Classification, Parameter Estimation and State Estimation

Classification, Parameter Estimation and State Estimation: An Engineering Approach using MATLAB
F. van der Heijden, R.P.W. Duin, D. de Ridder and D.M.J. Tax
© 2004 John Wiley & Sons, Ltd ISBN: 0-470-09013-8

Classification, Parameter Estimation and State Estimation

An Engineering Approach using MATLAB®

F. van der Heijden
Faculty of Electrical Engineering, Mathematics and Computer Science
University of Twente
The Netherlands

R.P.W. Duin
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
The Netherlands

D. de Ridder
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
The Netherlands

D.M.J. Tax
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
The Netherlands

John Wiley & Sons, Ltd

Copyright © 2004 John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wileyeurope.com or www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The Publisher is not associated with any product or vendor mentioned in this book.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 33 Park Road, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 22 Worcester Road, Etobicoke, Ontario, Canada M9W 1L1

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging in Publication Data

Classification, parameter estimation and state estimation : an engineering approach using MATLAB / F. van der Heijden . . . [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 0-470-09013-8 (cloth : alk. paper)
1. Engineering mathematics—Data processing. 2. MATLAB. 3. Mensuration—Data processing. 4. Estimation theory—Data processing. I. Heijden, Ferdinand van der.
TA331.C53 2004
681′.2—dc22

2004011561

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-470-09013-8

Typeset in 10.5/13pt Sabon by Integra Software Services Pvt. Ltd, Pondicherry, India
Printed and bound in Great Britain by TJ International Ltd, Padstow, Cornwall
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

Contents

Preface xi
Foreword xv

1 Introduction **1**
1.1 The scope of the book 2
1.1.1 Classification 3
1.1.2 Parameter estimation 4
1.1.3 State estimation 5
1.1.4 Relations between the subjects 6
1.2 Engineering 9
1.3 The organization of the book 11
1.4 References 12

2 Detection and Classification **13**
2.1 Bayesian classification 16
2.1.1 Uniform cost function and minimum error rate 23
2.1.2 Normal distributed measurements; linear and quadratic classifiers 25
2.2 Rejection 32
2.2.1 Minimum error rate classification with reject option 33
2.3 Detection: the two-class case 35
2.4 Selected bibliography 43
2.5 Exercises 43

3 Parameter Estimation **45**
3.1 Bayesian estimation 47
3.1.1 MMSE estimation 54

3.1.2 MAP estimation 55
3.1.3 The Gaussian case with linear sensors 56
3.1.4 Maximum likelihood estimation 57
3.1.5 Unbiased linear MMSE estimation 59
3.2 Performance of estimators 62
3.2.1 Bias and covariance 63
3.2.2 The error covariance of the unbiased linear MMSE estimator 67
3.3 Data fitting 68
3.3.1 Least squares fitting 68
3.3.2 Fitting using a robust error norm 72
3.3.3 Regression 74
3.4 Overview of the family of estimators 77
3.5 Selected bibliography 79
3.6 Exercises 79

4 State Estimation 81
4.1 A general framework for online estimation 82
4.1.1 Models 83
4.1.2 Optimal online estimation 86
4.2 Continuous state variables 88
4.2.1 Optimal online estimation in linear-Gaussian systems 89
4.2.2 Suboptimal solutions for nonlinear systems 100
4.2.3 Other filters for nonlinear systems 112
4.3 Discrete state variables 113
4.3.1 Hidden Markov models 113
4.3.2 Online state estimation 117
4.3.3 Offline state estimation 120
4.4 Mixed states and the particle filter 128
4.4.1 Importance sampling 128
4.4.2 Resampling by selection 130
4.4.3 The condensation algorithm 131
4.5 Selected bibliography 135
4.6 Exercises 136

5 Supervised Learning 139
5.1 Training sets 140
5.2 Parametric learning 142
5.2.1 Gaussian distribution, mean unknown 143

5.2.2 Gaussian distribution, covariance matrix unknown 144
5.2.3 Gaussian distribution, mean and covariance matrix both unknown 145
5.2.4 Estimation of the prior probabilities 147
5.2.5 Binary measurements 148
5.3 Nonparametric learning 149
5.3.1 Parzen estimation and histogramming 150
5.3.2 Nearest neighbour classification 155
5.3.3 Linear discriminant functions 162
5.3.4 The support vector classifier 168
5.3.5 The feed-forward neural network 173
5.4 Empirical evaluation 177
5.5 References 181
5.6 Exercises 181

