LINEAR ESTIMATION

T. Kailath
Stanford University

A. H. Sayed
UCLA

B. Hassibi
Bell Laboratories

©2000 All rights reserved.

To our parents and our families.
CONTENTS

PREFACE xix

SYMBOLS xxiii

1 OVERVIEW 1
   1.1 The Asymptotic Observer 2
   1.2 The Optimum Transient Observer 4
       1.2.1 The Mean-Square-Error Criterion 6
       1.2.2 Minimization via Completion of Squares 8
       1.2.3 The Optimum Transient Observer 10
       1.2.4 The Kalman Filter 11
   1.3 Coming Attractions 13
       1.3.1 Smoothed Estimators 13
       1.3.2 Extensions to Time-Variant Models 13
       1.3.3 Fast Algorithms for Time-Invariant Systems 14
       1.3.4 Numerical Issues 15
       1.3.5 Array Algorithms 16
       1.3.6 Other Topics 17
   1.4 The Innovations Process 17
       1.4.1 Whiteness of the Innovations Process 18
       1.4.2 Innovations Representations 20
       1.4.3 Canonical Covariance Factorization 21
       1.4.4 Exploiting State-Space Structure for Matrix Problems 22
   1.5 Steady-State Behavior 23
       1.5.1 Appropriate Solutions of the DARE 23
       1.5.2 Wiener Filters 25

v
1.5.3 Convergence Results 28
1.6 Several Related Problems 29
  1.6.1 Adaptive RLS Filtering 29
  1.6.2 Linear Quadratic Control 30
  1.6.3 $H_\infty$ Estimation 31
  1.6.4 $H_\infty$ Adaptive Filtering 34
  1.6.5 $H_\infty$ Control 36
  1.6.6 Linear Algebra and Matrix Theory 38
1.7 Complements 38
1.8 Problems 40

2 DETERMINISTIC LEAST-SQUARES PROBLEMS 43
  2.1 The Deterministic Least-Squares Criterion 44
  2.2 The Classical Solution 45
    2.2.1 The Normal Equations 45
    2.2.2 Weighted Least-Squares Problems 47
    2.2.3 Statistical Assumptions on the Noise 47
  2.3 A Geometric Formulation: The Orthogonality Condition 48
    2.3.1 The Projection Theorem in Inner Product Spaces 50
    2.3.2 Geometric Insights 51
    2.3.3 Projection Matrices 52
    2.3.4 An Application: Order-Recursive Least-Squares 53
  2.4 Regularized Least-Squares Problems 54
  2.5 An Array Algorithm: The QR Method 56
  2.6 Updating Least-Squares Solutions: RLS Algorithms 59
    2.6.1 The RLS Algorithm 59
    2.6.2 An Array Algorithm for RLS 62
  2.7 Downdating Least-Squares Solutions 63
  2.8 Some Variations of Least-Squares Problems 65
    2.8.1 The Total Least-Squares Criterion 66
    2.8.2 Criteria with Bounds on Data Uncertainties 67
  2.9 Complements 70
  2.10 Problems 71
  2.A On Systems of Linear Equations 78

3 STOCHASTIC LEAST-SQUARES PROBLEMS 82
  3.1 The Problem of Stochastic Estimation 83
  3.2 Linear Least-Mean-Squares Estimators 84
3.2.1 The Fundamental Equations 84
3.2.2 Stochastic Interpretation of Triangular Factorization 86
3.2.3 Singular Data Covariance Matrices 88
3.2.4 Nonzero-Mean Values and Centering 89
3.2.5 Estimators for Complex-Valued Random Variables 89
3.3 A Geometric Formulation 93
  3.3.1 The Orthogonality Condition 93
  3.3.2 Examples 97
3.4 Linear Models 100
  3.4.1 Information Forms When \( R_x > 0 \) and \( R_e > 0 \) 100
  3.4.2 The Gauss-Markov Theorem 101
  3.4.3 Combining Estimators 103
3.5 Equivalence to Deterministic Least-Squares 104
3.6 Complements 107
3.7 Problems 108
3.A Least-Mean-Squares Estimation 118
3.B Gaussian Random Variables 119
3.C Optimal Estimation for Gaussian Variables 121

