

**Mathematical Methods and Algorithms
for
Signal Processing**

Todd K. Moon
Utah State University

Wynn C. Stirling
Brigham Young University

PRENTICE HALL
Upper Saddle River, NJ 07458

This previously included a CD. The
CD contents can now be accessed
at www.prenhall.com/moon. Thank You.

Contents

I	Introduction and Foundations	1
1	Introduction and Foundations	3
1.1	What is signal processing?	3
1.2	Mathematical topics embraced by signal processing	5
1.3	Mathematical models	6
1.4	Models for linear systems and signals	7
1.4.1	Linear discrete-time models	7
1.4.2	Stochastic MA and AR models	12
1.4.3	Continuous-time notation	20
1.4.4	Issues and applications	21
1.4.5	Identification of the modes	26
1.4.6	Control of the modes	28
1.5	Adaptive filtering	28
1.5.1	System identification	29
1.5.2	Inverse system identification	29
1.5.3	Adaptive predictors	29
1.5.4	Interference cancellation	30
1.6	Gaussian random variables and random processes	31
1.6.1	Conditional Gaussian densities	36
1.7	Markov and Hidden Markov Models	37
1.7.1	Markov models	37
1.7.2	Hidden Markov models	39
1.8	Some aspects of proofs	41
1.8.1	Proof by computation: direct proof	43
1.8.2	Proof by contradiction	45
1.8.3	Proof by induction	46
1.9	An application: LFSRs and Massey's algorithm	48
1.9.1	Issues and applications of LFSRs	50
1.9.2	Massey's algorithm	52
1.9.3	Characterization of LFSR length in Massey's algorithm	53
1.10	Exercises	58
1.11	References	67
II	Vector Spaces and Linear Algebra	69
2	Signal Spaces	71
2.1	Metric spaces	72
2.1.1	Some topological terms	76
2.1.2	Sequences, Cauchy sequences, and completeness	78

2.1.3	Technicalities associated with the L_p and L_∞ spaces	82
2.2	Vector spaces	84
2.2.1	Linear combinations of vectors	87
2.2.2	Linear independence	88
2.2.3	Basis and dimension	90
2.2.4	Finite-dimensional vector spaces and matrix notation	93
2.3	Norms and normed vector spaces	93
2.3.1	Finite-dimensional normed linear spaces	97
2.4	Inner products and inner-product spaces	97
2.4.1	Weak convergence	99
2.5	Induced norms	99
2.6	The Cauchy–Schwarz inequality	100
2.7	Direction of vectors: Orthogonality	101
2.8	Weighted inner products	103
2.8.1	Expectation as an inner product	105
2.9	Hilbert and Banach spaces	106
2.10	Orthogonal subspaces	107
2.11	Linear transformations: Range and nullspace	108
2.12	Inner-sum and direct-sum spaces	110
2.13	Projections and orthogonal projections	113
2.13.1	Projection matrices	115
2.14	The projection theorem	116
2.15	Orthogonalization of vectors	118
2.16	Some final technicalities for infinite dimensional spaces	121
2.17	Exercises	121
2.18	References	129
3	Representation and Approximation in Vector Spaces	130
3.1	The approximation problem in Hilbert space	130
3.1.1	The Grammian matrix	133
3.2	The orthogonality principle	135
3.2.1	Representations in infinite-dimensional space	136
3.3	Error minimization via gradients	137
3.4	Matrix representations of least-squares problems	138
3.4.1	Weighted least-squares	140
3.4.2	Statistical properties of the least-squares estimate	140
3.5	Minimum error in Hilbert-space approximations	141
	<i>Applications of the orthogonality theorem</i>	
3.6	Approximation by continuous polynomials	143
3.7	Approximation by discrete polynomials	145
3.8	Linear regression	147
3.9	Least-squares filtering	149
3.9.1	Least-squares prediction and AR spectrum estimation	154
3.10	Minimum mean-square estimation	156
3.11	Minimum mean-squared error (MMSE) filtering	157
3.12	Comparison of least squares and minimum mean squares	161
3.13	Frequency-domain optimal filtering	162
3.13.1	Brief review of stochastic processes and Laplace transforms	162

