INTERNATIONAL EDITION

ADAPTIVE FILTER THEORY

FOURTH EDITION

SIMON HAYKIN

PRENTICE HALL INFORMATION AND SYSTEM SCIENCES SERIES

Thomas Kailath Series Editor

Contents

Preface x

Acknowledgments xv

Background and Preview 1

1. The Filtering Problem 1
2. Linear Optimum Filters 3
3. Adaptive Filters 4
4. Linear Filter Structures 6
5. Approaches to the Development of Linear Adaptive Filters 14
6. Adaptive Beamforming 18
7. Four Classes of Applications 22
8. Historical Notes 25

Chapter 1 Stochastic Processes and Models 35

1.1 Partial Characterization of a Discrete-Time Stochastic Process 35
1.2 Mean Ergodic Theorem 37
1.3 Correlation Matrix 39
1.4 Correlation Matrix of Sine Wave Plus Noise 43
1.5 Stochastic Models 45
1.6 Wold Decomposition 51
1.7 Asymptotic Stationarity of an Autoregressive Process 53
1.8 Yule–Walker Equations 55
1.9 Computer Experiment: Autoregressive Process of Order Two 57
1.10 Selecting the Model Order 65
1.11 Complex Gaussian Processes 67
1.12 Power Spectral Density 69
1.13 Properties of Power Spectral Density 71
1.14 Transmission of a Stationary Process Through a Linear Filter 73
1.15 Cramér Spectral Representation for a Stationary Process 76
1.16 Power Spectrum Estimation 78
1.17 Other Statistical Characteristics of a Stochastic Process 81
1.18 Polyspectra 82
1.19 Spectral-Correlation Density 85
1.20 Summary 88
Problems 89

Chapter 2 Wiener Filters 94

2.1 Linear Optimum Filtering: Statement of the Problem 94
2.2 Principle of Orthogonality 96

2.3 Minimum Mean-Square Error 100
2.4 Wiener–Hopf Equations 102
2.5 Error-Performance Surface 104
2.6 Multiple Linear Regression Model 108
2.7 Example 110
2.8 Linearly Constrained Minimum-Variance Filter 115
2.9 Generalized Sidelobe Cancellers 120
2.10 Summary 126
Problems 128

Chapter 3 Linear Prediction 136

3.1 Forward Linear Prediction 136
3.2 Backward Linear Prediction 142
3.3 Levinson–Durbin Algorithm 148
3.4 Properties of Prediction-Error Filters 156
3.5 Schur–Cohn Test 166
3.6 Autoregressive Modeling of a Stationary Stochastic Process 168
3.7 Cholesky Factorization 171
3.8 Lattice Predictors 174
3.9 All-Pole, All-Pass Lattice Filter 179
3.10 Joint-Process Estimation 181
3.11 Predictive Modeling of Speech 185
3.12 Summary 192
Problems 193

Chapter 4 Method of Steepest Descent 203

4.1 Basic Idea of the Steepest-Descent Algorithm 203
4.2 The Steepest-Descent Algorithm Applied to the Wiener Filter 204
4.3 Stability of the Steepest-Descent Algorithm 208
4.4 Example 213
4.5 The Steepest-Descent Algorithm as a Deterministic Search Method 225
4.6 Virtue and Limitation of the Steepest-Descent Algorithm 226
4.7 Summary 227
Problems 228

Chapter 5 Least-Mean-Square Adaptive Filters 231

5.1 Overview of the Structure and Operation of the Least-Mean-Square Algorithm 231
5.2 Least-Mean-Square Adaptation Algorithm 235
5.3 Applications 238
5.4 Statistical LMS Theory 258
5.5 Comparison of the LMS Algorithm with the Steepest-Descent Algorithm 278
5.6 Computer Experiment on Adaptive Prediction 279
5.7 Computer Experiment on Adaptive Equalization 285
5.8 Computer Experiment on a Minimum-Variance Distortionless-Response Beamformer 291
5.9 Directionality of Convergence of the LMS Algorithm for Nonwhite Inputs 293
5.10 Robustness of the LMS Filter: H^∞ Criterion 297
5.11 Upper Bounds on the Step-Size Parameter for Different Scenarios 306
5.12 Transfer Function Approach for Deterministic Inputs 307
5.13 Summary 311
Problems 312

