PATTERN CLASSIFICATION AND SCENE ANALYSIS

RICHARD O. DUDA
PETER E. HART

Stanford Research Institute,
Menlo Park, California

A WILEY-INTERSCIENCE PUBLICATION

JOHN WILEY & SONS
New York • Chichester • Brisbane • Toronto • Singapore

CONTENTS

Part I PATTERN CLASSIFICATION

1 INTRODUCTION 1

1.1 Machine Perception 1
1.2 An Example 2
1.3 The Classification Model 4
1.4 The Descriptive Approach 5
1.5 Summary of the Book by Chapters 6
1.6 Bibliographical Remarks 7

2 BAYES DECISION THEORY 10

2.1 Introduction 10
2.2 Bayes Decision Theory—The Continuous Case 13
2.3 Two-Category Classification 15
2.4 Minimum-Error-Rate Classification 16
2.5 Classifiers, Discriminant Functions and Decision Surfaces 17
2.5.1 The Multicategory Case 17
2.5.2 The Two-Category Case 20
2.6 Error Probabilities and Integrals 20
2.7 The Normal Density 22
2.7.1 The Univariate Normal Density 22
2.7.2 The Multivariate Normal Density 23
2.8 Discriminant Functions for the Normal Density 24
2.8.1 Case 1: $\Sigma_i = \sigma^2 I$ 26
2.8.2 Case 2: $\Sigma_i = \Sigma$ 27
2.8.3 Case 3: Σ_i Arbitrary 30
2.9 Bayesian Decision Theory—The Discrete Case 31
2.10 Independent Binary Features 32
2.11 Compound Bayes Decision Theory and Context 34

2.12 Remarks 35
2.13 Bibliographical and Historical Remarks 36
Problems 39

3 PARAMETER ESTIMATION AND SUPERVISED LEARNING 44

3.1 Parameter Estimation and Supervised Learning 44
3.2 Maximum Likelihood Estimation 45
3.2.1 The General Principle 45
3.2.2 The Multivariate Normal Case: Unknown Mean 47
3.2.3 The General Multivariate Normal Case 48
3.3 The Bayes Classifier 49
3.3.1 The Class-Conditional Densities 50
3.3.2 The Parameter Distribution 51
3.4 Learning the Mean of a Normal Density 52
3.4.1 The Univariate Case: $p(\mu \mid \mathcal{X})$ 52
3.4.2 The Univariate Case: $p(x \mid \mathcal{X})$ 55
3.4.3 The Multivariate Case 55
3.5 General Bayesian Learning 57
3.6 Sufficient Statistics 59
3.7 Sufficient Statistics and the Exponential Family 62
3.8 Problems of Dimensionality 66
3.8.1 An Unexpected Problem 66
3.8.2 Estimating a Covariance Matrix 67
3.8.3 The Capacity of a Separating Plane 69
3.8.4 The Problem-Average Error Rate 70
3.9 Estimating the Error Rate 73
3.10 Bibliographical and Historical Remarks 76
Problems 80

4 NONPARAMETRIC TECHNIQUES 85

4.1 Introduction 85
4.2 Density Estimation 85
4.3 Parzen Windows 88
4.3.1 General Discussion 88
4.3.2 Convergence of the Mean 90
4.3.3 Convergence of the Variance 91
4.3.4 Two Examples 91
4.4 *k*-Nearest Neighbor Estimation 95

4.5 Estimation of A Posteriori Probabilities 97
4.6 The Nearest-Neighbor Rule 98
4.6.1 General Considerations 98
4.6.2 Convergence of the Nearest-Neighbor 99
4.6.3 Error Rate for the Nearest-Neighbor Rule 100
4.6.4 Error Bounds 101
4.7 The *k*-Nearest-Neighbor Rule 103
4.8 Approximations by Series Expansions 105
4.9 Approximations for the Binary Case 108
4.9.1 The Rademacher-Walsh Expansion 108
4.9.2 The Bahadur-Lazarsfeld Expansion 111
4.9.3 The Chow Expansion 113
4.10 Fisher's Linear Discriminant 114
4.11 Multiple Discriminant Analysis 118
4.12 Bibliographical and Historical Remarks 121
Problems 126

5 LINEAR DISCRIMINANT FUNCTIONS 130

5.1 Introduction 130
5.2 Linear Discriminant Functions and Decision Surfaces 131
5.2.1 The Two-Category Case 131
5.2.2 The Multicategory Case 132
5.3 Generalized Linear Discriminant Functions 134
5.4 The Two-Category Linearly-Separable Case 138
5.4.1 Geometry and Terminology 138
5.4.2 Gradient Descent Procedures 140
5.5 Minimizing the Perceptron Criterion Function 141
5.5.1 The Perceptron Criterion Function 141
5.5.2 Convergence Proof for Single-Sample Correction 142
5.5.3 Some Direct Generalizations 146
5.6 Relaxation Procedures 147
5.6.1 The Descent Algorithm 147
5.6.2 Convergence Proof 148
5.7 Nonseparable Behavior 149
5.8 Minimum Squared Error Procedures 151
5.8.1 Minimum Squared Error and the Pseudoinverse 151
5.8.2 Relation to Fisher's Linear Discriminant 152
5.8.3 Asymptotic Approximation to an Optimal Discriminant 154
5.8.4 The Widrow-Hoff Procedure 155
5.8.5 Stochastic Approximation Methods 156

