

Contents

Notation	xiii
1 Introduction	1
1.1 Aims of modelling	1
1.2 System	1
1.3 Model	3
1.4 Criterion	4
1.5 Optimization	5
1.6 Parameter uncertainty	5
1.7 Critical analysis of the results	6
1.8 In summary	6
2 Structures	7
2.1 Phenomenological and behavioural models	7
2.2 Linear and nonlinear models	9
2.3 Continuous- and discrete-time models	11
2.3.1 Continuous-time models	11
2.3.2 Discrete-time models	12
2.3.3 Sampling	13
2.4 Deterministic and stochastic models	15
2.5 Choice of complexity	19
2.6 Structural properties of models	20
2.6.1 Identifiability	20
2.6.1.1 Laplace transform approach	22
2.6.1.2 Similarity transformation approach	24
2.6.1.3 Taylor series approach	26
2.6.1.4 Local state isomorphism approach	28
2.6.1.5 Use of elimination theory	30
2.6.1.6 Numerical local approach	31
2.6.2 Distinguishability	32
2.6.3 Relationship between identifiability and distinguishability	34
2.6.4 Chemical engineering example	34
2.7 Conclusions	36
3 Criteria	37
3.1 Least squares	37
3.2 Least modulus	39
3.3 Maximum likelihood	40
3.3.1 Output-additive independent random variables	42
3.3.2 Output-additive dependent random variables	49
3.3.3 Properties of maximum-likelihood estimators	51
3.3.4 Estimation of parameter distribution in a population	53

3.3.5 Nonparametric description of structural errors	58
3.4 Complexity	63
3.5 Bayesian criteria	66
3.5.1 Maximum <i>a posteriori</i>	67
3.5.2 Minimum risk	68
3.6 Constraints on parameters	71
3.6.1 Equality constraints	71
3.6.2 Inequality constraints	72
3.7 Robustness	74
3.7.1 Robustness to uncertainty on the noise distribution	74
3.7.2 Breakdown point	76
3.7.3 M-estimators	77
3.7.4 Image processing example	79
3.8 Tuning of hyperparameters	81
3.9 Conclusions	82
4 Optimization	83
4.1 LP structures and quadratic cost functions	84
4.1.1 LP structures	84
4.1.2 Quadratic cost functions	88
4.1.3 Least-squares estimator	88
4.1.3.1 Properties of the least-squares estimator	89
4.1.3.2 Numerical considerations	91
4.1.4 Data-recursive least squares	92
4.1.4.1 $\mathbf{p}^*$ assumed to be constant	92
4.1.4.2 $\mathbf{p}^*$ may drift	96
4.1.4.3 $\mathbf{p}^*$ may jump	97
4.1.4.4 Application to adaptive control	97
4.1.5 Parameter-recursive least squares	101
4.1.6 Kalman filter	102
4.1.6.1 Vector data-recursive least squares	104
4.1.6.2 Static system without process noise	105
4.1.6.3 Dynamic system with process noise	106
4.1.6.4 Off-line computation	108
4.1.6.5 On-line computation	108
4.1.6.6 Influence of the covariances of the process and measurement noise	108
4.1.6.7 Detection of divergence	109
4.1.6.8 Stationary filter	110
4.1.6.9 Use for the choice of sensors	111
4.1.6.10 Extended Kalman filter: real-time parameter estimation	112
4.1.6.11 Stochastic identification	114
4.1.7 Errors-in-variables approach	115
4.2 Least-squares based methods	117
4.2.1 Pseudolinear regression	118
4.2.1.1 Extended least squares	118
4.2.1.2 Properties of extended least squares	119
4.2.2 Multilinear regression	119
4.2.2.1 Generalized least squares	120
4.2.2.2 Properties of generalized least squares	121
4.2.3 Filtering	122
4.2.3.1 Steiglitz and McBride's method	122
4.2.3.2 Extended matrix method	125
4.2.4 First-order expansion of the error	126
4.2.5 Instrumental-variable method	127