4 THE INNOVATIONS PROCESS 123
4.1 Estimation of Stochastic Processes 124
  4.1.1 The Fixed Interval Smoothing Problem 125
  4.1.2 The Causal Filtering Problem 126
  4.1.3 The Wiener-Hopf Technique 127
  4.1.4 A Note on Terminology —— Vectors and Gramians 130
4.2 The Innovations Process 131
  4.2.1 A Geometric Approach 131
  4.2.2 An Algebraic Approach 133
  4.2.3 The Modified Gram-Schmidt Procedure 136
  4.2.4 Estimation Given the Innovations Process 136
  4.2.5 The Filtering Problem via the Innovations Approach 137
  4.2.6 Computational Issues 138
4.3 Innovations Approach to Deterministic Least-Squares Problems 139
4.4 The Exponentially Correlated Process 140
  4.4.1 Triangular Factorization of \( R_y \) 140
  4.4.2 Finding \( L^{-1} \) and the Innovations 143
  4.4.3 Innovations via the Gram-Schmidt Procedures 143
5 STATE-SPACE MODELS
5.1 The Exponentially Correlated Process 158
  5.1.1 Finite Interval Problems; Initial Conditions for Stationarity 158
  5.1.2 Innovations from the Process Model 160
5.2 Going Beyond the Stationary Case 161
  5.2.1 Stationary Processes 162
  5.2.2 Nonstationary Processes 163
5.3 Higher-Order Processes and State-Space Models 163
  5.3.1 Autoregressive Processes 164
  5.3.2 Handling Initial Conditions 164
  5.3.3 State-Space Descriptions 165
  5.3.4 The Standard State-Space Model 167
  5.3.5 Examples of Other State-Space Models 169
5.4 Wide-Sense Markov Processes 170
  5.4.1 Forwards Markovian Models 171
  5.4.2 Backwards Markovian Models 173
  5.4.3 Backwards Models from Forwards Models 175
  5.4.4 Markovian Representations and the Standard Model 178
5.5 Complements 180
5.6 Problems 181
5.A Some Global Formulas 186

6 INNOVATIONS FOR STATIONARY PROCESSES 190
6.1 Innovations via Spectral Factorization 190
  6.1.1 Stationary Processes 191
  6.1.2 Generating Functions and z-Spectra 193
6.2 Signals and Systems 196
  6.2.1 The z-Transform 196
  6.2.2 Linear Time-Invariant Systems 198
  6.2.3 Causal, Anticausal, and Minimum-Phase Systems 199
6.3 Stationary Random Processes 200
  6.3.1 Properties of the z-Spectrum 201
  6.3.2 Linear Operations on Stationary Stochastic Processes 202
6.4 Canonical Spectral Factorization 204
6.5 Scalar Rational $z$-Spectra 208
6.6 Vector-Valued Stationary Processes 211
6.7 Complements 214
6.8 Problems 215
6.9 Continuous-Time Systems and Processes 224

7 WIENER THEORY FOR SCALAR PROCESSES 229
7.1 Continuous-Time Wiener Smoothing 229
  7.1.1 The Geometric Formulation 231
  7.1.2 Solution via Fourier Transforms 232
  7.1.3 The Minimum Mean-Square Error 233
  7.1.4 Filtering Signals out of Noisy Measurements 234
  7.1.5 Comparison with the Ideal Filter 234
7.2 The Continuous-Time Wiener-Hopf Equation 235
7.3 Discrete-Time Problems 237
  7.3.1 The Discrete-Time Wiener Smoother 237
  7.3.2 The Discrete-Time Wiener-Hopf Equation 238
7.4 The Discrete Time Wiener-Hopf Technique 240
7.5 Causal Parts via Partial Fractions 243
7.6 Important Special Cases and Examples 246
  7.6.1 Pure Prediction 246
  7.6.2 Additive White Noise 249
7.7 Innovations Approach to the Wiener Filter 252
  7.7.1 The Pure Prediction Problem 254
  7.7.2 Additive White-Noise Problems 255
7.8 Vector Processes 256
7.9 Extensions of Wiener Filtering 257
7.10 Complements 259
7.11 Problems 260
7.12 The Continuous-Time Wiener-Hopf Technique 272

