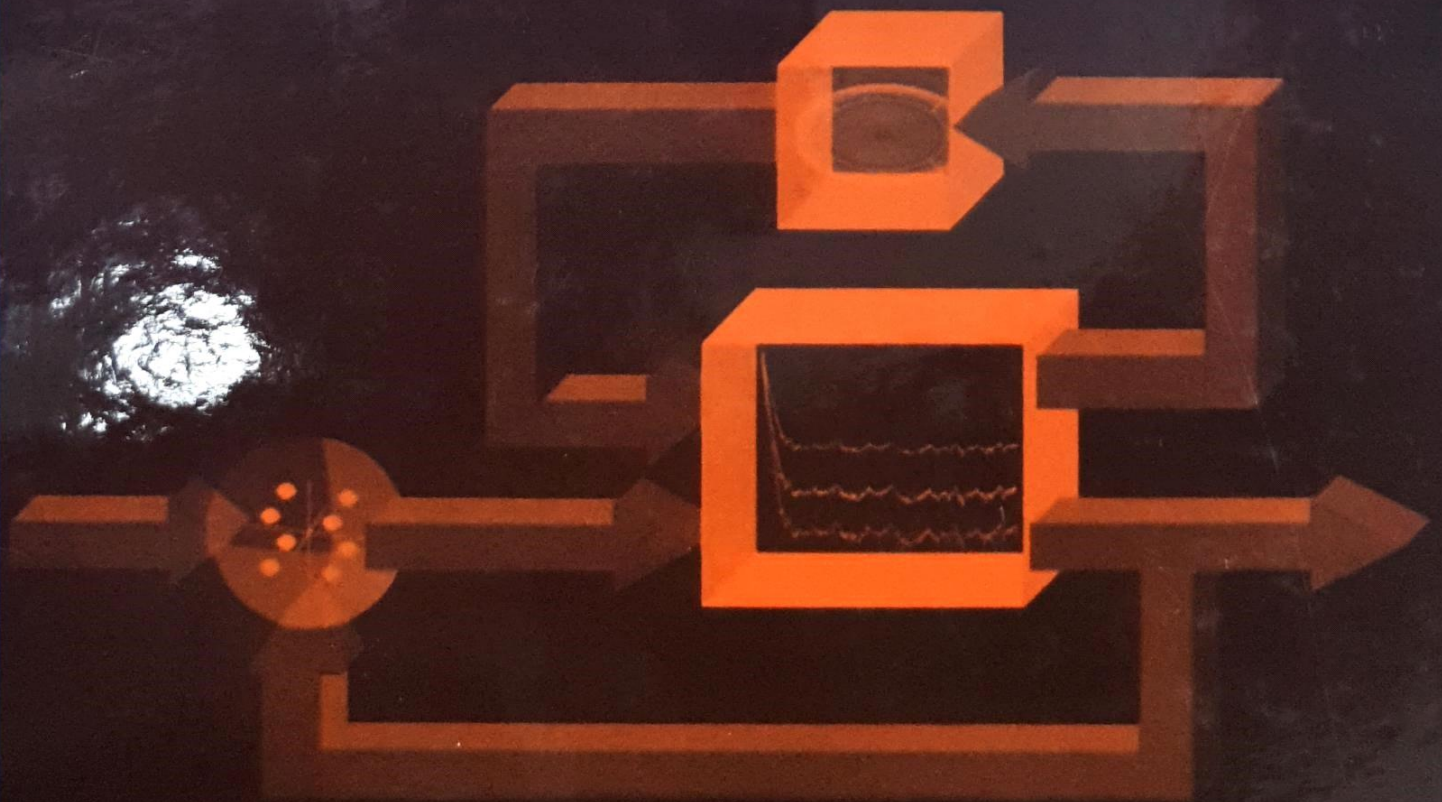


Fundamentals ^{of} Adaptive Filtering

ALI H. SAYED

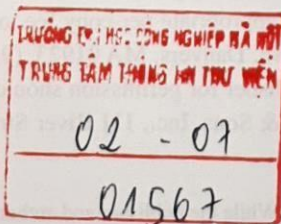


Fundamentals of Adaptive Filtering

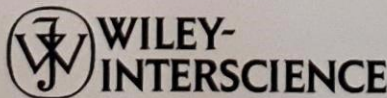
To Laila, Faten, and Samiya

Ali H. Sayed

University of California, Los Angeles



IEEE PRESS



A JOHN WILEY & SONS, INC., PUBLICATION

CONTENTS

PREFACE	xix
ACKNOWLEDGMENTS	xxix
NOTATION	xxxii
SYMBOLS	xxxv
1 OPTIMAL ESTIMATION	1
1.1 Variance of a Random Variable	1
1.2 Estimation Given No Observations	5
1.3 Estimation Given Dependent Observations	6
1.3.1 Mean-Square-Error Criterion	8
1.3.2 Orthogonality Principle	12
1.3.3 Gaussian Random Variables	15
1.4 Estimation in the Complex and Vector Cases	18
1.4.1 Complex-Valued Random Variables	18
1.4.2 Vector-Valued Random Variables	20
1.4.3 Optimal Estimator in the Vector Case	22
1.4.4 Equivalent Optimization Criterion	26
1.4.5 Spherically Invariant Gaussian Variables	28
1.5 Summary of Main Results	30
1.6 Bibliographic Notes	31
1.7 Problems	33
1.8 Computer Project	37
1.A Hermitian and Positive-Definite Matrices	39
1.B Gaussian Random Vectors	42
2 LINEAR ESTIMATION	47
2.1 Normal Equations	48
2.1.1 Affine Estimation	48
2.1.2 Mean-Square-Error Criterion	49
2.1.3 Minimization by Differentiation	50

