Statistical Signal Processing
Detection, Estimation, and Time Series Analysis

Louis L. Scharf
University of Colorado at Boulder

with Cédric Demeure collaborating on Chapters 10 and 11
Contents

CHAPTER 1

Introduction 1
1.1 Statistical Signal Processing and Related Topics  1
1.2 The Structure of Statistical Reasoning  3
1.3 A Detection Problem  4
1.4 An Estimation Problem  7
1.5 A Time Series Problem  9
1.6 Notation and Terminology  11
   Probability Distributions  11
   Linear Models  12
References and Comments  18
Problems  19

CHAPTER 2

Rudiments of Linear Algebra and Multivariate Normal Theory 23
2.1 Vector Spaces  24
   Euclidean Space  24
   Hilbert Space  25
   Matrices  25
2.2 Linear Independence  25
   Gram Determinant  26
   Sequences of Gram Determinants  27
   Cholesky Factors of the Gram Matrix  29
2.3 QR Factors 31
   Gram-Schmidt Procedure 32
   Householder Transformation 33
   Givens Transformation 36

2.4 Linear Subspaces 37
   Basis 38
   Direct Subspaces 38
   Unicity 38
   Dimension, Rank, and Nullity 39
   Linear Equations 40
   Decomposition of $\mathbb{R}^n$ 41

2.5 Hermitian Matrices 42
   The Eigenvalues of a Hermitian Matrix Are Real 42
   The Eigenvectors of a Hermitian Matrix Are Orthogonal 42
   Hermitian Matrices Are Diagonalizable 42

2.6 Singular Value Decomposition 43
   Range and Null Space 45
   Low Rank Approximation 45
   Resolution (or Decomposition) of Identity 46

2.7 Projections, Rotations, and Pseudoinverses 46
   Projections 47
   Rotations 48
   Pseudoinverse 49
   Orthogonal Representations 49

2.8 Quadratic Forms 51

2.9 Matrix Inversion Formulas 52

2.10 The Multivariate Normal Distribution 55
   Characteristic Function 56
   The Bivariate Normal Distribution 57
   Linear Transformations 59
   Analysis and Synthesis 60
   Diagonalizing Transformations 61

2.11 Quadratic Forms in MVN Random Variables 62
   Quadratic Forms Using Projection Matrices 64
   Asymptotics 66

References and Comments 66
Problems 67

CHAPTER 3

Sufficiency and MVUB Estimators 77

3.1 Sufficiency 78
   Discrete Random Variables 79
   Continuous Random Variables 82
   Nonsingular Transformations and Sufficiency 85
Invariant Tests and Maximal Invariant Statistics  132
Uniformly Most Powerful Invariant Test  135
Reduction by Sufficiency and Invariance  135

4.9 Matched Filters (Normal)  136
4.10 CFAR Matched Filters (t)  140
4.11 Matched Subspace Filters ($\chi^2$)  145
4.12 CFAR Matched Subspace Filters (F)  148
A Comparative Summary and a Partial Ordering of Performance  149

4.13 Signal Design  153
  Signal Design for Detection  153
  Constrained Signal Design  154

4.14 Detectors for Gaussian Random Signals  157
  Likelihood Ratios and Quadratic Detectors  157
  Orthogonal Decomposition  158
  Distribution of Log Likelihood  158
  Rank Reduction  160

4.15 Linear Discriminant Functions  162
  Linear Discrimination  163
  An Extremization Problem  163
  Maximizing Divergence  164

References and Comments  166
Problems  167
Appendix: The $t$, $\chi^2$, and $F$ Distributions  174
  Central $\chi^2$  174
  Central $t$  175
  Central $F$  175
  Noncentral $\chi^2$  176
  Noncentral $t$  177
  Noncentral $F$  178
  Size and Power  178

CHAPTER 5
Bayes Detectors  179

5.1 Bayes Risk for Hypothesis Testing  180
  Loss  181
  Risk  181
  Bayes Risk  183

5.2 Bayes Tests of Simple Hypotheses  183

5.3 Minimax Tests  185
  Risk Set  185
  Bayes Tests  187
  Minimax and Maximin Tests  188
  Computing Minimax Tests  189
  Least Favorable Prior  190
5.4 Bayes Tests of Multiple Hypotheses 191
5.5 $M$-Orthogonal Signals 193
5.6 Composite Matched Filters and Associative Memories 198
  Application to Associative Memories 201
  Summary 202
5.7 Likelihood Ratios, Posterior Probabilities, and Odds 202
  Bayes Tests 203
5.8 Balanced Tests 204
References and Comments 205
Problems 206

CHAPTER 6

Maximum Likelihood Estimators 209

6.1 Maximum Likelihood Principle 210
  Random Parameters 213
6.2 Sufficiency 216
6.3 Invariance 217
6.4 The Fisher Matrix and the Cramer-Rao Bound 221
  Cramer-Rao Bound 222
  Concentration Ellipses 225
  Efficiency 226
  Cramer-Rao Bound for Functions of Parameters 229
  Numerical Maximum Likelihood and the Stochastic Fisher Matrix 230
6.5 Nuisance Parameters 231
6.6 Entropy, Likelihood, and Nonlinear Least Squares 233
  Entropy 233
  Likelihood 234
  Nonlinear Least Squares 234
  Comments 235
6.7 The Multivariate Normal Model 235
6.8 The Linear Statistical Model 238
6.9 Mode Identification in the Linear Statistical Model 239
  Maximum Likelihood Equations 240
  The Fisher Information Matrix 241
6.10 Maximum Likelihood Identification of ARMA Parameters 242
  Maximum Likelihood Equations 245
  The Projector $P(a)$ 245
  Interpretations 246
  KiSS 247
  The Fisher Information Matrix 248
  Mode Identification 250
6.11 Maximum Likelihood Identification of a Signal Subspace 252
CHAPTER 8
Minimum Mean-Squared Error Estimators 323