3.13.2	Two-sided Laplace transforms and their decompositions	165
3.13.3	The Wiener–Hopf equation	169
3.13.4	Solution to the Wiener–Hopf equation	171
3.13.5	Examples of Wiener filtering	174
3.13.6	Mean-square error	176
3.13.7	Discrete-time Wiener filters	176
3.14	A dual approximation problem	179
3.15	Minimum-norm solution of underdetermined equations	182
3.16	Iterative Reweighted LS (IRLS) for L_p optimization	183
3.17	Signal transformation and generalized Fourier series	186
3.18	Sets of complete orthogonal functions	190
3.18.1	Trigonometric functions	190
3.18.2	Orthogonal polynomials	190
3.18.3	Sinc functions	193
3.18.4	Orthogonal wavelets	194
3.19	Signals as points: Digital communications	208
3.19.1	The detection problem	210
3.19.2	Examples of basis functions used in digital communications	212
3.19.3	Detection in nonwhite noise	213
3.20	Exercises	215
3.21	References	228
4	Linear Operators and Matrix Inverses	229
4.1	Linear operators	230
4.1.1	Linear functionals	231
4.2	Operator norms	232
4.2.1	Bounded operators	233
4.2.2	The Neumann expansion	235
4.2.3	Matrix norms	235
4.3	Adjoint operators and transposes	237
4.3.1	A dual optimization problem	239
4.4	Geometry of linear equations	239
4.5	Four fundamental subspaces of a linear operator	242
4.5.1	The four fundamental subspaces with non-closed range	246
4.6	Some properties of matrix inverses	247
4.6.1	Tests for invertibility of matrices	248
4.7	Some results on matrix rank	249
4.7.1	Numeric rank	250
4.8	Another look at least squares	251
4.9	Pseudoinverses	251
4.10	Matrix condition number	253
4.11	Inverse of a small-rank adjustment	258
4.11.1	An application: the RLS filter	259
4.11.2	Two RLS applications	261
4.12	Inverse of a block (partitioned) matrix	264
4.12.1	Application: Linear models	267
4.13	Exercises	268
4.14	References	274

5	Some Important Matrix Factorizations	275
5.1	The LU factorization	275
5.1.1	Computing the determinant using the LU factorization	277
5.1.2	Computing the LU factorization	278
5.2	The Cholesky factorization	283
5.2.1	Algorithms for computing the Cholesky factorization	284
5.3	Unitary matrices and the QR factorization	285
5.3.1	Unitary matrices	285
5.3.2	The QR factorization	286
5.3.3	QR factorization and least-squares filters	286
5.3.4	Computing the QR factorization	287
5.3.5	Householder transformations	287
5.3.6	Algorithms for Householder transformations	291
5.3.7	QR factorization using Givens rotations	293
5.3.8	Algorithms for QR factorization using Givens rotations	295
5.3.9	Solving least-squares problems using Givens rotations	296
5.3.10	Givens rotations via CORDIC rotations	297
5.3.11	Recursive updates to the QR factorization	299
5.4	Exercises	300
5.5	References	304
6	Eigenvalues and Eigenvectors	305
6.1	Eigenvalues and linear systems	305
6.2	Linear dependence of eigenvectors	308
6.3	Diagonalization of a matrix	309
6.3.1	The Jordan form	311
6.3.2	Diagonalization of self-adjoint matrices	312
6.4	Geometry of invariant subspaces	316
6.5	Geometry of quadratic forms and the minimax principle	318
6.6	Extremal quadratic forms subject to linear constraints	324
6.7	The Gershgorin circle theorem	324
	<i>Application of Eigendecomposition methods</i>	
6.8	Karhunen–Loève low-rank approximations and principal methods	327
6.8.1	Principal component methods	329
6.9	Eigenfilters	330
6.9.1	Eigenfilters for random signals	330
6.9.2	Eigenfilter for designed spectral response	332
6.9.3	Constrained eigenfilters	334
6.10	Signal subspace techniques	336
6.10.1	The signal model	336
6.10.2	The noise model	337
6.10.3	Pisarenko harmonic decomposition	338
6.10.4	MUSIC	339
6.11	Generalized eigenvalues	340
6.11.1	An application: ESPRIT	341
6.12	Characteristic and minimal polynomials	342
6.12.1	Matrix polynomials	342
6.12.2	Minimal polynomials	344
6.13	Moving the eigenvalues around: Introduction to linear control	344
6.14	Noiseless constrained channel capacity	347