2.1.4	Minimization by Completion-of-Squares	51
2.1.5	Minimization of the Error-Covariance Matrix	52
2.1.6	Statement of Main Results	53
2.2	Design Examples	54
2.3	Existence of Solutions	60
2.4	Orthogonality Principle	63
2.5	Nonzero-Mean Variables	65
2.6	Linear Models	66
2.7	Applications	68
2.7.1	Channel Estimation	69
2.7.2	Block Data Estimation	70
2.7.3	Linear Channel Equalization	71
2.7.4	Multiple-Antenna Receiver	75
2.8	Summary of Main Results	76
2.9	Bibliographic Notes	77
2.10	Problems	79
2.11	Computer Project	95
2.A	Range Spaces and Nullspaces of Matrices	103
2.B	Complex Gradients	105
2.C	Kalman Filter	108
3	CONSTRAINED LINEAR ESTIMATION	114
3.1	Minimum-Variance Unbiased Estimation	115
3.1.1	Problem Formulation	116
3.1.2	Interpretation and Solution	116
3.1.3	Example: Mean Estimation	118
3.2	Application: Channel and Noise Estimation	119
3.3	Application: Decision Feedback Equalization	120
3.4	Application: Antenna Beamforming	128
3.5	Summary of Main Results	131
3.6	Bibliographic Notes	131
3.7	Problems	133
3.8	Two Computer Projects	143
3.A	Schur Complements	155
3.B	Primer on Channel Equalization	159
3.C	Causal Wiener-Hopf Filtering	167
4	STEEPEST-DESCENT ALGORITHMS	170
4.1	Linear Estimation Problem	171
4.2	Steepest-Descent Method	174
4.2.1	Choice of the Search Direction	175
4.2.2	Condition on the Step-Size for Convergence	176
4.2.3	More General Cost Functions	178
4.3	Transient Behavior	179

4.3.1	Modes of Convergence	179
4.3.2	Time Constants	183
4.3.3	Learning Curve	184
4.3.4	Contour Curves of the Error Surface	186
4.4	Iteration-Dependent Step-Sizes	187
4.4.1	Condition for Convergence	187
4.4.2	Optimal Iteration-Dependent Step-Size	189
4.5	Newton's Method	191
4.5.1	Convergence Properties	192
4.5.2	Learning Curve	192
4.6	Summary of Main Results	193
4.7	Bibliographic Notes	194
4.8	Problems	196
4.9	Two Computer Projects	204
5	STOCHASTIC-GRADIENT ALGORITHMS	212
5.1	Motivation	213
5.2	LMS Algorithm	214
5.2.1	Instantaneous Approximations	214
5.2.2	Computational Cost	215
5.2.3	Least-Perturbation Property of LMS	216
5.3	Application: Adaptive Channel Estimation	218
5.4	Application: Adaptive Channel Equalization	220
5.5	Application: Decision-Feedback Equalization	223
5.6	Normalized LMS Algorithm	225
5.6.1	Instantaneous Approximations	225
5.6.2	Variations and Cost	226
5.6.3	Least-Perturbation Property of ϵ -NLMS	229
5.7	Other LMS-type Algorithms	233
5.7.1	Non-Blind Algorithms	233
5.7.2	Blind Algorithms	235
5.7.3	Some Properties	236
5.8	Affine Projection Algorithms	238
5.8.1	Instantaneous Approximations	238
5.8.2	Least-Perturbation Property of ϵ -APA	240
5.8.3	Affine Projection Interpretation	241
5.8.4	Variations	242
5.9	RLS Algorithm	245
5.10	Ensemble-Average Learning Curves	248
5.11	Summary of Main Results	251
5.12	Bibliographic Notes	252
5.13	Problems	256
5.14	Three Computer Projects	267

6	STEADY-STATE PERFORMANCE OF ADAPTIVE FILTERS	281
6.1	Performance Measure	282
6.1.1	Stochastic Equations	283
6.1.2	Excess Mean-Square Error and Misadjustment	283
6.2	Stationary Data Model	284
6.2.1	Linear Regression Model	284
6.2.2	Useful Independence Results	285
6.2.3	Alternative Expression for the EMSE	286
6.2.4	Error Quantities	286
6.3	Fundamental Energy-Conservation Relation	287
6.3.1	Algebraic Derivation	288
6.4	Fundamental Variance Relation	290
6.4.1	Steady-State Filter Operation	290
6.4.2	Variance Relation for Steady-State Performance	291
6.4.3	Relevance to Mean-Square Performance Analysis	292
6.5	Mean-Square Performance of LMS	292
6.5.1	Sufficiently Small Step-Sizes	293
6.5.2	Separation Principle	293
6.5.3	White Gaussian Input	294
6.5.4	Statement of Results	298
6.5.5	Simulation Results	298
6.6	Mean-Square Performance of ϵ -NLMS	300
6.6.1	Separation Principle	300
6.6.2	Simulation Results for ϵ -NLMS	302
6.6.3	ϵ -NLMS with Power Normalization	302
6.7	Mean-Square Performance of Sign-Error LMS	305
6.7.1	Real-Valued Data	306
6.7.2	Complex-Valued Data	307
6.7.3	Simulation Results	308
6.8	Mean-Square Performance of LMF and LMMN	308
6.8.1	Real-Valued Data	310
6.8.2	Complex-Valued Data	312
6.8.3	Simulation Results	314
6.9	Mean-Square Performance of RLS	317
6.9.1	Energy-Conservation Relation	317
6.9.2	Steady-State Approximation	318
6.9.3	Filter Performance	319
6.9.4	Simulation Results	321
6.10	Mean-Square Performance of ϵ -APA	322
6.10.1	Energy Conservation Relation	322
6.10.2	Simulation Results	325
6.11	Mean-Square Performance of Other Filters	325
6.12	Performance Table for Small Step-Sizes	327

6.13	Summary of Main Results	327
6.14	Bibliographic Notes	329
6.15	Problems	332
6.16	Computer Project	343
6.A	Interpretations of the Energy Relation	348
6.B	Relating ϵ -NLMS to LMS	353
6.C	Affine Projection Performance Condition	355
7	TRACKING PERFORMANCE OF ADAPTIVE FILTERS	357
7.1	Motivation	357
7.2	Nonstationary Data Model	358
7.2.1	Random-Walk Model	359
7.2.2	Useful Independence Results	361
7.2.3	Alternative Expression for the EMSE	362
7.2.4	Degree of Nonstationarity	362
7.2.5	Error Quantities	363
7.3	Fundamental Energy-Conservation Relation	364
7.4	Fundamental Variance Relation	364
7.4.1	Variance Relation for Steady-State Performance	365
7.4.2	Relevance to Tracking Analysis	366
7.5	Tracking Performance of LMS	367
7.5.1	Small Step-Sizes	367
7.5.2	Separation Principle	368
7.5.3	White Gaussian Input	368
7.6	Tracking Performance of ϵ -NLMS	370
7.6.1	Separation Principle	370
7.6.2	ϵ -NLMS with Power Normalization	371
7.7	Tracking Performance of Sign-Error LMS	372
7.7.1	Real-Valued Data	372
7.7.2	Complex-Valued Data	373
7.8	Tracking Performance of LMF and LMMN	374
7.8.1	Real-Valued Data	374
7.8.2	Complex-Valued Data	376
7.9	Comparison of Tracking Performance	378
7.10	Tracking Performance of RLS	380
7.11	Tracking Performance of ϵ -APA	384
7.12	Tracking Performance of Other Filters	386
7.13	Performance Table for Small Step-Sizes	387
7.14	Summary of Main Results	387
7.15	Bibliographic Notes	389
7.16	Problems	391
7.17	Computer Project	401
8	FINITE PRECISION EFFECTS	408