8 RECURSIVE WIENER FILTERING 275
8.1 Time-Invariant State-Space Models 275
  8.1.1 Covariance Functions for Time-Invariant Models 276
  8.1.2 The Special Case of Stationary Processes 276
  8.1.3 Expressions for the $z$-Spectrum 278
8.2 An Equivalence Class for Input Gramians 279
8.3 Canonical Spectral Factorization 282
8.3.1 Unit-Circle Controllability Condition 282
8.3.2 An Inertia Property 284
8.3.3 Algebraic Riccati Equations and Spectral Factorization 285
8.3.4 Appropriate Solutions of the DARE 286
8.3.5 Canonical Spectral Factorization and Innovations Models 288
8.3.6 A Digression: A Criterion for Positivity 290
8.4 Recursive Estimation Given State-Space Models 291
8.4.1 Recursive Predictors 291
8.4.2 Recursive State Predictors 291
8.4.3 Recursive Smoothed Estimators 292
8.5 Factorization Given Covariance Data: Recursive Wiener Filters 294
8.6 Extension to Time-Variant Models 296
8.7 The Appendices 298
8.8 Complements 298
8.9 Problems 299
8.A The Popov function 303
8.C The KYP and Related Lemmas 310
8.D Vector Spectral Factorization in Continuous Time 313

9 THE KALMAN FILTER 321
9.1 The Standard State-Space Model 321
9.2 The Kalman Filter Recursions for the Innovations 323
9.2.1 Recursions for the Innovations 323
9.2.2 $R_{i,i}$ and $K_{p,i}$ in Terms of $P_i$ 325
9.2.3 Recursion for $P_i$ 327
9.2.4 The Kalman Filter Recursions for the Innovations 328
9.2.5 Innovations Models for the Output Process 329
9.3 Recursions for Predicted and Filtered State Estimators 330
9.3.1 The Predicted Estimators 330
9.3.2 Schmidt’s Modification: Measurement and Time Updates 331
9.3.3 Recursions for Filtered Estimators 333
9.3.4 An Alternative Innovations Model 335
9.4 Triangular Factorizations of $R_y$ and $R_y^{-1}$ 335
9.5 An Important Special Assumption: $R_i > 0$ 337
9.5.1 Simplifications for Correlated Noise Processes 337
9.5.2 Measurement Updates in Information Form 339
9.5.3 Existence of $P_i^{-1}$ 341
9.5.4 Sequential Processing 342
9.5.5 Time Updates in Information Form ($Q_i > 0$) 343
9.5.6 A Recursion for $P_i^{-1}$ 344
9.5.7 Summary of Results under Invertibility Conditions 344
9.6 Covariance-Based Filters 346
9.7 Approximate Nonlinear Filtering 349
  9.7.1 A Linearized Kalman Filter 351
  9.7.2 The Schmidt Extended Kalman Filter (EKF) 352
  9.7.3 The Iterated Schmidt EKF 353
  9.7.4 Performance of the Approximate Filters 354
  9.7.5 Other Schemes 354
9.8 Backwards Kalman Recursions 355
  9.8.1 Backwards Markovian Representations of $\{y_i\}$ 355
  9.8.2 Recursions for the Backwards Innovations Process 356
  9.8.3 The Filtered Version of the Backwards Kalman Recursions 357
  9.8.4 $UDU^*$ Factorization of $R_y$ 358
9.9 Complements 359
9.10 Problems 363
9.A Factorization of $R_y$ Using the MGS Procedure 374
9.B Factorization via Gramian Equivalence Classes 377