6.15	Computation of eigenvalues and eigenvectors	350
6.15.1	Computing the largest and smallest eigenvalues	350
6.15.2	Computing the eigenvalues of a symmetric matrix	351
6.15.3	The QR iteration	352
6.16	Exercises	355
6.17	References	368
7	The Singular Value Decomposition	369
7.1	Theory of the SVD	369
7.2	Matrix structure from the SVD	372
7.3	Pseudoinverses and the SVD	373
7.4	Numerically sensitive problems	375
7.5	Rank-reducing approximations: Effective rank	377
	<i>Applications of the SVD</i>	
7.6	System identification using the SVD	378
7.7	Total least-squares problems	381
7.7.1	Geometric interpretation of the TLS solution	385
7.8	Partial total least squares	386
7.9	Rotation of subspaces	389
7.10	Computation of the SVD	390
7.11	Exercises	392
7.12	References	395
8	Some Special Matrices and Their Applications	396
8.1	Modal matrices and parameter estimation	396
8.2	Permutation matrices	399
8.3	Toeplitz matrices and some applications	400
8.3.1	Durbin's algorithm	402
8.3.2	Predictors and lattice filters	403
8.3.3	Optimal predictors and Toeplitz inverses	407
8.3.4	Toeplitz equations with a general right-hand side	408
8.4	Vandermonde matrices	409
8.5	Circulant matrices	410
8.5.1	Relations among Vandermonde, circulant, and companion matrices	412
8.5.2	Asymptotic equivalence of the eigenvalues of Toeplitz and circulant matrices	413
8.6	Triangular matrices	416
8.7	Properties preserved in matrix products	417
8.8	Exercises	418
8.9	References	421
9	Kronecker Products and the Vec Operator	422
9.1	The Kronecker product and Kronecker sum	422
9.2	Some applications of Kronecker products	425
9.2.1	Fast Hadamard transforms	425
9.2.2	DFT computation using Kronecker products	426
9.3	The vec operator	428
9.4	Exercises	431
9.5	References	433