8.1	Quantization Model	409
8.2	Data Model and Quantization Error Sources	410
8.3	Fundamental Energy-Conservation Relation	413
8.3.1	Finite-Precision Implementation	413
8.3.2	Energy Conservation	415
8.4	Fundamental Variance Relation	416
8.5	Performance Degradation of LMS	419
8.6	Performance Degradation of ϵ -NLMS	421
8.7	Performance Degradation of Sign-Error LMS	423
8.8	Performance Degradation of LMF and LMMN	424
8.9	Performance Degradation of Other Filters	425
8.10	Summary of Main Results	426
8.11	Bibliographic Notes	428
8.12	Problems	430
8.13	Computer Project	437
9	TRANSIENT PERFORMANCE OF ADAPTIVE FILTERS	441
9.1	Data Model	442
9.2	Data-Normalized Adaptive Filters	442
9.3	Weighted Energy-Conservation Relation	443
9.4	Weighted Variance Relation	445
9.4.1	Variance Relation	446
9.4.2	Independence Assumption	448
9.4.3	Convenient Change of Coordinates	450
9.5	Transient Performance of LMS	452
9.5.1	Gaussian Regressors	452
9.5.2	Non-Gaussian Regressors	465
9.6	Transient Performance of ϵ -NLMS	471
9.7	Performance of Data-Normalized Filters	474
9.8	Summary of Main Results	477
9.9	Bibliographic Notes	481
9.10	Problems	487
9.11	Computer Project	516
9.A	Stability Bound	522
9.B	Stability of ϵ -NLMS	524
9.C	Adaptive Filters with Error Nonlinearities	526
9.D	Convergence Time of Adaptive Filters	538
9.E	Learning Behavior of Adaptive Filters	545
9.F	Independence and Averaging Analysis	559
9.G	Interpretation of Weighted Energy Relation	568
9.H	Kronecker Products	570
10	BLOCK ADAPTIVE FILTERS	572
10.1	Transform-Domain Adaptive Filters	573

10.1.1	Pre-Whitening Filters	573
10.1.2	Unitary Transformations	577
10.1.3	DFT-Domain LMS	581
10.1.4	DCT-Domain LMS	582
10.2	Motivation for Block Adaptive Filters	584
10.3	Efficient Block Convolution	586
10.3.1	Block Data Formulation	586
10.3.2	Block Convolution	589
10.4	DFT-Based Block Adaptive Filters	597
10.4.1	Unconstrained Filter Implementation	599
10.4.2	Constrained Filter Implementations	601
10.4.3	Computational Complexity	603
10.5	Subband Adaptive Filters	605
10.5.1	Analysis Filter Bank	607
10.5.2	Synthesis Filter Bank	609
10.5.3	Structures for Subband Filtering	611
10.6	Summary of Main Results	612
10.7	Bibliographic Notes	614
10.8	Problems	616
10.9	Computer Project	620
10.A	DCT-Transformed Regressors	626
10.B	More Constrained DFT Block Filters	628
10.C	Overlap-Add DFT-Based Block Adaptive Filter	632
10.D	DCT-Based Block Adaptive Filters	640
10.E	DHT-Based Block Adaptive Filters	648
11	THE LEAST-SQUARES CRITERION	657
11.1	Least-Squares Problem	658
11.1.1	Vector Formulation	659
11.1.2	Geometric Argument	660
11.1.3	Differentiation Argument	661
11.1.4	Completion-of-Squares Argument	661
11.1.5	Statement of Solution	663
11.1.6	Projection Matrices	665
11.2	Weighted Least-Squares	666
11.3	Regularized Least-Squares	669
11.4	Weighted Regularized Least-Squares	671
11.5	Order-Update Relations	672
11.5.1	Backward Projection	674
11.5.2	Forward Projection	684
11.6	Summary of Main Results	688
11.7	Bibliographic Notes	689
11.8	Problems	693

11.9 Three Computer Projects	703
11.A Equivalence Results in Linear Estimation	724
11.B QR Decomposition	726
11.C Singular Value Decomposition	728
12 RECURSIVE LEAST-SQUARES	732
12.1 Motivation	732
12.2 RLS Algorithm	733
12.2.1 Derivation	734
12.2.2 Regularization	736
12.2.3 Conversion Factor	737
12.2.4 Time-Update of the Minimum Cost	737
12.3 Exponentially-Weighted RLS Algorithm	739
12.3.1 Derivation	739
12.4 General Time-Update Result	741
12.5 Summary of Main Results	745
12.6 Bibliographic Notes	745
12.7 Problems	748
12.8 Two Computer Projects	755
12.A Kalman Filtering and Recursive Least-Squares	763
12.B Extended RLS Algorithms	768
13 RLS ARRAY ALGORITHMS	775
13.1 Some Difficulties	775
13.2 Square-Root Factors	776
13.2.1 Definition	777
13.2.2 Cholesky Factor	778
13.2.3 Array Algorithms	778
13.3 Norm and Angle Preservation	778
13.4 Motivation for Array Methods	780
13.5 RLS Algorithm	784
13.6 Inverse QR Algorithm	785
13.7 QR Algorithm	788
13.8 Extended QR Algorithm	793
13.9 Summary of Main Results	794
13.10 Bibliographic Notes	795
13.11 Problems	797
13.12 Computer Project	802
13.A Unitary Transformations	804
13.A.1 Givens Rotations	804
13.A.2 Householder Transformations	808
13.B Array Algorithms for Kalman Filtering	812
14 FAST FIXED-ORDER FILTERS	816