6 Feature Extraction and Selection 183
6.1 Criteria for selection and extraction 185
6.1.1 Inter/intra class distance 186
6.1.2 Chernoff–Bhattacharyya distance 191
6.1.3 Other criteria 194
6.2 Feature selection 195
6.2.1 Branch-and-bound 197
6.2.2 Suboptimal search 199
6.2.3 Implementation issues 201
6.3 Linear feature extraction 202
6.3.1 Feature extraction based on the Bhattacharyya distance with Gaussian distributions 204
6.3.2 Feature extraction based on inter/intra class distance 209
6.4 References 213
6.5 Exercises 214

7 Unsupervised Learning 215
7.1 Feature reduction 216
7.1.1 Principal component analysis 216
7.1.2 Multi-dimensional scaling 220
7.2 Clustering 226
7.2.1 Hierarchical clustering 228
7.2.2 K-means clustering 232

7.2.3 Mixture of Gaussians 234
7.2.4 Mixture of probabilistic PCA 240
7.2.5 Self-organizing maps 241
7.2.6 Generative topographic mapping 246
7.3 References 250
7.4 Exercises 250

8 State Estimation in Practice 253
8.1 System identification 256
8.1.1 Structuring 256
8.1.2 Experiment design 258
8.1.3 Parameter estimation 259
8.1.4 Evaluation and model selection 263
8.1.5 Identification of linear systems with a random input 264
8.2 Observability, controllability and stability 266
8.2.1 Observability 266
8.2.2 Controllability 269
8.2.3 Dynamic stability and steady state solutions 270
8.3 Computational issues 276
8.3.1 The linear-Gaussian MMSE form 280
8.3.2 Sequential processing of the measurements 282
8.3.3 The information filter 283
8.3.4 Square root filtering 287
8.3.5 Comparison 291
8.4 Consistency checks 292
8.4.1 Orthogonality properties 293
8.4.2 Normalized errors 294
8.4.3 Consistency checks 296
8.4.4 Fudging 299
8.5 Extensions of the Kalman filter 300
8.5.1 Autocorrelated noise 300
8.5.2 Cross-correlated noise 303
8.5.3 Smoothing 303
8.6 References 306
8.7 Exercises 307

9 Worked Out Examples 309
9.1 Boston Housing classification problem 309
9.1.1 Data set description 309
9.1.2 Simple classification methods 311

9.1.3 Feature extraction 312
9.1.4 Feature selection 314
9.1.5 Complex classifiers 316
9.1.6 Conclusions 319
9.2 Time-of-flight estimation of an acoustic tone burst 319
9.2.1 Models of the observed waveform 321
9.2.2 Heuristic methods for determining the ToF 323
9.2.3 Curve fitting 324
9.2.4 Matched filtering 326
9.2.5 ML estimation using covariance models for the reflections 327
9.2.6 Optimization and evaluation 332
9.3 Online level estimation in an hydraulic system 339
9.3.1 Linearized Kalman filtering 341
9.3.2 Extended Kalman filtering 343
9.3.3 Particle filtering 344
9.3.4 Discussion 350
9.4 References 352

Appendix A Topics Selected from Functional Analysis 353
A.1 Linear spaces 353
A.1.1 Normed linear spaces 355
A.1.2 Euclidean spaces or inner product spaces 357
A.2 Metric spaces 358
A.3 Orthonormal systems and Fourier series 360
A.4 Linear operators 362
A.5 References 366

Appendix B Topics Selected from Linear Algebra and Matrix Theory 367
B.1 Vectors and matrices 367
B.2 Convolution 370
B.3 Trace and determinant 372
B.4 Differentiation of vector and matrix functions 373
B.5 Diagonalization of self-adjoint matrices 375
B.6 Singular value decomposition (SVD) 378
B.7 References 381

Appendix C Probability Theory 383
C.1 Probability theory and random variables 383
C.1.1 Moments 386

C.1.2 Poisson distribution 387
C.1.3 Binomial distribution 387
C.1.4 Normal distribution 388
C.1.5 The Chi-square distribution 389
C.2 Bivariate random variables 390
C.3 Random vectors 395
C.3.1 Linear operations on Gaussian random vectors 396
C.3.2 Decorrelation 397
C.4 Reference 398

Appendix D Discrete-time Dynamic Systems 399
D.1 Discrete-time dynamic systems 399
D.2 Linear systems 400
D.3 Linear time invariant systems 401
D.3.1 Diagonalization of a system 401
D.3.2 Stability 402
D.4 References 403

Appendix E Introduction to PRTools 405
E.1 Motivation 405
E.2 Essential concepts in PRTools 406
E.3 Implementation 407
E.4 Some details 410
E.4.1 Data sets 410
E.4.2 Classifiers and mappings 411
E.5 How to write your own mapping 414

Appendix F MATLAB Toolboxes Used 417

Index 419