III	Detection, Estimation, and Optimal Filtering	435
10	Introduction to Detection and Estimation, and Mathematical Notation	437
10.1	Detection and estimation theory	437
10.1.1	Game theory and decision theory	438
10.1.2	Randomization	440
10.1.3	Special cases	441
10.2	Some notational conventions	442
10.2.1	Populations and statistics	443
10.3	Conditional expectation	444
10.4	Transformations of random variables	445
10.5	Sufficient statistics	446
10.5.1	Examples of sufficient statistics	450
10.5.2	Complete sufficient statistics	451
10.6	Exponential families	453
10.7	Exercises	456
10.8	References	459
11	Detection Theory	460
11.1	Introduction to hypothesis testing	460
11.2	Neyman–Pearson theory	462
11.2.1	Simple binary hypothesis testing	462
11.2.2	The Neyman–Pearson lemma	463
11.2.3	Application of the Neyman–Pearson lemma	466
11.2.4	The likelihood ratio and the receiver operating characteristic (ROC)	467
11.2.5	A Poisson example	468
11.2.6	Some Gaussian examples	469
11.2.7	Properties of the ROC	480
11.3	Neyman–Pearson testing with composite binary hypotheses	483
11.4	Bayes decision theory	485
11.4.1	The Bayes principle	486
11.4.2	The risk function	487
11.4.3	Bayes risk	489
11.4.4	Bayes tests of simple binary hypotheses	490
11.4.5	Posterior distributions	494
11.4.6	Detection and sufficiency	498
11.4.7	Summary of binary decision problems	498
11.5	Some M -ary problems	499
11.6	Maximum-likelihood detection	503
11.7	Approximations to detection performance: The union bound	503
11.8	Invariant Tests	504
11.8.1	Detection with random (nuisance) parameters	507
11.9	Detection in continuous time	512
11.9.1	Some extensions and precautions	516
11.10	Minimax Bayes decisions	520
11.10.1	Bayes envelope function	520
11.10.2	Minimax rules	523
11.10.3	Minimax Bayes in multiple-decision problems	524

11.10.4	Determining the least favorable prior	528
11.10.5	A minimax example and the minimax theorem	529
11.11	Exercises	532
11.12	References	541
12	Estimation Theory	542
12.1	The maximum-likelihood principle	542
12.2	ML estimates and sufficiency	547
12.3	Estimation quality	548
12.3.1	The score function	548
12.3.2	The Cramér–Rao lower bound	550
12.3.3	Efficiency	552
12.3.4	Asymptotic properties of maximum-likelihood estimators	553
12.3.5	The multivariate normal case	556
12.3.6	Minimum-variance unbiased estimators	559
12.3.7	The linear statistical model	561
12.4	Applications of ML estimation	561
12.4.1	ARMA parameter estimation	561
12.4.2	Signal subspace identification	565
12.4.3	Phase estimation	566
12.5	Bayes estimation theory	568
12.6	Bayes risk	569
12.6.1	MAP estimates	573
12.6.2	Summary	574
12.6.3	Conjugate prior distributions	574
12.6.4	Connections with minimum mean-squared estimation	577
12.6.5	Bayes estimation with the Gaussian distribution	578
12.7	Recursive estimation	580
12.7.1	An example of non-Gaussian recursive Bayes	582
12.8	Exercises	584
12.9	References	590
13	The Kalman Filter	591
13.1	The state-space signal model	591
13.2	Kalman filter I: The Bayes approach	592
13.3	Kalman filter II: The innovations approach	595
13.3.1	Innovations for processes with linear observation models	596
13.3.2	Estimation using the innovations process	597
13.3.3	Innovations for processes with state-space models	598
13.3.4	A recursion for $P_{t t-1}$	599
13.3.5	The discrete-time Kalman filter	601
13.3.6	Perspective	602
13.3.7	Comparison with the RLS adaptive filter algorithm	603
13.4	Numerical considerations: Square-root filters	604
13.5	Application in continuous-time systems	606
13.5.1	Conversion from continuous time to discrete time	606
13.5.2	A simple kinematic example	606
13.6	Extensions of Kalman filtering to nonlinear systems	607