14.1	Fast Array Algorithm	817
14.1.1	Time-Update for the Gain Vector	818
14.1.2	Time-Update for the Conversion Factor	819
14.1.3	Initial Conditions	819
14.1.4	Array Algorithm	821
14.2	Regularized Prediction Problems	825
14.2.1	Regularized Backward Prediction	827
14.2.2	Regularized Forward Prediction	829
14.2.3	Low-Rank Factorization	831
14.3	Fast Transversal Filter	832
14.4	FAEST Filter	836
14.5	Fast Kalman Filter	838
14.6	Stability Issues	839
14.6.1	Array Implementation	840
14.6.2	Rescue Mechanisms	841
14.6.3	Feedback Stabilization	842
14.6.4	Incorporating Leakage	844
14.7	Summary of Main Results	845
14.8	Bibliographic Notes	846
14.9	Problems	848
14.10	Computer Project	857
14.A	Hyperbolic Rotations	860
14.B	Hyperbolic Basis Rotations	867
14.C	Backward Consistency and Minimality	869
14.D	Chandrasekhar Filter	871
15	LATTICE FILTERS	874
15.1	Motivation and Notation	875
15.1.1	Notation for Order-Recursive Problems	875
15.1.2	Motivation for Lattice Filters	877
15.2	Joint Process Estimation	878
15.3	Backward Estimation Problem	880
15.4	Forward Estimation Problem	883
15.5	Time and Order-Update Relations	885
15.5.1	Order-Update of Estimation Errors	885
15.5.2	Time-Update Relations	888
15.6	Significance of Data Structure	891
15.6.1	Shift Structure	891
15.6.2	Interpretation of the Estimation Errors	893
15.7	A Posteriori-Based Lattice Filter	894
15.8	A Priori-Based Lattice Filter	895
15.9	A Priori Error-Feedback Lattice Filter	897
15.10	A Posteriori Error-Feedback Lattice Filter	902

15.11 Normalized Lattice Filter	904
15.11.1 Time-Update for the Normalized Reflection Coefficients	905
15.11.2 Order-Update for the Normalized Estimation Errors	907
15.11.3 Significance of Data Structure	909
15.12 Array-Based Lattice Filter	910
15.12.1 Order-Update of the Output Estimation Error	910
15.12.2 Order-Update of the Backward Estimation Error	912
15.12.3 Order-Update of the Forward Estimation Error	913
15.12.4 Significance of Data Structure	914
15.13 Relation Between RLS and Lattice Filters	915
15.14 Summary of Main Results	917
15.15 Bibliographic Notes	918
15.16 Problems	920
15.17 Computer Project	925
16 LAGUERRE ADAPTIVE FILTERS	931
16.1 Orthonormal Filter Structures	932
16.2 Data Structure	934
16.3 Fast Array Algorithm	936
16.4 Regularized Projection Problems	942
16.5 Extended Fast Transversal Filter	954
16.6 Extended FAEST Filter	957
16.7 Extended Fast Kalman Filter	958
16.8 Stability Issues	959
16.9 Order-Recursive Filters	960
16.10 A Posteriori-Based Lattice Filter	968
16.11 A Priori-Based Lattice Filter	970
16.12 A Priori Error-Feedback Lattice Filter	972
16.13 A Posteriori Error-Feedback Lattice Filter	976
16.14 Normalized Lattice Filter	978
16.15 Array Lattice Filter	982
16.16 Summary of Main Results	985
16.17 Bibliographic Notes	986
16.18 Problems	989
16.19 Computer Project	994
16.A Modeling with Orthonormal Basis Functions	999
16.B Efficient Matrix-Vector Multiplication	1007
16.C Lyapunov Equations	1009
17 ROBUST ADAPTIVE FILTERS	1012
17.1 Indefinite Least-Squares	1013
17.2 Recursive Minimization Algorithm	1018
17.2.1 Derivation of the Algorithm	1019
17.2.2 Time-Update of the Minimum Cost	1022

17.2.3	Minimization Conditions	1024
17.2.4	Singular Weighting Matrices	1026
17.3	A Posteriori-Based Robust Filters	1027
17.3.1	Robustness Criterion	1028
17.3.2	Relation to Indefinite Least-Squares	1029
17.3.3	Solving the Minimization Step	1030
17.3.4	Enforcing Positivity	1030
17.3.5	Statement of the Robust Solution	1032
17.3.6	Gauss-Newton Algorithm	1033
17.3.7	ϵ -NLMS Algorithm	1035
17.4	A Priori-Based Robust Filters	1036
17.4.1	Gauss-Newton Algorithm	1040
17.4.2	LMS Algorithm	1041
17.5	Energy Conservation Arguments	1043
17.5.1	Robustness of LMS	1043
17.5.2	Cauchy-Schwartz Interpretation	1045
17.5.3	Robustness Interpretation	1046
17.5.4	Robustness of ϵ -NLMS	1047
17.5.5	Time-Dependent Step-Sizes	1048
17.5.6	Robustness of Gauss-Newton	1049
17.5.7	Robustness of RLS	1050
17.6	Summary of Main Results	1052
17.7	Bibliographic Notes	1052
17.8	Problems	1056
17.9	Computer Project	1072
17.A	Arbitrary Coefficient Matrices	1078
17.B	Total-Least-Squares	1081
17.C	\mathcal{H}^∞ Filters	1085
17.D	Stationary Points	1089
BIBLIOGRAPHY		1090
AUTHOR INDEX		1113
SUBJECT INDEX		1118

PREFACE

Adaptive filtering is a topic of immense practical relevance and deep theoretical challenges that persist even to this date. There are several notable texts on the subject that describe many of the features that have marvelled students and researchers over the years. In this textbook, we choose to step back and to take a broad look at the field. In so doing, we feel that we are able to bring forth, to the benefit of the reader, extensive commonalities that exist among different classes of adaptive algorithms and even among different filtering theories. We are also able to provide a uniform treatment of the subject in a manner that addresses some existing limitations, provides additional insights, and allows for extensions of current theory.

We do not have any illusions about the difficulties that arise in any attempt at understanding adaptive filters more fully. This is because adaptive filters are, by design, time-variant, nonlinear, and stochastic systems. Anyone of these qualifications alone would have resulted in a formidable system to study. Put them together and you face an almost impossible task. It is no wonder then that current practice tends to study different adaptive schemes separately, with techniques and assumptions that are usually more suitable for one adaptation form over another. It is also no surprise that most treatments of adaptive filters, including the one adopted in this textbook, need to rely on some simplifying assumptions in order to make filter analysis and design a more tractable objective.