10 SMOOTHED ESTIMATORS 382
  10.1 General Smoothing Formulas 382
  10.2 Exploiting State-Space Structure 385
    10.2.1 The Bryson-Frazier (BF) Formulas 385
    10.2.2 Stochastic Interpretation of the Adjoint Variable 387
  10.3 The Rauch-Tung-Striebel (RTS) Recursions 388
    10.3.1 First Form of RTS Recursions 388
    10.3.2 The Smoothing Errors are Backwards Markov 390
    10.3.3 The Original Rauch-Tung-Striebel (RTS) Formulas 391
  10.4 Two-Filter Formulas 392
    10.4.1 General Two Filter Formulas 393
    10.4.2 The Mayne and Fraser-Potter Formulas 394
    10.4.3 Combined Estimators Derivation 396
  10.5 The Hamiltonian Equations ($R_i > 0$) 398
  10.6 Variational Origin of Hamiltonian Equations 400
10.7 Applications of Equivalence 402
   10.7.1 The Equivalent Stochastic Problem 402
   10.7.2 Solving the Stochastic Problem 403
   10.7.3 Solving the Deterministic Problem 404
   10.7.4 An Alternative Direct Solution 405
   10.7.5 MAP Estimation and a Deterministic Interpretation for the
        Kalman Filter 406
   10.7.6 The Deterministic Approach of Whittle 407
10.8 Complements 410
10.9 Problems 411

11 FAST ALGORITHMS 419
   11.1 The Fast (CKMS) Recursions 419
   11.2 Two Important Cases 426
      11.2.1 Zero Initial Conditions 426
      11.2.2 Stationary Processes 427
   11.3 Structured Time-Variant Systems 428
   11.4 CKMS Recursions Given Covariance Data 430
   11.5 Relation to Displacement Rank 432
   11.6 Complements 435
   11.7 Problems 436

12 ARRAY ALGORITHMS 441
   12.1 Review and Notations 442
      12.1.1 Notation 443
      12.1.2 Normalizations 444
      12.1.3 A Demonstration of Round-Off Error Effects 445
   12.2 Potter’s Explicit Algorithm for Scalar Measurement Update 446
   12.3 Several Array Algorithms 448
      12.3.1 A Standing Assumption 448
      12.3.2 Time Updates 448
      12.3.3 Measurement Updates 450
      12.3.4 Predicted Estimators 451
      12.3.5 Filtered Estimators 452
      12.3.6 Estimator Update 453
      12.3.7 Operation Counts and Condensed Forms 454
   12.4 Numerical Examples 454
      12.4.1 Triangularization via Givens Rotations 455
12.4.2 Triangularization via Householder Transformations 456
12.4.3 Triangularization via Square-Root Free Rotations 458
12.5 Derivations of the Array Algorithms 459
12.5.1 The Time-Update Algorithm 459
12.5.2 The Measurement-Update Algorithm 460
12.5.3 Algorithm for the State Predictors 460
12.6 A Geometric Explanation of the Arrays 462
12.6.1 Predicted Form of the Arrays 463
12.6.2 Measurement Updates 465
12.6.3 Time Updates 466
12.7 Paige's Form of the Array Algorithm 466
12.8 Array Algorithms for the Information Forms 468
12.8.1 Information Array for the Measurement Update 468
12.8.2 Information Array for the Time Update 469
12.8.3 Alternative Derivation via Inversion of Covariance Forms 470
12.8.4 Derivation via Dualities When $R_i > 0$ and $Q_i > 0$ 470
12.8.5 The General Information Filter Form 472
12.8.6 A Geometric Derivation of the Information Filter Form 474
12.9 Array Algorithms for Smoothing 475
12.9.1 Bryson-Frazier Formulas in Array Form 475
12.9.2 Rauch-Tung-Striebel Formulas in Array Form 476
12.9.3 Two-Filter (or Mayne-Fraser) Array Formulas 477
12.10 Complements 478
12.11 Problems 479
12.A The UD Algorithm 485
12.B The Use of Schur and Condensed Forms 487
12.C Paige's Array Algorithm 488