13.7	Smoothing	613
13.7.1	The Rauch–Tung–Streibel fixed-interval smoother	613
13.8	Another approach: H_∞ smoothing	616
13.9	Exercises	617
13.10	References	620
IV	Iterative and Recursive Methods in Signal Processing	621
14	Basic Concepts and Methods of Iterative Algorithms	623
14.1	Definitions and qualitative properties of iterated functions	624
14.1.1	Basic theorems of iterated functions	626
14.1.2	Illustration of the basic theorems	627
14.2	Contraction mappings	629
14.3	Rates of convergence for iterative algorithms	631
14.4	Newton’s method	632
14.5	Steepest descent	637
14.5.1	Comparison and discussion: Other techniques	642
	<i>Some Applications of Basic Iterative Methods</i>	
14.6	LMS adaptive Filtering	643
14.6.1	An example LMS application	645
14.6.2	Convergence of the LMS algorithm	646
14.7	Neural networks	648
14.7.1	The backpropagation training algorithm	650
14.7.2	The nonlinearity function	653
14.7.3	The forward–backward training algorithm	654
14.7.4	Adding a momentum term	654
14.7.5	Neural network code	655
14.7.6	How many neurons?	658
14.7.7	Pattern recognition: ML or NN?	659
14.8	Blind source separation	660
14.8.1	A bit of information theory	660
14.8.2	Applications to source separation	662
14.8.3	Implementation aspects	664
14.9	Exercises	665
14.10	References	668
15	Iteration by Composition of Mappings	670
15.1	Introduction	670
15.2	Alternating projections	671
15.2.1	An applications: bandlimited reconstruction	675
15.3	Composite mappings	676
15.4	Closed mappings and the global convergence theorem	677
15.5	The composite mapping algorithm	680
15.5.1	Bandlimited reconstruction, revisited	681
15.5.2	An example: Positive sequence determination	681
15.5.3	Matrix property mappings	683
15.6	Projection on convex sets	689
15.7	Exercises	693
15.8	References	694

16	Other Iterative Algorithms	695
16.1	Clustering	695
16.1.1	An example application: Vector quantization	695
16.1.2	An example application: Pattern recognition	697
16.1.3	k -means Clustering	698
16.1.4	Clustering using fuzzy k -means	700
16.2	Iterative methods for computing inverses of matrices	701
16.2.1	The Jacobi method	702
16.2.2	Gauss–Seidel iteration	703
16.2.3	Successive over-relaxation (SOR)	705
16.3	Algebraic reconstruction techniques (ART)	706
16.4	Conjugate-direction methods	708
16.5	Conjugate-gradient method	710
16.6	Nonquadratic problems	713
16.7	Exercises	713
16.8	References	715
17	The EM Algorithm in Signal Processing	717
17.1	An introductory example	718
17.2	General statement of the EM algorithm	721
17.3	Convergence of the EM algorithm	723
17.3.1	Convergence rate: Some generalizations	724
	<i>Example applications of the EM algorithm</i>	
17.4	Introductory example, revisited	725
17.5	Emission computed tomography (ECT) image reconstruction	725
17.6	Active noise cancellation (ANC)	729
17.7	Hidden Markov models	732
17.7.1	The E- and M-steps	734
17.7.2	The forward and backward probabilities	735
17.7.3	Discrete output densities	736
17.7.4	Gaussian output densities	736
17.7.5	Normalization	737
17.7.6	Algorithms for HMMs	738
17.8	Spread-spectrum, multiuser communication	740
17.9	Summary	743
17.10	Exercises	744
17.11	References	747
V	Methods of Optimization	749
18	Theory of Constrained Optimization	751
18.1	Basic definitions	751
18.2	Generalization of the chain rule to composite functions	755
18.3	Definitions for constrained optimization	757
18.4	Equality constraints: Lagrange multipliers	758
18.4.1	Examples of equality-constrained optimization	764
18.5	Second-order conditions	767
18.6	Interpretation of the Lagrange multipliers	770
18.7	Complex constraints	773
18.8	Duality in optimization	773