Still, in our view, three desirable features of any study of adaptive filters would be (1) to attempt to keep the number of simplifying assumptions to a minimum, (2) to delay their use until necessary, and (3) to apply similar assumptions uniformly across different classes of adaptive algorithms. This last feature enables us to evaluate and compare the performance of adaptive schemes under similar assumptions on the data, while delaying the use of assumptions enables us to extract the most information possible about actual filter performance. In our discussions in this book we pay particular attention to these three features throughout the presentation.

In addition, we share the conviction that a thorough understanding of the performance and limitations of adaptive filters requires a solid grasp of the fundamentals of least-mean-squares estimation theory. These fundamentals help the designer understand what it is that an adaptive filter is trying to accomplish and how well it performs in this regard. For this reason, the first three chapters of the book are designed to provide the reader with a self-contained and easy-to-follow exposition of estimation theory, with a focus on topics that are relevant to the subject matter of the book. In these initial chapters, special emphasis is placed on geometric interpretations of several fundamental results. The reader is advised to pay close attention to these interpretations since it will become clear, time and again, that cumbersome algebraic manipulations can often be simplified by recourse to geometric constructions. These constructions not only provide a more lasting appreciation for the results of the book, but they also expose the reader to powerful tools that can be useful in other contexts as well, other than adaptive filtering and estimation theory.

The reader is further advised to master the convenience of the vector notation, which is used extensively throughout this book. Besides allowing a compact exposition of ideas and a compact representation of results, the vector notation also allows us to exploit to great effect several important results from linear algebra and matrix theory and to capture, in elegant ways, many revealing characteristics of adaptive filters. We cannot emphasize strongly enough the importance of linear algebraic and matrix tools in our presentation, as well as the elegance that they bring to the subject. The combined power of the geometric point of view and the vector notation are perhaps best exemplified by our detailed treatment later in this book of least-squares theory and its algorithmic variants. Of course, the reader is exposed to geometric and vector formulations in the early chapters of the book already, including the first chapter.

Style of the Book

Each chapter in the book consists generally of five distinctive parts in the following order: i) concepts, ii) bibliographic notes, iii) problems, iv) computer projects, and v) appendices.

- i) **Concepts.** In the early chapters, each concept is motivated from first principles; starting from the obvious and ending with the more advanced. We follow this route of presentation until the reader develops enough maturity in the field. As the book progresses, we expect the reader to become more sophisticated and, therefore, we cut back on the “obvious”. While for some advanced readers and researchers the “obvious” part in the initial chapters might seem at first unnecessary, please keep in mind that the primary readers of any textbook are novices to the field. From our experience over the years, teaching from early drafts of this manuscript, students have been particularly receptive to this line of presentation. In addition, for ease of reference, we have collected at the end of each chapter a summary of the key concepts and results.
- ii) **Bibliographic Notes.** In the remarks at the end of each chapter we provide a wealth of references on the main contributors to the results discussed in the text. Rather than scatter references throughout the chapter, we find it useful to collect all references at the concluding section of each chapter in the form of a narrative. We believe that this way of presentation gives the reader a more focused perspective on how the references and the contributions relate to each other both in time and context.
- iii) **Problems.** The book contains a significant number of problems, some more challenging than others and some more applied than others. The problems should be viewed as an *integral* part of the text, especially since many additional and interesting results appear in them. It was for this reason, and also for the benefit of the reader, that we have chosen to formulate and design all problems in a guided manner. Usually, and especially in the more challenging cases, a problem starts by stating its objective followed by a sequence of guided steps until the final answer is attained. The answer to each step appears stated in the body of the problem. In this way, a reader would know what the answer should be, even if he fails to solve the problem. Thus rather than ask the reader to “find an expression for x ”, we would instead ask him to “verify that x is given by $x = \dots$ ” and then give the expression for x .

All instructors can request copies of a complete solutions manual from the publisher.

Moreover, several problems in the book have been designed to introduce readers to ideas of interest from related fields, such as multi-antenna receivers, cyclic-prefixing,

maximal ratio combining, OFDM receivers, CDMA receivers, and so forth. Students are usually surprised to learn how classical concepts and ideas form the underpinnings of seemingly advanced techniques.

- iv) **Computer Projects.** We have included several computer projects to show students, and also practitioners, how the results developed in the book can be useful in situations of practical interest (e.g., linear equalization, decision feedback equalization, channel estimation, beamforming, tracking fading channels, line echo cancellation, acoustic echo cancellation, active noise control, OFDM receivers, CDMA receivers, finite-precision implementations). In designing these projects, we have made an effort at choosing topics that are relevant to practitioners. We have also made an effort at illustrating to students how a solid theoretical understanding can guide them in challenging situations. All computer projects in the book are followed by extensive commentary and typical performance plots. Complete MATLAB¹ programs are available for solving all computer projects.

Detailed MATLAB programs that solve all computer projects in the book can be downloaded by all readers from the publisher's website:
ftp://ftp.wiley.com/public/sci_tech_med/filtering/

- v) **Appendices.** Rather than collect all appendices at the end of the book, we have opted to place each appendix at the end of the chapter where it is called upon. In this way, the usefulness of the material in an appendix, and its relation to the discussion in the chapter, would become more evident to the reader. For example, although most students would have had some exposure to linear algebra and matrix theory before a course on adaptive filtering, we provide a handful of self-contained appendices that explain all the required concepts for the purposes of this book (e.g., rank and range spaces of matrices, solutions of linear equations, Schur complements, singular value decomposition, Cholesky decomposition, etc). Since each appendix is placed right where the concepts are first needed, students will be able to appreciate firsthand the elegance that such concepts bring to the presentation. Actually, after progressing sufficiently enough in the book, students will be able to master many useful concepts from linear algebra and matrix theory, in addition to adaptive filtering.

Organization of the Book

The material in the book can be categorized into five broad areas, as listed in Table P.1. Area I covers the fundamentals of least-mean-squares estimation theory with several application examples. Areas II and III deal mainly with LMS-type adaptive filters, while areas IV and V deal with least-squares-type adaptive filters. If an instructor wishes to focus mostly on LMS-type filters, then he can do so by covering only material from within areas II and III. Even in this case, students will still be exposed to the recursive-least-squares (RLS) algorithm and its performance results from the discussions in Chapter 5 and Area III. However, for a more-in-depth treatment of RLS and its many variants, instructors will need to select chapters from within Area IV as well.

