167
6.7	Model validation	184
6.7.1	Test of data homogeneity	185
6.7.2	Learning results	187
6.7.3	Forgetting-based validation	190
6.7.4	Inspection by a human designer	191
6.7.5	Operating modes	191
7	Solution and principles of its approximation: design part ..	193
7.1	Common tools	194
7.1.1	Model projections in design	194
7.1.2	Dynamic predictors in advising	197
7.1.3	Advices and their influence	199
7.1.4	Fully probabilistic design in advising	202
7.1.5	Approximations of the KL divergence	204
7.2	Design of advising strategies	210
7.2.1	Academic design	210
7.2.2	Choice of user ideal on pointers	217
7.2.3	Industrial design	222

7.2.4	Simultaneous academic and industrial design	225
7.3	Interaction with an operator	229
7.3.1	Assigning priorities	229
7.3.2	Stimulating the operator	236
7.4	Design validation	241
8	Learning with normal factors and components	243
8.1	Common tools	244
8.1.1	Selected matrix operations	244
8.1.2	$L'DL$ decomposition	245
8.1.3	GiW pdf as a conjugate prior	251
8.1.4	KL divergence of normal pdfs	258
8.1.5	Estimation and prediction with normal factors	259
8.1.6	Estimation and prediction with log-normal factors	261
8.1.7	Relationships of a component to its factors	262
8.1.8	Prediction and model selection	264
8.1.9	Likelihood on variants	265
8.1.10	Branch-and-bound techniques	265
8.2	Data preprocessing	266
8.2.1	Use of physical boundaries	266
8.2.2	Removal of high-frequency noise	266
8.2.3	Suppression of outliers	266
8.3	Use of prior knowledge at the factor level	266
8.3.1	Internally consistent fictitious data blocks	267
8.3.2	Translation of input–output characteristics into data	267
8.3.3	Merging of knowledge pieces	269
8.4	Construction of the prior estimate	270
8.4.1	Common bounding mapping	270
8.4.2	Flattening mapping	271
8.4.3	Geometric mean as branching mapping	272
8.4.4	Random branching of statistics	273
8.4.5	Prior-posterior branching	275
8.4.6	Branching by forgetting	277
8.4.7	Branching by factor splitting	279
8.4.8	Hierarchical and optimization-based splitting	285
8.4.9	Techniques applicable to static mixtures	289
8.5	Approximate parameter estimation	289
8.5.1	Quasi-Bayes estimation	289
8.5.2	EM estimation	291
8.5.3	Batch quasi-Bayes estimation	293
8.6	Structure estimation	295
8.6.1	Estimation of factor structure	295
8.6.2	Structure estimation in factor splitting	297
8.6.3	Estimation of component structure	298
8.6.4	Merging and cancelling of components	299

8.7	Model validation	303
8.7.1	Test of data homogeneity	303
8.7.2	Learning results	305
8.7.3	Forgetting-based validation	306
8.7.4	Inspection by a human designer	307
8.7.5	Operating modes	307
9	Design with normal mixtures	309
9.1	Common tools	311
9.1.1	Model projections in design	311
9.1.2	Dynamic predictors in advising	318
9.1.3	Advices and their influence	324
9.1.4	Practical presentation aspects	324
9.1.5	Quadratic forms in a fully probabilistic design	329
9.2	Design of the advising strategy	338
9.2.1	Academic design	338
9.2.2	Industrial design	349
9.2.3	Simultaneous academic and industrial design	356
9.3	Interaction with an operator	366
9.3.1	Assigning priorities	366
9.3.2	Stimulating the operator	371
10	Learning with Markov-chain factors and components	377
10.1	Common tools	378
10.1.1	Dirichlet pdf as a conjugate prior	378
10.1.2	Estimation and prediction with Markov-chain factors ..	381
10.1.3	Likelihood on variants	382
10.1.4	Branch-and-bound techniques	382
10.2	Data preprocessing	383
10.2.1	Use of physical boundaries	383
10.2.2	Removal of high-frequency noise	385
10.2.3	Suppression of outliers	385
10.2.4	Coding of regression vectors	385
10.2.5	Coding of signal values	386
10.3	Use of prior knowledge at the factor level	387
10.3.1	Internally consistent fictitious data blocks	387
10.3.2	Translation of input-output characteristics into data ..	387
10.3.3	Merging of knowledge pieces	389
10.4	Construction of the prior estimate	389
10.4.1	Iterative construction of prior pdf	390
10.4.2	Common bounding mapping	390
10.4.3	Flattening mapping	390
10.4.4	Geometric mean as branching mapping	391
10.4.5	Random branching of statistics	392
10.4.6	Prior-posterior branching	393