Chapter 6 Normalized Least-Mean-Square Adaptive Filters 320

6.1 Normalized LMS Filter as the Solution to a Constrained Optimization Problem 320
6.2 Stability of the Normalized LMS Filter 324
6.3 Step-Size Control for Acoustic Echo Cancellation 327
6.4 Geometric Considerations Pertaining to the Convergence Process for Real-Valued Data 331
6.5 Affine Projection Adaptive Filters 334
6.6 Summary 340
Problems 341

Chapter 7 Frequency-Domain and Subband Adaptive Filters 344

7.1 Block-Adaptive Filters 345
7.2 Fast Block-LMS Algorithm 350
7.3 Unconstrained Frequency-Domain Adaptive Filters 355
7.4 Self-Orthogonalizing Adaptive Filters 356
7.5 Computer Experiment on Adaptive Equalization 367
7.6 Subband Adaptive Filters 372
7.7 Classification of Adaptive Filtering Algorithms 380
7.8 Summary 381
Problems 382

Chapter 8 Method of Least Squares 385

8.1 Statement of the Linear Least-Squares Estimation Problem 385
8.2 Data Windowing 388
8.3 Principle of Orthogonality Revisited 389
8.4 Minimum Sum of Error Squares 392
8.5 Normal Equations and Linear Least-Squares Filters 393
8.6 Time-Average Correlation Matrix Φ 396
8.7 Reformulation of the Normal Equations in Terms of Data Matrices 398
8.8 Properties of Least-Squares Estimates 402
8.9 MVDR Spectrum Estimation 406
8.10 Regularized MVDR Beamforming 409
8.11 Singular-Value Decomposition 414
8.12 Pseudoinverse 421
8.13 Interpretation of Singular Values and Singular Vectors 423
8.14 Minimum-Norm Solution to the Linear Least-Squares Problem 424
8.15 Normalized LMS Algorithm Viewed as the Minimum-Norm Solution to an Underdetermined Least-Squares Estimation Problem 427
8.16 Summary 429
Problems 430

Chapter 9 Recursive Least-Squares Adaptive Filters 436

9.1 Some Preliminaries 436
9.2 The Matrix Inversion Lemma 440
9.3 The Exponentially Weighted Recursive Least-Squares Algorithm 440
9.4 Selection of the Regularizing Parameter 444
9.5 Update Recursion For the Sum of Weighted Error Squares 446
9.6 Example: Single-Weight Adaptive Noise Canceller 447
9.7 Convergence Analysis of the RLS Algorithm 448
9.8 Computer Experiment on Adaptive Equalization 454
9.9 Robustness of RLS Filters 457
9.10 Summary 463
Problems 463

Chapter 10 Kalman Filters 466
10.1 Recursive Minimum Mean-Square Estimation for Scalar Random Variables 466
10.2 Statement of the Kalman Filtering Problem 470
10.3 The Innovations Process 472
10.4 Estimation of the State Using the Innovations Process 474
10.5 Filtering 479
10.6 Initial Conditions 483
10.7 Summary of the Kalman Filter 483
10.8 Kalman Filter as the Unifying Basis for RLS Filters 485
10.9 Variants of the Kalman Filter 491
10.10 The Extended Kalman Filter 496
10.11 Summary 501
Problems 501

Chapter 11 Square-Root Adaptive Filters 506
11.1 Square-Root Kalman Filters 506
11.2 Building Square-Root Adaptive Filters on their Kalman Filter Counterparts 512
11.3 QR–RLS Algorithm 513
11.4 Adaptive Beamforming 521
11.5 Inverse QR–RLS Algorithm 528
11.6 Summary 531
Problems 531

Chapter 12 Order-Recursive Adaptive Filters 535
12.1 Gradient-Adaptive Lattice Filter 536
12.2 Order-Recursive Adaptive Filters Using Least-Squares Estimation: An Overview 543
12.3 Adaptive Forward Linear Prediction 544
12.4 Adaptive Backward Linear Prediction 548
12.5 Conversion Factor 550
12.6 Least-Squares Lattice Predictor 553
12.7 Angle-Normalized Estimation Errors 563
12.8 First-Order State-Space Models for Lattice Filtering 565
12.9 QR-Decomposition-Based Least-Squares Lattice Filters 571
12.10 Fundamental Properties of the QRD–LSL Filter 579
12.11 Computer Experiment on Adaptive Equalization 581
12.12 Recursive Least-Squares Lattice Filters Using a Posteriori Estimation Errors 586
12.13 Recursive LSL Filters Using a Priori Estimation Errors with Error Feedback 592
12.14 Relation Between Recursive LSL and RLS filters 596
12.15 Summary 598
Problems 600