8.1 Conditional Expectation and Orthogonality 324
8.2 Minimum Mean-Squared Error Estimators 326
8.3 Linear Minimum Mean-Squared Error (LMMSE) Estimators 327
   Wiener-Hopf Equation 327
   Summary and Interpretations 328
8.4 Low-Rank Wiener Filter 330
8.5 Linear Prediction 331
8.6 The Kalman Filter 333
   Prediction 333
   Estimation 335
   Covariance Recursions 336
   The Kalman Recursions 337
8.7 Low-Rank Approximation of Random Vectors 337
   Interpretation 338
   Order Selection 339
8.8 Optimum Scalar Quantizers 339
   Scalar Quantizers 340
   Designing the Optimum Quantizer 342
8.9 Optimum Block Quantizers 346
8.10 Rank Reduction and Rate Distortion 349
References and Comments 351
Problems 352

CHAPTER 9
Least Squares 359

9.1 The Linear Model 360
   Interpretations 360
   The Normal Error Model 364
9.2 Least Squares Solutions 365
   Projections 365
   Signal and Orthogonal Subspaces 366
   Orthogonality 368
9.3 Structure of Subspaces in Least Squares Problems 371
9.4 Singular Value Decomposition 372
   Synthesis Representation 373
   Analysis Representations 373
9.5 Solving Least Squares Problems 374
   Cholesky-Factoring the Gram Matrix 374
   QR-Factoring the Model Matrix 375
   Singular Value Decomposition 376
9.6 Performance of Least Squares 377
   Posterior Model 377
   Performance 378

9.7 Goodness of Fit 378
   Statistician's Pythagorean Theorem 379

9.8 Improvement in SNR 379

9.9 Rank Reduction 380

9.10 Order Selection in the Linear Statistical Model 381

9.11 Sequential Least Squares 384

9.12 Weighted Least Squares 386

9.13 Constrained Least Squares 387
   Interpretations 388
   Condition Adjustment 388

9.14 Quadratic Minimization with Linear Constraints 389

9.15 Total Least Squares 392

9.16 Inverse Problems and Underdetermined Least Squares 393
   Characterizing the Class of Solutions 394
   Minimum-Norm Solution 395
   Reducing Rank 395
   Bayes 397
   Maximum Entropy 398
   Image Formation 398
   Newton-Raphson 400

9.17 Mode Identification in the Linear Statistical Model 401

9.18 Identification of Autoregressive Moving Average Signals and Systems 402

9.19 Linear Prediction and Prony's Method 405
   Modified Least Squares 405
   Linear Prediction 406
   Prony's Method 407

9.20 Least Squares Estimation of Structured Correlation Matrices 409
   Linear Structure 409
   Toeplitz Matrix 410
   Low-Rank Matrix 411
   Orthonormal Case 413
   More on the Orthonormal Case 413

References and Comments 415
Problems 415

CHAPTER 10

Linear Prediction 423

10.1 Stationary Time Series 424
   Wold Representation 426
   Kolmogorov Representation 427
   Filtering Interpretations 428
10.2 Stationary Prediction Theory 429
   Prediction Error Variance 430
   Prediction Error Variance and Poles and Zeros 431
   Spectral Flatness 433
   Filtering Interpretations 434
10.3 Maximum Entropy and Linear Prediction 436
10.4 Nonstationary Linear Prediction 438
   Synthesis 440
   Nonstationary Innovations Representation 441
   Analysis 441
   Nonstationary Predictor Representation 442
10.5 Fast Algorithms of the Levinson Type 442
   Interpretation 446
   Backward Form 447
   Filtering Interpretations 447
   AR Synthesis Lattice 448
   ARMA Lattice 450
10.6 Fast Algorithms of the Schur Type 451
10.7 Least Squares Theory of Linear Prediction 452
   QR Factors and Sliding Windows 456
   Summary and Interpretations 459
10.8 Lattice Algorithms 460
   Initialization 462
   Recursions for \( k_t \) 462
   Solving for \( \sigma_i^2 \) 462
   Algebraic Interpretations 463
   Lattice Interpretations 463
10.9 Prediction in Autoregressive Moving Average Time Series 464
   Stationary State-Space Representations 464
   Markovian State-Space Model 466
   Nonstationary (or Innovations) State-Space Representations 468
10.10 Fast Algorithms of the Morf-Sidhu-Kailath Type 470
10.11 Linear Prediction and Likelihood Formulas 472
10.12 Differential PCM 473
References and Comments 475
Problems 480

CHAPTER 11
Modal Analysis 483
11.1 Signal Model 484
   Modal Decomposition 485
   ARMA Impulse Response 486
   Linear Prediction 488
11.2 The Original Prony Method 489
  Filter Coefficients 489
  Modes 490
11.3 Least Squares Prony Method 491
  Choice of the Data Matrix 492
  Fast Algorithms 492
  Solving for the Mode Weights 493
  Solving for the Filter Coefficients 493
11.4 Maximum Likelihood and Exact Least Squares 493
  Compressed Likelihood 494
11.5 Total Least Squares 495
11.6 Principal Component Method (Tufts and Kumaresan) 496
  Information Criteria 496
  Overfitting 497
  Order Selection 498
  Real Arithmetic 498
11.7 Pisarenko and MUSIC Methods 499
  The Pisarenko Method 500
  MUSIC 502
11.8 Pencil Methods 503
11.9 A Frequency-Domain Version of Prony's Method (Kumaresan) 505
  Divided Differences 506
  Solving for $A(z)$ 506
  Solving for $B(z)$ 507

References 508

Index 515