10.4.7	Branching by forgetting	396
10.4.8	Branching by factor splitting	397
10.4.9	Hierarchical selection of split factors	398
10.4.10	Techniques applicable to static mixtures	400
10.5	Approximate parameter estimation	400
10.5.1	Quasi-Bayes estimation	400
10.5.2	EM estimation	401
10.5.3	Batch quasi-Bayes estimation	403
10.6	Structure estimation	404
10.6.1	Estimation of factor structure	404
10.6.2	Estimation of component structure	406
10.6.3	Merging and cancelling of components	406
10.7	Model validation with Markov-chain components	409
10.7.1	Test of data homogeneity	409
10.7.2	Learning results and forgetting-based validation	410
10.7.3	Other indicators of model validity	410
11	Design with Markov-chain mixtures	411
11.1	Common tools	411
11.1.1	Model projections in design	411
11.1.2	Basic operations for fully probabilistic design	415
11.1.3	Dangerous components	418
11.2	Design of the advising strategy	419
11.2.1	Academic design	419
11.2.2	Industrial design	424
11.2.3	Simultaneous academic and industrial design	427
11.3	Interaction with an operator	432
11.3.1	Assigning priorities	432
11.3.2	Stimulating the operator	434
12	Sandwich BMTB for mixture initiation	437
12.1	Common tools	438
12.1.1	Properties of the MT factor	438
12.1.2	KL divergence of MT factors	439
12.1.3	Estimation and prediction with MT factors	440
12.2	Conceptual BMTB algorithm	441
12.3	B-step: preparation of MT parameters	443
12.3.1	Simple choice of box width	444
12.3.2	Centers and box widths via shadow cancelling	446
12.4	MT step: make the MT algorithm feasible	453
12.4.1	Initialization	453
12.4.2	Merging and cancelling of centers	454
12.4.3	Recognition of a false local maximum	454
12.4.4	Improved MT algorithm	455
12.5	-B step: MT results as initial mixture	458

12.5.1	Position and noise covariance	458
12.5.2	Remaining statistics	458
12.5.3	Conversion of static components to dynamic factors....	458
13	Mixed mixtures	463
13.1	Learning with factors on mixed-type quantities	464
13.1.1	Factors in EF with a discrete regressor	464
13.1.2	MT normal factors	465
13.2	An approximate estimation of mixture ratios.....	468
13.2.1	Mixture modelling of stationary pdf	468
13.2.2	Extended quasi-Bayes estimation	470
14	Applications of the advisory system	481
14.1	Operation of a rolling mill	481
14.1.1	Problem description	481
14.1.2	Problem and its solution	483
14.1.3	Implementation	489
14.1.4	Results.....	489
14.1.5	Conclusions.....	491
14.2	Treatment of thyroid gland cancer	492
14.2.1	Problem description	492
14.2.2	Problem and its solution	494
14.2.3	Implementation	497
14.2.4	Results.....	498
14.2.5	Conclusions.....	499
14.3	Prediction of traffic quantities.....	499
14.3.1	Problem description	499
14.3.2	Problem and its solution	500
14.3.3	Experiments	503
14.3.4	Results.....	504
14.3.5	Conclusions.....	505
14.4	Conclusions on the applications	505
14.4.1	Lessons learned	505
14.4.2	Other application areas	506
15	Concluding remarks	507
	References	511
	Index	523



<http://www.springer.com/978-1-85233-928-9>

Optimized Bayesian Dynamic Advising
Theory and Algorithms

Karny, M. (Ed.)

2006, XVII, 529 p., Hardcover

ISBN: 978-1-85233-928-9