13 FAST ARRAY ALGORITHMS 495
13.1 A Special Case: $P_0 = 0$ 495
13.1.1 Unitary Equivalence and an Alternative Derivation 496
13.2 A General Fast Array Algorithm 498
13.3 From Explicit Equations to Array Algorithms 500
13.4 Structured Time-Variant Systems 502
13.5 Complements 504
13.6 Problems 504
13.A Combining Displacement and State-Space Structures 509
14 ASYMPTOTIC BEHAVIOR

14.1 Introduction 512
  14.1.1 Time-Invariant State-Space Models 512
  14.1.2 Convergence for Indefinite Initial Conditions 513
  14.1.3 Convergence for Unstable $F$ 515
  14.1.4 Why Study Models with Unstable $F$? 515

14.2 Solutions of the DARE 519
14.3 Summary of Results 522
14.4 Riccati Solutions for Different Initial Conditions 526
14.5 Convergence Results 528
  14.5.1 A Sufficiency Result 528
  14.5.2 Simplified Convergence Conditions 541
  14.5.3 The Dual DARE and Stabilizability 546
14.6 The Case of Stable Systems 549
14.7 The Case of $S \neq 0$ 557
14.8 Exponential Convergence of the Fast Recursions 559
14.9 Complements 562
14.10 Problems 563

15 DUALITY AND EQUIVALENCE IN ESTIMATION AND CONTROL

15.1 Dual Bases 574
  15.1.1 Algebraic Specification 575
  15.1.2 Geometric Specification 575
  15.1.3 Some Reasons for Introducing Dual Bases 576
  15.1.4 Estimators via the Dual Basis 577
15.2 Application to Linear Models 578
  15.2.1 Linear Models and Dual Bases 579
  15.2.2 Application to the Measurement Update Problem 580
  15.2.3 Application to State-Space Models 581
15.3 Duality and Equivalence Relationships 583
  15.3.1 Equivalent Stochastic and Deterministic Problems 583
  15.3.2 Dual Stochastic and Deterministic Problems 584
  15.3.3 Summary of Duality and Equivalence Results 586
  15.3.4 A Deterministic Optimization Problem via Duality 589
  15.3.5 Application to Linear Quadratic Tracking 593
  15.3.6 Application to Linear Quadratic Regulation 596
15.4 Duality under Causality Constraints 597
15.4.1 Causal Estimation 597
15.4.2 Anticausal Dual Problem 598
15.4.3 Anticausal Estimation and Causal Duality 599
15.4.4 Application to Stochastic Quadratic Control 601
15.5 Measurement Constraints and a Separation Principle 606
15.5.1 A Separation Principle with Causal Dependence on Data 606
15.5.2 A Separation Principle with Anticausal Dependence on Data 611
15.5.3 Application to Measurement Feedback Control 612
15.6 Duality in the Frequency Domain 614
15.6.1 Duality without Constraints 614
15.6.2 Duality with Causality Constraints 616
15.6.3 Application to the Infinite-Horizon LQR Problem 618
15.7 Complementary State-Space Models 619
15.7.1 The Standard State-Space Model 619
15.7.2 Backwards Complementary Models 620
15.7.3 Direct Derivation of the Hamiltonian Equations 624
15.7.4 Forwards Complementary Models 625
15.7.5 The Mixed Complementary Model 629
15.7.6 An Application to Smoothing 629
15.8 Complements 631
15.9 Problems 632