Dependencies among chapters. Figure P.1 illustrates the dependencies among the chapters in the book. In the figure, the material in a chapter that is at the receiving end of an arrow requires some (but not necessarily all) of the material from the chapter at the origin of the

¹MATLAB is a trademark of the MathWorks Inc.

Table P.1. A breakdown of the book chapters into general topic areas.

Category	Chapters
I. Introduction and Foundations	1. Optimal estimation. 2. Linear estimation. 3. Constrained linear estimation.
II. Stochastic-Gradient Methods	4. Steepest-descent algorithms. 5. Stochastic-gradient algorithms. 10. Block adaptive filters.
III. Performance Analyses	6. Steady-state performance of adaptive filters. 7. Tracking performance of adaptive filters. 8. Finite-precision effects. 9. Transient performance of adaptive filters.
IV. Least-Squares Methods	11. The least-squares criterion. 12. Recursive least-squares. 13. RLS array algorithms. 14. Fast fixed-order filters. 15. Lattice filters. 16. Laguerre adaptive filters.
V. Indefinite Least-Squares	17. Robust adaptive filters.

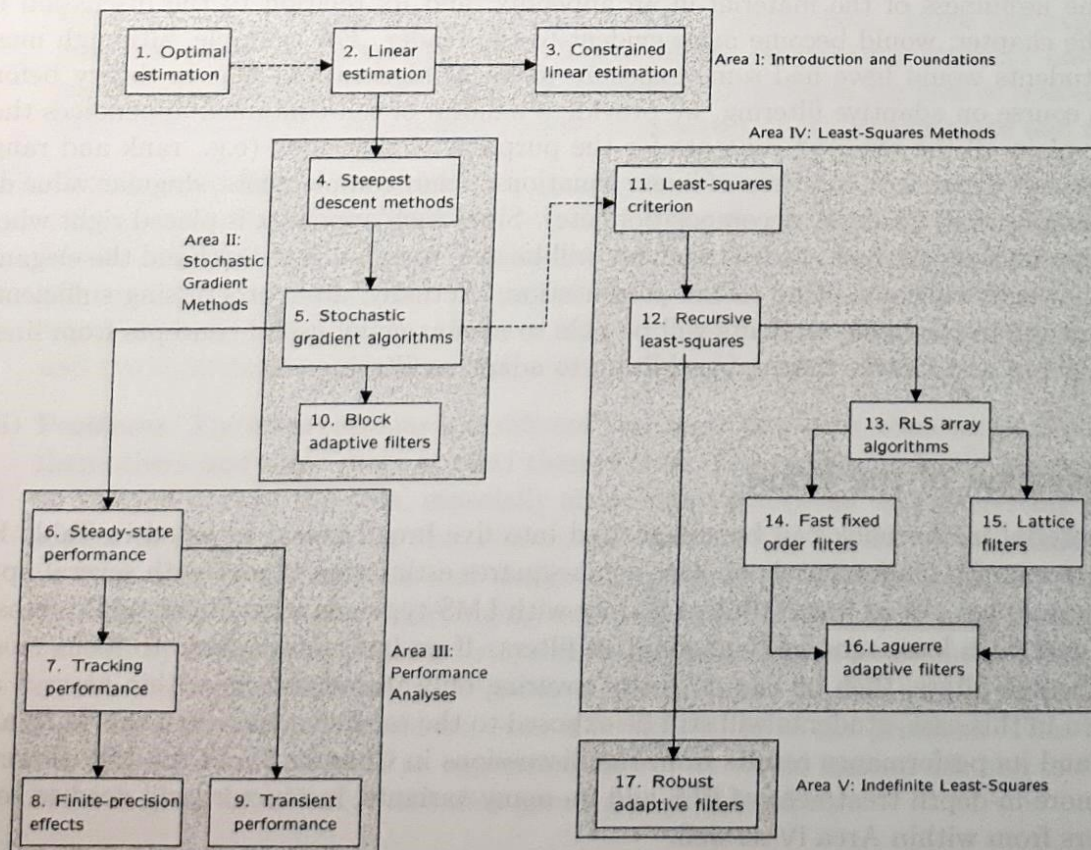


Figure P.1. Dependencies among the chapters. Instructors can design different course sequences in accordance with their needs and interests.

arrow. A dashed arrow indicates that the dependency between the respective chapters is weak and, if desired, the chapters can be covered independently of each other. For example, in order to cover Chapter 6, the instructor would need to cover Chapter 5, which, in turn, relies on Chapters 4 and 2. The material in Chapter 1 is not necessary for Chapter 2 but it is useful for a better understanding of it; actually, the material in Chapter 1 is presented in such a way that it also provides a useful review of basic probability theory concepts.

Figure P.1 can be used by instructors to design different course sequences according to their needs and interests. For example, if the instructor is interested in covering only LMS-type adaptive filters and in studying their performance, then one possibility is to cover material from within Chapters 2, 4, 5, 6, 7, and 9. Later, in Table P.4, we list the key sections from within each chapter; the other sections usually contain more advanced material, which students can read once they understand the key concepts from the main sections.

List of appendices. The book contains over 40 appendices that complement the material in the chapters and provide useful reviews and further analyses and connections. All appendices are listed in Table P.2. Readers interested in a quick review of basic linear algebra and matrix theory concepts may consult initially Apps. 1.A, 2.A, and 3.A, and subsequently Apps. 9.H, 11.B, 11.C, 13.A, 14.A, and 14.B for more advanced topics. These appendices are highlighted by the symbol * in Table P.2. Readers interested in a quick review of basic probability theory concepts should consult Secs. 1.1–1.4 and App. 1.B.

Computer projects. The book contains 24 computer projects that have been designed to reinforce the concepts discussed in the chapters. The projects are listed in Table P.3, and most of them cover topics of interest in communications and signal processing such as channel estimation, linear equalization (adaptive and channel-estimation based), decision-feedback equalization (also adaptive and channel-estimation based), adaptive blind equalization, CDMA and RAKE receivers, OFDM receivers, tracking of Rayleigh fading channels, line echo cancellation, acoustic echo cancellation, active noise control, beamforming, finite-precision effects, etc. Detailed MATLAB programs that solve all projects can be downloaded by all readers from the publisher's website (ftp://ftp.wiley.com/public/sci_tech_med/filtering/). These programs are offered without any guarantees. While we have found them to be effective for the instructional purposes of this textbook, the programs are not intended to be examples of full-blown or optimized designs; practitioners should use them at their own risk. For example, in order to keep the codes at a level that is easy to understand by students, we have often decided to sacrifice performance in lieu of simplicity.