Chapter 13 Finite-Precision Effects 607
13.1 Quantization Errors 608
13.2 Least-Mean-Square Algorithm 610
13.3 Recursive Least-Squares Algorithm 619
13.4 Square-Root Adaptive Filters 625
13.5 Order-Recursive Adaptive Filters 627
13.6 Fast Transversal Filters 629
13.7 Summary 633
Problems 635

Chapter 14 Tracking of Time-Varying Systems 637
14.1 Markov Model for System Identification 637
14.2 Degree of Nonstationarity 640

14.3 Criteria for Tracking Assessment 642
14.4 Tracking Performance of the LMS Algorithm 643
14.5 Tracking Performance of the RLS Algorithm 647
14.6 Comparison of the Tracking Performance of LMS and RLS Algorithms 651
14.7 How to Improve the Tracking Behavior of the RLS Algorithm 654
14.8 Computer Experiment on System Identification 657
14.9 Automatic Tuning of The Adaptation Constants 659
14.10 Summary 664
Problems 665

Chapter 15 Adaptive Filters Using Infinite-Duration Impulse Response Structures 666

15.1 IIR Adaptive Filters: Output Error Method 666
15.2 IIR Adaptive Filters: Equation Error Method 671
15.3 Some Practical Considerations 673
15.4 Laguerre Transversal Filters 674
15.5 Adaptive Laguerre Lattice Filters 677
15.6 Summary 681
Problems 682

Chapter 16 Blind Deconvolution 684

16.1 An Overview of the Blind Deconvolution Problem 684
16.2 Channel Identifiability Using Cyclostationary Statistics 688
16.3 Subspace Decomposition for Fractionally Spaced Blind Identification 689
16.4 Bussgang Algorithm for Blind Equalization 703
16.5 Extension of the Bussgang Algorithm to Complex Baseband Channels 720
16.6 Special Cases of the Bussgang Algorithm 721
16.7 Fractionally Spaced Bussgang Equalizers 725
16.8 Summary 729
Problems 732

Chapter 17 Back-Propagation Learning 736

17.1 Sigmoid Neuronal Model 736
17.2 Multilayer Perceptron 738
17.3 Complex Back-Propagation Algorithm 740
17.4 Universal Approximation Theorem 753
17.5 Network Complexity 755
17.6 Temporal Processing: How to Account for "Time" 756
17.7 Virtues and Limitations of Back-Propagation Learning 758
17.8 Summary 759
Problems 760

Epilogue 762

1. Proportionate Adaptation 762
2. Robust Statistics 764
3. Blind Source Separation 766
4. Recursive Neural Networks 770
5. Nonlinear Dynamical Systems: Derivative-Free State Estimation 773

Appendix A Complex Variables 779

A.1 Cauchy–Reimann Equations 779
A.2 Cauchy's Integral Formula 781
A.3 Laurent's Series 783

A.4 Singularities and Residues 785
A.5 Cauchy's Residue Theorem 786
A.6 Principle of the Argument 787
A.7 Inversion Integral for the z-Transform 790
A.8 Parseval's Theorem 792

Appendix B Differentiation with Respect to a Vector 794
B.1 Basic Definitions 794
B.2 Examples 796
B.3 Relation Between the Derivative with Respect to a Vector and the Gradient Vector 798

Appendix C Method of Lagrange Multipliers 799
C.1 Optimization Involving a Single Equality Constraint 799
C.2 Optimization Involving Multiple Equality Constraints 800
C.3 Optimum Beamformer 801

Appendix D Estimation Theory 802
D.1 Likelihood Function 802
D.2 Cramér–Rao Inequality 803
D.3 Properties of Maximum-Likelihood Estimators 804
D.4 Conditional Mean Estimator 805

Appendix E Eigenanalysis 807
E.1 The Eigenvalue Problem 807
E.2 Properties of Eigenvalues and Eigenvectors 809
E.3 Low-Rank Modeling 823
E.4 Eigenfilters 827
E.5 Eigenvalue Computations 829

Appendix F Rotations and Reflections 833
F.1 Plane Rotations 833
F.2 Two-Sided Jacobi Algorithm 835
F.3 Cyclic Jacobi Algorithm 841
F.4 Householder Transformation 844
F.5 The QR Algorithm 847

Appendix G Complex Wishart Distribution 854
G.1 Definition 854
G.2 The Chi-Square Distribution as a Special Case 855
G.3 Properties of the Complex Wishart Distribution 856
G.4 Expectation of the Inverse Correlation Matrix $\Phi^{-1}(n)$ 857

Glossary 858

Bibliography 870

Index 912