5.9 The Ho-Kashyap Procedures 159
5.9.1 The Descent Procedure 159
5.9.2 Convergence Proof 161
5.9.3 Nonseparable Behavior 163
5.9.4 Some Related Procedures 163
5.10 Linear Programming Procedures 166
5.10.1 Linear Programming 166
5.10.2 The Linearly Separable Case 167
5.10.3 Minimizing the Perceptron Criterion Function 168
5.10.4 Remarks 169
5.11 The Method of Potential Functions 172
5.12 Multicategory Generalizations 174
5.12.1 Kesler's Construction 174
5.12.2 The Fixed-Increment Rule 176
5.12.3 Generalization for MSE Procedures 177
5.13 Bibliographical and Historical Remarks 179
Problems 186

6 UNSUPERVISED LEARNING AND CLUSTERING 189

6.1 Introduction 189
6.2 Mixture Densities and Identifiability 190
6.3 Maximum Likelihood Estimates 192
6.4 Application to Normal Mixtures 193
6.4.1 Case 1: Unknown Mean Vectors 194
6.4.2 An Example 195
6.4.3 Case 2: All Parameters Unknown 198
6.4.4 A Simple Approximate Procedure 201
6.5 Unsupervised Bayesian Learning 203
6.5.1 The Bayes Classifier 203
6.5.2 Learning the Parameter Vector 204
6.5.3 An Example 207
6.5.4 Decision-Directed Approximations 210
6.6 Data Description and Clustering 211
6.7 Similarity Measures 213
6.8 Criterion Functions for Clustering 217
6.8.1 The Sum-of-Squared-Error Criterion 217
6.8.2 Related Minimum Variance Criteria 219
6.8.3 Scattering Criteria 221
6.8.3.1 The Scatter Matrices 221

6.8.3.2 The Trace Criterion 222
6.8.3.3 The Determinant Criterion 222
6.8.3.4 Invariant Criteria 223
6.9 Iterative Optimization 225
6.10 Hierarchical Clustering 228
6.10.1 Definitions 228
6.10.2 Agglomerative Hierarchical Clustering 230
6.10.2.1 The Nearest-Neighbor Algorithm 233
6.10.2.2 The Furthest-Neighbor Algorithm 233
6.10.2.3 Compromises 235
6.10.3 Stepwise-Optimal Hierarchical Clustering 235
6.10.4 Hierarchical Clustering and Induced Metrics 236
6.11 Graph Theoretic Methods 237
6.12 The Problem of Validity 239
6.13 Low-Dimensional Representations and Multidimensional Scaling 243
6.14 Clustering and Dimensionality Reduction 246
6.15 Bibliographical and Historical Remarks 248
Problems 256

Part II SCENE ANALYSIS

7 REPRESENTATION AND INITIAL SIMPLIFICATIONS 263

7.1 Introduction 263
7.2 Representations 264
7.3 Spatial Differentiation 267
7.4 Spatial Smoothing 272
7.5 Template Matching 276
7.5.1 Template Matching—Metric Interpretation 276
7.5.2 Template Matching—Statistical Interpretation 282
7.6 Region Analysis 284
7.6.1 Basic Concepts 284
7.6.2 Extensions 288
7.7 Contour Following 290
7.8 Bibliographical and Historical Remarks 293
Problems 297

8 THE SPATIAL FREQUENCY DOMAIN 298

8.1 Introduction 298
8.2 The Sampling Theorem 302
8.3 Template Matching and the Convolution Theorem 305
8.4 Spatial Filtering 308
8.5 Mean Square Estimation 318
8.6 Bibliographical and Historical Remarks 322
Problems 325

9 DESCRIPTIONS OF LINE AND SHAPE 327

9.1 Introduction 327
9.2 Line Description 328
9.2.1 Minimum-Squared-Error Line Fitting 328
9.2.2 Eigenvector Line Fitting 332
9.2.3 Line Fitting by Clustering 335
9.2.4 Line Segmentation 337
9.2.5 Chain Encoding 339
9.3 Shape Description 341
9.3.1 Topological Properties 342
9.3.2 Linear Properties 345
9.3.3 Metric Properties 348
9.3.4 Descriptions Based on Irregularities 352
9.3.5 The Skeleton of a Figure 356
9.3.6 Analytic Descriptions of Shape 362
9.3.7 Integral Geometric Descriptions 367
9.4 Bibliographical and Historical Remarks 372
Problems 377

10 PERSPECTIVE TRANSFORMATIONS 379

10.1 Introduction 379
10.2 Modelling Picture Taking 380
10.3 The Perspective Transformation in Homogeneous Coordinates 382
10.4 Perspective Transformations With Two Reference Frames 386
10.5 Illustrative Applications 392
10.5.1 Camera Calibration 392
10.5.2 Object Location 393
10.5.3 Vertical Lines: Perspective Distortion 394
10.5.4 Horizontal Lines and Vanishing Points 396

10.6 Stereoscopic Perception 398
10.7 Bibliographical and Historical Remarks 401
Problems 404

11 PROJECTIVE INVARIANTS 405

11.1 Introduction 405
11.2 The Cross Ratio 407
11.3 Two-Dimensional Projective Coordinates 411
11.4 The Inter-Lens Line 414
11.5 An Orthogonal Projection Approximation 418
11.6 Object Reconstruction 421
11.7 Bibliographical and Historical Remarks 422
Problems 424

12 DESCRIPTIVE METHODS IN SCENE ANALYSIS 425

12.1 Introduction 425
12.2 Descriptive Formalisms 426
12.2.1 Syntactic Descriptions 426
12.2.2 Relational Graphs 434
12.3 Three-Dimensional Models 436
12.4 The Analysis of Polyhedra 441
12.4.1 Line Semantics 442
12.4.2 Grouping Regions into Objects 449
12.4.3 Monocular Determination of Three-Dimensional Structure 456
12.5 Bibliographical and Historical Remarks 462
Problems 465

AUTHOR INDEX 467

SUBJECT INDEX 472