Audience

The book is intended for a graduate-level course on adaptive filtering. Although it is beneficial that students have some familiarity with basic concepts from matrix theory, linear algebra, and random variables, the book includes several appendices on background material in these areas. The review is done in a motivated manner and is tailored to the needs of the presentation. From our experience, these reviews are sufficient for a thorough understanding of the discussions in the book. In addition, several of the problems reinforce the linear algebraic and matrix concepts, so much so that students will get valuable training in linear algebra and matrix theory, in addition to adaptive filtering, from reading (and understanding) this book.

The book is also intended to be a reference for researchers, which explains why we have chosen to include some advanced topics in several places. As a result, the book contains ample material for instructors to design courses according to their interests. Clearly, we do not expect instructors to cover all the material in the book in a typical course offering; such an objective would be counter-productive and even impossible. In our own teaching of

Table P.2. A listing of all appendices in the book. Appendices highlighted by the symbol * provide reviews of linear algebra and matrix theory concepts.

Appendix	Title
1.A*	Hermitian and positive-definite matrices.
1.B	Gaussian random variables.
2.A*	Range spaces and nullspaces of matrices.
2.B	Complex gradients and Hessians.
2.C	The Kalman filter.
3.A*	Schur complements.
3.B	A primer on channel equalization.
3.C	Causal Wiener-Hopf filtering.
6.A	Interpretations of the energy relation.
6.B	Relating ϵ -NLMS to LMS.
6.C	Affine projection performance condition.
9.A	Stability bound.
9.B	Stability of ϵ -NLMS.
9.C	Adaptive filters with error nonlinearities.
9.D	Convergence time of adaptive filters.
9.E	Learning behavior of adaptive filters.
9.F	Independence and averaging analysis.
9.G	Physical interpretation of energy relation.
9.H*	Kronecker products.
10.A	DCT-transformed regressors.
10.B	More constrained DFT block filters.
10.C	Overlap-add DFT-based block adaptive filters.
10.D	DCT-based block adaptive filters.
10.E	DHT-based block adaptive filters.
11.A	Equivalence results in linear estimation.
11.B*	The QR decomposition.
11.C*	The singular value decomposition.
12.A	Kalman filtering and recursive least-squares.
12.B	Extended RLS algorithms.
13.A*	Unitary transformations.
13.B	Array algorithms for Kalman filtering.
14.A*	Hyperbolic rotations.
14.B*	Hyperbolic basis rotations.
14.C	Backward consistency and minimality.
14.D	The Chandrasekhar filter.
16.A	Modeling with orthonormal basis functions.
16.B	Efficient matrix-vector multiplication.
16.C	Lyapunov equations.
17.A	Arbitrary coefficient matrices.
17.B	Total-least-squares.
17.C	\mathcal{H}^∞ filters.
17.D	Stationary points.

Table P.3. A listing of all computer projects in the book. MATLAB programs that solve these projects can be downloaded by all readers from the publisher's website (ftp://ftp.wiley.com/public/sci_tech_med/filtering/).

Computer project	Topic
1	Comparing optimal and suboptimal estimators.
2	Linear equalization and decision devices.
3.1	Beamforming.
3.2	Decision-feedback equalization.
4.1	Constant-modulus criterion.
4.2	Linear prediction.
5.1	Constant-modulus algorithm.
5.2	Adaptive channel equalization.
5.3	Blind adaptive equalization.
6	Line echo cancellation.
7	Tracking Rayleigh fading channels.
8	Quantization effects in adaptive filtering.
9	Transient behavior of LMS and LMF.
10	Acoustic echo cancellation.
11.1	Amplitude tone detection.
11.2	An OFDM receiver.
11.3	CDMA and RAKE receivers.
12.1	Channel estimation with insufficient excitation.
12.2	Tracking a Rayleigh fading channel by extended RLS.
13	Performance of array implementations in finite precision.
14	Stability issues in fast least-squares.
15	Performance of lattice filters in finite precision.
16	Laguerre and FIR implementations.
17	Active noise control.

the material, we instead *focus on some key sections and request that students complement the discussions by means of reading and problem solving*. As explained below, several key sections in the chapters have been designed to convey the main concepts; while the remaining sections tend to include more advanced material and also illustrative examples. Once students understand the basic principles, you will be amazed at how well they can follow the other sections on their own and even solve the pertinent problems.

Guidelines to Instructors

As we explained before, instructors can use Fig. P.1 to design different course sequences according to their interests. For example, a course that is focused solely on LMS-type filters and their performance can be designed by covering only material from within Chapters 2 and 4–9. Even then, instructors do not need to cover the entire material from each one of these chapters. Instead, they need only cover some key sections and, if desired, ask students to complement the discussions in class with reading material from the other more advanced sections. To facilitate such a course planning, Table P.4 lists in boldface the key sections for the different chapters in the book for both lecturing and reading purposes.

For example, the key sections in Chapter 2 are Sec. 2.1 (Normal Equations), Sec. 2.4 (Orthogonality Condition), and Sec. 2.6 (Linear Models). These sections formulate and solve

the linear least-mean-squares estimation problem and specialize the results to the important class of linear models (which is frequent in applications). The other sections in Chapter 2 complement the discussions with design examples, among other things.

As a second example, Chapter 9 studies the transient performance of a large family of adaptive filters in a uniform manner. The main idea is captured by the transient analysis of the LMS algorithm in Sec. 9.5, which uses the machinery developed in Sec. 9.4. Once students understand the framework as applied to LMS, they will be able to study the transient analysis of the other filters mostly on their own. This is one key advantage of adopting and emphasizing a uniform treatment of adaptive filter performance throughout our presentation. Similar remarks hold for the steady-state, tracking, and finite-precision performance analyses of Chapters 6–8. It is sufficient to illustrate how the methodology applies to the special case of LMS, for example, by covering Secs. 6.5, 7.5, and 8.5, which in turn rely on the machinery developed in Secs. 6.4, 7.4, and 8.4. The remaining sections in Chapters 6–8 extend the same type of analysis to other (more demanding) adaptive filters. Here again, students can do well in studying the extensions on their own if desired.

Table P.4. A suggested list of key sections (in boldface) for both lecturing and reading in all chapters along with relevant complementary sections (in normal font). At the instructor's discretion, some of the key sections for reading could, of course, be covered during lecturing as well; especially those dealing with basic review material on linear algebraic and matrix theory concepts.