16 CONTINUOUS-TIME STATE-SPACE ESTIMATION 638
16.1 Continuous-Time Models 638
16.1.1 Standard Continuous-Time Models 639
16.1.2 Discrete-Time Approximations 639
16.1.3 An Application: State-Variance Recursions 642
16.2 The Kalman Filter Equations Given State-Space and Covariance Models 643
16.3 Some Examples 649
16.4 Direct Solution Using the Innovations Process 651
16.4.1 The Innovations Process 652
16.4.2 The Innovations Approach 655
16.5 Smoothed Estimators 658
16.6 Fast Algorithms for Time-Invariant Models 662
16.7 Asymptotic Behavior 665
16.7.1 Positive-Semi-Definite Solutions of the CARE 665
16.7.2 Convergence Results 666
16.7.3 The Dual CARE 670
16.7.4 Exponential Convergence of the Fast Filtering Equations 670
16.8 The Steady-State Filter 671
16.9 Complements 672
16.10 Problems 679
16.A Backwards Markovian Models 694
16.A.1 Backwards Models via Time Reversal 694
16.A.2 The Backwards-Time Kalman Filters 696
16.A.3 Application to Smoothing Problems 696

17 A SCATTERING THEORY APPROACH 699
17.1 A Generalized Transmission–Line Model 700
17.1.1 Identifying the Macroscopic Scattering Operators 703
17.1.2 Identifying the Signals 705
17.2 Backward Evolution 706
17.3 The Star Product 709
17.3.1 Evolution Equations 711
17.3.2 General Initial Conditions 712
17.3.3 Chain Scattering or Transmission Matrices 714
17.4 Various Riccati Formulas 716
17.4.1 Incorporating Boundary Conditions 716
17.4.2 Partitioned Formulas 718
17.4.3 General Changes in the Boundary Conditions 718
17.4.4 Smoothing as an Extended Filtering Problem 721
17.5 Homogeneous Media: Time-Invariant Models 725
17.5.1 A Doubling Algorithm 725
17.5.2 Generalized Stokes Identities 726
17.6 Discrete-Time Scattering Formulation 729
17.6.1 Some Features of Discrete-Time Scattering 731
17.6.2 The Scattering Parameters 732
17.6.3 The Kalman Filter and Related Identities 735
17.6.4 General Change of Initial Conditions 737
17.6.5 Backward Evolution 738
17.6.6 Homogeneous Media 739
17.7 Further Work 742
17.8 Complements 742
17.9 Problems 743
17. A Complementary State-Space Model 746

A USEFUL MATRIX RESULTS 748
A.1 Some Matrix Identities 748
A.2 Kronecker Products 754
A.3 The Reduced and Full QR Decompositions 755
A.4 The Singular Value Decomposition and Applications 756
A.5 Basis Rotations 760
A.6 Complex Gradients and Hessians 763
A.7 Further Reading 764

B UNITARY AND J-UNITARY TRANSFORMATIONS 765
B.1 Householder Transformations 765
B.2 Circular or Givens Rotations 769
B.3 Fast Givens Transformations 771
B.4 J-Unitary Householder Transformations 774
B.5 Hyperbolic Givens Rotations 775
B.6 Some Alternative Implementations 777

C SOME SYSTEM THEORY CONCEPTS 780
C.1 Linear State-Space Models 780
C.2 State-Transition Matrices 780
C.3 Controllability and Stabilizability 783
C.4 Observability and Detectability 785
C.5 Minimal Realizations 786

D LYAPUNOV EQUATIONS 787
D.1 Discrete-Time Lyapunov Equations 787
D.2 Continuous-Time Lyapunov Equations 789
D.3 Internal Stability 791

E ALGEBRAIC RICCATI EQUATIONS 794
E.1 Overview of DARE 794
E.2 A Linear Matrix Inequality 798
E.3 Existence of Solutions to the DARE 799
E.4 Properties of the Maximal Solution 801
E.5 Main Result 804
E.6 Further Remarks 805
E.7 The Invariant Subspace Method 808
E.8  The Dual DARE  817
E.9  The CARE  819
E.10 Complements  826

F  DISPLACEMENT STRUCTURE  827
  F.1  Motivation  827
  F.2  Two Fundamental Properties  829
  F.3  A Generalized Schur Algorithm  831
  F.4  The Classical Schur Algorithm  833
  F.5  Combining Displacement and State-Space Structures  835

REFERENCES  836