
PARAMETER ESTIMATION FOR SCIENTISTS AND ENGINEERS

Adriaan van den Bos



WILEY-INTERSCIENCE
A John Wiley & Sons, Inc., Publication

CONTENTS

Preface	xiii
1 Introduction	1
2 Parametric Models of Observations	7
2.1 Introduction	7
2.2 Purposes of model parameter estimation	8
2.3 Traditional deterministic parameter estimation methods	9
2.4 Statistical parametric models of observations	12
2.4.1 The expectation model	12
2.4.2 Advantages of statistical parametric models of observations	16
2.5 Conclusions	20
2.6 Comments and references	20
3 Distributions of Observations	21
3.1 Introduction	21
3.2 Expectation, covariance, and Fisher score	21
3.3 The joint real normal distribution	24
3.3.1 The joint real normal probability density function	24
3.3.2 The Fisher score of normal observations	25
	vii

3.4	The Poisson distribution	25
3.4.1	The Poisson probability function	25
3.4.2	The Fisher score of Poisson observations	26
3.5	The multinomial distribution	27
3.5.1	The multinomial probability function	27
3.5.2	The Fisher score of multinomial observations	29
3.6	Exponential families of distributions	30
3.6.1	Definition and examples	30
3.6.2	Properties of exponential families of distributions	32
3.7	Statistical properties of Fisher scores	35
3.8	Complex stochastic variables	37
3.8.1	Scalar complex stochastic variables	37
3.8.2	Vectors of complex stochastic variables	38
3.8.3	Vectors of real and complex stochastic variables	39
3.9	The joint real–complex normal distribution	40
3.10	Comments and references	41
3.11	Problems	42
4	Precision and Accuracy	45
4.1	Introduction	45
4.2	Properties of estimators	46
4.3	Properties of covariance matrices	48
4.3.1	Real covariance matrices	48
4.3.2	Complex covariance matrices	50
4.4	Fisher information	51
4.4.1	Definition of the Fisher information matrix	51
4.4.2	The Fisher information matrix for exponential families of distributions	55
4.4.3	Inflow of Fisher information	56
4.5	Limits to precision: The Cramér–Rao lower bound	60
4.5.1	The Cramér–Rao lower bound for scalar functions of scalar parameters	60
4.5.2	The Cramér–Rao lower bound for vector functions of vector parameters	63
4.6	Properties of the Cramér–Rao lower bound	66
4.6.1	Interpretation of the expression for the Cramér–Rao lower bound	66
4.6.2	The Cramér–Rao lower bound as a measure of efficiency of estimation	67
4.6.3	Monotonicity with the number of observations	70
4.6.4	Propagation of standard deviation	70
4.6.5	Influence of estimation of additional parameters	71
4.6.6	The Cramér–Rao lower bound for biased estimators	72

4.7	The Cramér–Rao lower bound for complex parameters or functions of parameters	74
4.7.1	Introduction	74
4.7.2	The Cramér–Rao lower bound for vectors of real and complex functions of real parameters	75
4.7.3	The Cramér–Rao lower bound for vectors of real and complex functions of real and complex parameters	76
4.8	The Cramér–Rao lower bound for exponential families of distributions	78
4.9	The Cramér–Rao lower bound and identifiability	79
4.10	The Cramér–Rao lower bound and experimental design	81
4.10.1	Introduction	81
4.10.2	Experimental design for nonlinear vector parameters	84
4.11	Comments and references	91
4.12	Problems	91
5	Precise and Accurate Estimation	99
5.1	Introduction	99
5.2	Maximum likelihood estimation	100
5.3	Properties of maximum likelihood estimators	105
5.3.1	The invariance property of maximum likelihood estimators	105
5.3.2	Connection of efficient unbiased estimators and maximum likelihood estimators	106
5.3.3	Consistency of maximum likelihood estimators	106
5.3.4	Asymptotic normality of maximum likelihood estimators	110
5.3.5	Asymptotic efficiency of maximum likelihood estimators	112
5.4	Maximum likelihood for normally distributed observations	113
5.4.1	The likelihood function for normally distributed observations	113
5.4.2	Properties of maximum likelihood estimators for normally distributed observations	115
5.4.3	Maximum likelihood estimation of the parameters of linear models from normally distributed observations	121
5.5	Maximum likelihood for Poisson distributed observations	123
5.6	Maximum likelihood for multinomially distributed observations	124
5.7	Maximum likelihood for exponential family distributed observations	125
5.8	Testing the expectation model: The likelihood ratio test	126
5.8.1	Model testing for arbitrary distributions	126
5.8.2	Model testing for exponential families of distributions	132
5.9	Least squares estimation	134
5.10	Nonlinear least squares estimation	136
5.11	Linear least squares estimation	139
5.12	Weighted linear least squares estimation	140
5.13	Properties of the linear least squares estimator	144