Key sections for lecturing	Key sections for reading
Secs. 1.1, 1.2, 1.3 , 1.4	App. 1.A
Secs. 2.1 , 2.2, 2.3, 2.4 , 2.5, 2.6	Apps. 2.A , 2.B, 2.C
Secs. 3.1 , 3.2, 3.3 , 3.4	Apps. 3.A , 3.B
Secs. 4.1, 4.2 , 4.3 , 4.4, 4.5	
Secs. 5.1, 5.2 , 5.3–5.6, 5.9	Secs. 5.6 , 5.7, 5.8, 5.10
Secs. 6.1, 6.2, 6.3–6.5	Secs. 6.6–6.10
Secs. 7.1,7.2, 7.3–7.5	Secs. 7.6–7.11
Secs. 8.1–8.5	Secs. 8.6–8.9
Secs. 9.1–9.3, 9.4 , 9.5	Secs. 9.6–9.7 Apps. 9.A, 9.B, 9.C , 9.D , 9.E
Secs. 10.1 , 10.2, 10.3, 10.4 , 10.5	Apps. 10.A–10.E
Secs. 11.1–11.4	Sec. 11.5 , Apps. 11.A–11.C
Secs. 12.1–12.3	Sec. 12.4 , Apps. 12.A–12.B
Secs. 13.1–13.3, 13.5 , 13.6 , App. 13.A	Sec. 13.7
Secs. 14.1	Secs. 14.2 , 14.3–14.6 Apps. 14.A–14.B
Secs. 15.1, 15.2–15.4 , 15.5, 15.6 , 15.7	Secs. 15.8, 15.9–15.13
Secs. 16.1–16.2 , 16.8, 16.9	Secs. 16.3–16.7, 16.10–14 Apps. 16.A–16.C
Secs. 17.1–17.5	Apps. 17.A–17.D

Some Features of Our Treatment

There are some distinctive features in our treatment of adaptive filtering. Among other features, experts will be able to notice the following contributions:

- (a) We treat a large variety of adaptive algorithms, as listed in Tables P.5 and P.6 for

both LMS-type and RLS-type filters.

Table P.5. A list of LMS-type adaptive algorithms covered in the book.

Algorithm	Description
LMS	Least-mean-squares algorithm
NLMS	Normalized LMS
NLMS with power normalization	
leaky-LMS	
constrained LMS	
LMF	Least-mean-fourth algorithm
LMMN	Least-mean-mixed norm algorithm
sign-error LMS	
sign-regressor LMS	
sign-sign LMS	
FxLMS	Filtered-x LMS
FeLMS	Filtered-error LMS
APA	Affine projection algorithm
CMA	Constant modulus algorithm
NCMA	Normalized CMA
RCA	Reduced constellation algorithm
MMA	Multi-modulus algorithm
DFT-domain LMS	Transform-domain LMS
DCT-domain LMS	
DFT-based block LMS	
DCT-based block LMS	
DHT-based block LMS	
closed and open-loop subband LMS	
Robust filters	<i>A priori</i> and <i>a posteriori</i> forms

- (b) Chapters 6–9 study the mean-square performance of adaptive filters by resorting to energy-conservation arguments. While the performance of different adaptive filters is usually studied separately in the literature, the framework adopted in these chapters applies uniformly across different classes of adaptive filters. In addition, the same framework is used for steady-state analysis, transient analysis, tracking analysis, fixed-point analysis, and robustness analysis.
- (c) Chapter 10 studies block adaptive filters, and the related class of subband adaptive filters, in a manner that clarifies the connections between these two families more directly than prior treatments. Our presentation also indicates how to move beyond DFT-based transforms and how to use other classes of orthogonal transforms for block adaptive filtering.
- (d) Chapters 11–15 provide a detailed treatment of least-squares adaptive filters that is distinct from prevailing approaches in a handful of respects. First, we focus on regularized least-squares problems from the onset and take the regularization factor into account in all derivations. Second, we insist on deriving time- and order-update relations independent of any structure in the regression data (e.g., we do not require the regressors to arise from a tapped-delay-line implementation). In this way, we are able to develop efficient least-squares filtering even for some non-FIR structures. Third, we emphasize the role and benefits of array-based schemes. And, finally, we highlight the role of geometric constructions and the insights they bring into least-squares theory.

Table P.6. A list of RLS-type adaptive algorithms covered in the book.

Algorithm	Description
RLS	Recursive least-squares algorithm
GN	Gauss-Newton algorithm
sliding-window RLS	
block RLS	Multi-channel RLS
extended RLS	
QR-RLS (square-root information RLS)	Array-based recursive least-squares
inverse-QR RLS (square-root RLS)	
Fast array RLS	
FTF	Fast transversal filter
FAEST	Fast <i>a posteriori</i> error sequential technique
Fast Kalman	
<i>a posteriori</i> -based LSL	Least-squares lattice
<i>a priori</i> -based LSL	Least-squares lattice
<i>a posteriori</i> error-feedback LSL	
<i>a priori</i> error-feedback LSL	
Normalized LSL	
Array-based LSL	
Laguerre FTF	Extended FTF
Laguerre FAEST	Extended FAEST
Laguerre fast Kalman	
<i>a posteriori</i> -based Laguerre LSL	Least-squares Laguerre lattice
<i>a priori</i> -based Laguerre LSL	
<i>a posteriori</i> error-feedback Laguerre LSL	
<i>a priori</i> error-feedback Laguerre LSL	
Normalized Laguerre LSL	
Array-based Laguerre LSL	
Kalman filter	
Robust filters	<i>A priori</i> and <i>a posteriori</i> forms

- (e) Chapter 16 shows how the theory of fast least-squares methods (both for fixed-order and order-recursive problems) is not limited to tapped delay lines; an observation that extends classical derivations and developments. In the chapter, we illustrate this fact by studying Laguerre adaptive filters; they are obtained by replacing the delay operators in an FIR structure by first-order all-pass functions. Although the resulting regressors no longer possess shift-structure, it turns out that fast least-squares filters are still possible.
- (f) Chapter 17 develops the theory of robust adaptive filters by studying indefinite least-squares problems and by relying on energy arguments as well. In the process, the robustness and optimality properties of several adaptive filters are clarified. The presentation in this chapter is developed in a manner that parallels our treatment of least-squares problems in Chapters 11–12 so that readers can appreciate the similarities and distinctions between both theories (classical least-squares versus indefinite least-squares).