4.2.6	Least squares on correlations	129
4.3	General methods	130
4.3.1	Quadratic cost and partially LP structure	131
4.3.2	One-dimensional optimization	131
4.3.2.1	Definition of a search interval	132
4.3.2.2	Dichotomy	133
4.3.2.3	Fibonacci's and golden-section methods	133
4.3.2.4	Parabolic interpolation	136
4.3.2.5	Which method?	137
4.3.2.6	Combining one-dimensional optimizations	137
4.3.3	Limited expansions of the cost	142
4.3.3.1	Gradient method	142
4.3.3.2	Computation of the gradient	149
4.3.3.3	Newton method	167
4.3.3.4	Gauss-Newton method	171
4.3.3.5	Levenberg-Marquardt method	173
4.3.3.6	Quasi-Newton methods	174
4.3.3.7	Heavy-ball method	178
4.3.3.8	Conjugate-gradient methods	178
4.3.3.9	Choice of step size	182
4.3.4	Constrained optimization	184
4.3.4.1	Linear programming	185
4.3.4.2	Quadratic programming	187
4.3.4.3	Constrained gradient	189
4.3.4.4	Gradient-projection method	190
4.3.4.5	Constrained Newton and quasi-Newton	192
4.3.4.6	Method of centres	194
4.3.4.7	Method of feasible directions	196
4.3.5	Non-differentiable cost functions	197
4.3.5.1	Subgradient method	197
4.3.5.2	Cutting-plane method	199
4.3.5.3	Ellipsoidal method	200
4.3.5.4	Application to L_1 estimation	201
4.3.6	Initialisation	203
4.3.7	Termination	204
4.3.8	Recursive techniques	206
4.3.9	Global optimization	211
4.3.9.1	Eliminating parasitic local optima	211
4.3.9.2	Random search	216
4.3.9.3	Deterministic search	219
4.4	Optimization of a measured response	226
4.4.1	Model-free optimization	226
4.4.2	Response-surface methodology	227
4.5	Conclusions	229
5	Uncertainty	231
5.1	Cost contours in parameter space	231
5.1.1	Normal noise: cost contours, confidence regions	231
5.1.1.1	Noise with known variance	232
5.1.1.2	Noise with unknown variance	235
5.1.1.3	Noise with independently estimated variance	237
5.1.2	Determination of points on a cost contour	238
5.1.3	Characterization of non-connected domains	240
5.1.4	Representation of cost contours	240

5.2 Monte-Carlo methods	242
5.2.1 Principle	242
5.2.2 Number of significant digits of the estimate: the CESTAC method	243
5.2.3 Generating fictitious data by jack-knife and bootstrap	243
5.2.3.1 Jack-knife	244
5.2.3.2 Bootstrap	244
5.3 Methods based on the density of the estimator	245
5.3.1 Non-Bayesian estimators	245
5.3.1.1 Cramér-Rao inequality	246
5.3.1.2 LP model structure and normal noise with known covariance	246
5.3.1.3 LP model structure and normal noise with unknown variance	249
5.3.1.4 Other cases	250
5.3.2 Bayesian estimators	253
5.3.3 Approximation of the probability density of the estimator	254
5.4 Bounded-error set estimation	257
5.4.1 LP model structures	259
5.4.1.1 Recursive determination of outer ellipsoids	260
5.4.1.2 Non-recursive determination of outer boxes.....	269
5.4.1.3 Exact description	270
5.4.2 Non-LP model structures	272
5.4.2.1 Errors in variables	274
5.4.2.2 Outlier minimal number estimator	276
5.4.2.3 Set inversion	280
5.5 Conclusions	283
6 Experiments	285
6.1 Criteria	287
6.2 Local design	291
6.2.1 Exact design	292
6.2.1.1 Fedorov's algorithm	292
6.2.1.2 DETMAX algorithm	294
6.2.2 Distribution of experimental effort	295
6.2.2.1 Continuous design	295
6.2.2.2 Approximate design	296
6.2.2.3 Properties of optimal experiments	297
6.2.2.4 Algorithms	303
6.3 Applications	306
6.3.1 Optimal measurement times	306
6.3.2 Optimal inputs	306
6.3.2.1 Parametric inputs	307
6.3.2.2 Nonparametric inputs	308
6.3.3 Simultaneous choice of inputs and sampling times	326
6.4 Robust design	329
6.4.1 Limitations of local design	329
6.4.2 Sequential design	331
6.4.3 Average optimality	333
6.4.3.1 Criteria	333
6.4.3.2 Algorithms	336
6.4.4 Minimax optimality	338
6.4.4.1 Criteria	338
6.4.4.2 Algorithms	340
6.5 Design for Bayesian estimation	342
6.5.1 Exact design	343
6.5.2 Approximate design	344

- 6.6 Influence of model structure 345
 - 6.6.1 Robustness through Bayesian estimation 346
 - 6.6.2 Robust estimation and design 347
 - 6.6.2.1 Minimax approach 348
 - 6.6.2.2 Bayesian approach 349
 - 6.6.3 Structure discrimination 349
 - 6.6.3.1 Discriminating by prediction discrepancy 350
 - 6.6.3.2 Discriminating *via* entropy 352
 - 6.6.3.3 Discriminating *via* D_S -optimality 354
 - 6.6.3.4 Possible extensions 355
- 6.7 Conclusions 356
- 7 Falsification** 359
 - 7.1 Simple inspection 359
 - 7.2 Statistical analysis of residuals 360
 - 7.2.1 Testing for normality 362
 - 7.2.2 Testing for stationarity 372
 - 7.2.3 Testing for independence 376
 - 7.3 Conclusions 380
- References** 381
- Index** 405