5.14	The best linear unbiased estimator	145
5.15	Special cases of the best linear unbiased estimator and a related result	147
5.15.1	The Gauss–Markov theorem	147
5.15.2	Normally distributed observations	148
5.15.3	Exponential family distributed observations	148
5.16	Complex linear least squares estimation	149
5.17	Summary of properties of linear least squares estimators	151
5.18	Recursive linear least squares estimation	152
5.19	Recursive linear least squares estimation with forgetting	155
5.20	Comments and references	157
5.21	Problems	158
6	Numerical Methods for Parameter Estimation	163
6.1	Introduction	163
6.2	Numerical optimization	164
6.2.1	Key notions in numerical optimization	164
6.2.2	Reference log-likelihood functions and least squares criteria	166
6.3	The steepest descent method	168
6.3.1	Definition of the steepest descent step	168
6.3.2	Properties of the steepest descent step	171
6.4	The Newton method	174
6.4.1	Definition of the Newton step	174
6.4.2	Properties of the Newton step	175
6.4.3	The Newton step for maximizing log-likelihood functions	182
6.5	The Fisher scoring method	183
6.5.1	Definition of the Fisher scoring step	183
6.5.2	Properties of the Fisher scoring step	183
6.5.3	Fisher scoring step for exponential families	185
6.6	The Newton method for normal maximum likelihood and for nonlinear least squares	185
6.6.1	The Newton step for normal maximum likelihood	185
6.6.2	The Newton step for nonlinear least squares	187
6.7	The Gauss–Newton method	188
6.7.1	Definition of the Gauss–Newton step	188
6.7.2	Properties of the Gauss–Newton step	188
6.8	The Newton method for Poisson maximum likelihood	191
6.9	The Newton method for multinomial maximum likelihood	192
6.10	The Newton method for exponential family maximum likelihood	193
6.11	The generalized Gauss–Newton method for exponential family maximum likelihood	194
6.11.1	Definition of the generalized Gauss–Newton step	194
6.11.2	Properties of the generalized Gauss–Newton method	195

6.12	The iteratively reweighted least squares method	197
6.13	The Levenberg–Marquardt method	197
6.13.1	Definition of the Levenberg–Marquardt step	197
6.13.2	Properties of the Levenberg–Marquardt step	199
6.14	Summary of the described numerical optimization methods	200
6.14.1	Introduction	200
6.14.2	The steepest ascent (descent) method	200
6.14.3	The Newton method	200
6.14.4	The Fisher scoring method	201
6.14.5	The Gauss–Newton method	201
6.14.6	The generalized Gauss–Newton method	202
6.14.7	The iteratively reweighted least squares method	202
6.14.8	The Levenberg–Marquardt method	202
6.14.9	Conclusions	202
6.15	Parameter estimation methodology	203
6.15.1	Introduction	203
6.15.2	Investigating the feasibility of the observations	203
6.15.3	Preliminary simulation experiments	204
6.16	Comments and references	206
6.17	Problems	206
7	Solutions or Partial Solutions to Problems	211
	Appendix A: Statistical Results	247
A.1	Statistical properties of linear combinations of stochastic variables	247
A.2	The Cauchy–Schwarz inequality for expectations	249
	Appendix B: Vectors and Matrices	251
B.1	Vectors	251
B.2	Matrices	252
	Appendix C: Positive Semidefinite and Positive Definite Matrices	259
C.1	Real positive semidefinite and positive definite matrices	259
C.2	Complex positive semidefinite and positive definite matrices	263
	Appendix D: Vector and Matrix Differentiation	265
	References	269
	Topic Index	271