

SIMON J. D. PRINCE



COMPUTER VISION

MODELS, LEARNING,
AND INFERENCE

Contents

Acknowledgments	<i>page</i> xiii
Foreword by Andrew Fitzgibbon	xv
Preface	xvii
1 Introduction	1
Organization of the book	3
Other books	5
I Probability	
2 Introduction to probability	9
2.1 Random variables	9
2.2 Joint probability	10
2.3 Marginalization	10
2.4 Conditional probability	12
2.5 Bayes' rule	13
2.6 Independence	14
2.7 Expectation	14
3 Common probability distributions	17
3.1 Bernoulli distribution	18
3.2 Beta distribution	19
3.3 Categorical distribution	19
3.4 Dirichlet distribution	20
3.5 Univariate normal distribution	21
3.6 Normal-scaled inverse gamma distribution	21
3.7 Multivariate normal distribution	22
3.8 Normal inverse Wishart distribution	23
3.9 Conjugacy	24
4 Fitting probability models	28
4.1 Maximum likelihood	28
4.2 Maximum a posteriori	28
4.3 The Bayesian approach	29
4.4 Worked example 1: Univariate normal	30
4.5 Worked example 2: Categorical distribution	38

5	The normal distribution	44
5.1	Types of covariance matrix	44
5.2	Decomposition of covariance	45
5.3	Linear transformations of variables	47
5.4	Marginal distributions	47
5.5	Conditional distributions	48
5.6	Product of two normals	48
5.7	Change of variable	50
II	Machine learning for machine vision	
6	Learning and inference in vision	55
6.1	Computer vision problems	55
6.2	Types of model	56
6.3	Example 1: Regression	57
6.4	Example 2: Binary classification	60
6.5	Which type of model should we use?	63
6.6	Applications	64
7	Modeling complex data densities	71
7.1	Normal classification model	71
7.2	Hidden variables	74
7.3	Expectation maximization	75
7.4	Mixture of Gaussians	77
7.5	The t -distribution	82
7.6	Factor analysis	88
7.7	Combining models	93
7.8	Expectation maximization in detail	94
7.9	Applications	99
8	Regression models	108
8.1	Linear regression	108
8.2	Bayesian linear regression	111
8.3	Nonlinear regression	114
8.4	Kernels and the kernel trick	118
8.5	Gaussian process regression	119
8.6	Sparse linear regression	120
8.7	Dual linear regression	124
8.8	Relevance vector regression	127
8.9	Regression to multivariate data	128
8.10	Applications	128
9	Classification models	133
9.1	Logistic regression	133
9.2	Bayesian logistic regression	138
9.3	Nonlinear logistic regression	142
9.4	Dual logistic regression	144
9.5	Kernel logistic regression	146

9.6	Relevance vector classification	147
9.7	Incremental fitting and boosting	150
9.8	Classification trees	153
9.9	Multiclass logistic regression	156
9.10	Random trees, forests, and ferns	158
9.11	Relation to non-probabilistic models	159
9.12	Applications	160

III Connecting local models

10	Graphical models	173
10.1	Conditional independence	173
10.2	Directed graphical models	175
10.3	Undirected graphical models	178
10.4	Comparing directed and undirected graphical models	181
10.5	Graphical models in computer vision	181
10.6	Inference in models with many unknowns	184
10.7	Drawing samples	186
10.8	Learning	188
11	Models for chains and trees	195
11.1	Models for chains	196
11.2	MAP inference for chains	198
11.3	MAP inference for trees	202
11.4	Marginal posterior inference for chains	205
11.5	Marginal posterior inference for trees	211
11.6	Learning in chains and trees	212
11.7	Beyond chains and trees	213
11.8	Applications	216
12	Models for grids	227
12.1	Markov random fields	228
12.2	MAP inference for binary pairwise MRFs	231
12.3	MAP inference for multilabel pairwise MRFs	239
12.4	Multilabel MRFs with non-convex potentials	244
12.5	Conditional random fields	247
12.6	Higher order models	250
12.7	Directed models for grids	250
12.8	Applications	251

IV Preprocessing

13	Image preprocessing and feature extraction	269
13.1	Per-pixel transformations	269
13.2	Edges, corners, and interest points	279
13.3	Descriptors	283
13.4	Dimensionality reduction	287

V Models for geometry

14 The pinhole camera	297
14.1 The pinhole camera	297
14.2 Three geometric problems	304
14.3 Homogeneous coordinates	306
14.4 Learning extrinsic parameters	309
14.5 Learning intrinsic parameters	311
14.6 Inferring three-dimensional world points	312
14.7 Applications	314
15 Models for transformations	323
15.1 Two-dimensional transformation models	323
15.2 Learning in transformation models	330
15.3 Inference in transformation models	334
15.4 Three geometric problems for planes	335
15.5 Transformations between images	339
15.6 Robust learning of transformations	342
15.7 Applications	347
16 Multiple cameras	354
16.1 Two-view geometry	355
16.2 The essential matrix	357
16.3 The fundamental matrix	361
16.4 Two-view reconstruction pipeline	364
16.5 Rectification	368
16.6 Multiview reconstruction	372
16.7 Applications	376

VI Models for vision

17 Models for shape	387
17.1 Shape and its representation	388
17.2 Snakes	389
17.3 Shape templates	393
17.4 Statistical shape models	396
17.5 Subspace shape models	399
17.6 Three-dimensional shape models	405
17.7 Statistical models for shape and appearance	405
17.8 Non-Gaussian statistical shape models	410
17.9 Articulated models	414
17.10 Applications	415
18 Models for style and identity	424
18.1 Subspace identity model	427
18.2 Probabilistic linear discriminant analysis	433
18.3 Nonlinear identity models	437
18.4 Asymmetric bilinear models	438

18.5	Symmetric bilinear and multilinear models	443
18.6	Applications	446
19	Temporal models	453
19.1	Temporal estimation framework	453
19.2	Kalman filter	455
19.3	Extended Kalman filter	466
19.4	Unscented Kalman filter	467
19.5	Particle filtering	472
19.6	Applications	476
20	Models for visual words	483
20.1	Images as collections of visual words	483
20.2	Bag of words	484
20.3	Latent Dirichlet allocation	487
20.4	Single author–topic model	493
20.5	Constellation models	495
20.6	Scene models	499
20.7	Applications	500
VII Appendices		
A	Notation	507
B	Optimization	509
B.1	Problem statement	509
B.2	Choosing a search direction	511
B.3	Line search	515
B.4	Reparameterization	516
C	Linear algebra	519
C.1	Vectors	519
C.2	Matrices	520
C.3	Tensors	522
C.4	Linear transformations	522
C.5	Singular value decomposition	522
C.6	Matrix calculus	527
C.7	Common problems	528
C.8	Tricks for inverting large matrices	530
Bibliography		533
Index		567