18.9	Inequality constraints: Kuhn–Tucker conditions	777
18.9.1	Second-order conditions for inequality constraints	783
18.9.2	An extension: Fritz John conditions	783
18.10	Exercises	784
18.11	References	786
19	Shortest-Path Algorithms and Dynamic Programming	787
19.1	Definitions for graphs	787
19.2	Dynamic programming	789
19.3	The Viterbi algorithm	791
19.4	Code for the Viterbi algorithm	795
19.4.1	Related algorithms: Dijkstra's and Warshall's	798
19.4.2	Complexity comparisons of Viterbi and Dijkstra	799
	<i>Applications of path search algorithms</i>	
19.5	Maximum-likelihood sequence estimation	800
19.5.1	The intersymbol interference (ISI) channel	800
19.5.2	Code-division multiple access	804
19.5.3	Convolutional decoding	806
19.6	HMM likelihood analysis and HMM training	808
19.6.1	Dynamic warping	811
19.7	Alternatives to shortest-path algorithms	813
19.8	Exercises	815
19.9	References	817
20	Linear Programming	818
20.1	Introduction to linear programming	818
20.2	Putting a problem into standard form	819
20.2.1	Inequality constraints and slack variables	819
20.2.2	Free variables	820
20.2.3	Variable-bound constraints	822
20.2.4	Absolute value in the objective	823
20.3	Simple examples of linear programming	823
20.4	Computation of the linear programming solution	824
20.4.1	Basic variables	824
20.4.2	Pivoting	826
20.4.3	Selecting variables on which to pivot	828
20.4.4	The effect of pivoting on the value of the problem	829
20.4.5	Summary of the simplex algorithm	830
20.4.6	Finding the initial basic feasible solution	831
20.4.7	MATLAB® code for linear programming	834
20.4.8	Matrix notation for the simplex algorithm	835
20.5	Dual problems	836
20.6	Karmarker's algorithm for LP	838
20.6.1	Conversion to Karmarker standard form	842
20.6.2	Convergence of the algorithm	844
20.6.3	Summary and extensions	846
	<i>Examples and applications of linear programming</i>	
20.7	Linear-phase FIR filter design	846
20.7.1	Least-absolute-error approximation	847
20.8	Linear optimal control	849

20.9	Exercises	850
20.10	References	853
A	Basic Concepts and Definitions	855
A.1	Set theory and notation	855
A.2	Mappings and functions	859
A.3	Convex functions	860
A.4	O and o Notation	861
A.5	Continuity	862
A.6	Differentiation	864
A.6.1	Differentiation with a single real variable	864
A.6.2	Partial derivatives and gradients on \mathbb{R}^m	865
A.6.3	Linear approximation using the gradient	867
A.6.4	Taylor series	868
A.7	Basic constrained optimization	869
A.8	The Hölder and Minkowski inequalities	870
A.9	Exercises	871
A.10	References	876
B	Completing the Square	877
B.1	The scalar case	877
B.2	The matrix case	879
B.3	Exercises	879
C	Basic Matrix Concepts	880
C.1	Notational conventions	880
C.2	Matrix Identity and Inverse	882
C.3	Transpose and trace	883
C.4	Block (partitioned) matrices	885
C.5	Determinants	885
C.5.1	Basic properties of determinants	885
C.5.2	Formulas for the determinant	887
C.5.3	Determinants and matrix inverses	889
C.6	Exercises	889
C.7	References	890
D	Random Processes	891
D.1	Definitions of means and correlations	891
D.2	Stationarity	892
D.3	Power spectral-density functions	893
D.4	Linear systems with stochastic inputs	894
D.4.1	Continuous-time signals and systems	894
D.4.2	Discrete-time signals and systems	895
D.5	References	895
E	Derivatives and Gradients	896
E.1	Derivatives of vectors and scalars with respect to a real vector	896
E.1.1	Some important gradients	897
E.2	Derivatives of real-valued functions of real matrices	899
E.3	Derivatives of matrices with respect to scalars, and vice versa	901
E.4	The transformation principle	903
E.5	Derivatives of products of matrices	903

E.6	Derivatives of powers of a matrix	904
E.7	Derivatives involving the trace	906
E.8	Modifications for derivatives of complex vectors and matrices	908
E.9	Exercises	910
E.10	References	912
F	Conditional Expectations of Multinomial and Poisson r.v.s	913
F.1	Multinomial distributions	913
F.2	Poisson random variables	914
F.3	Exercises	914
	Bibliography	915
	Index	929