

PATTERN THEORY: FROM
REPRESENTATION
TO INFERENCE

Ulf Grenander and Michael I. Miller

OXFORD
UNIVERSITY PRESS

CONTENTS

1	Introduction	1
1.1	Organization	3
2	The Bayes Paradigm, Estimation and Information Measures	5
2.1	Bayes Posterior Distribution	5
2.1.1	Minimum Risk Estimation	6
2.1.2	Information Measures	7
2.2	Mathematical Preliminaries	8
2.2.1	Probability Spaces, Random Variables, Distributions, Densities, and Expectation	8
2.2.2	Transformations of Variables	10
2.2.3	The Multivariate Normal Distribution	10
2.2.4	Characteristic Function	11
2.3	Minimum Risk Hypothesis Testing on Discrete Spaces	12
2.3.1	Minimum Probability of Error via Maximum A Posteriori Hypothesis Testing	13
2.3.2	Neyman–Pearson and the Optimality of the Likelihood Ratio Test	14
2.4	Minimum Mean-Squared Error Risk Estimation in Vector Spaces	16
2.4.1	Normed Linear and Hilbert Spaces	17
2.4.2	Least-Squares Estimation	20
2.4.3	Conditional Mean Estimation and Gaussian Processes	22
2.5	The Fisher Information of Estimators	24
2.6	Maximum-Likelihood and its consistency	26
2.6.1	Consistency via Uniform Convergence of Empirical Log-likelihood	27
2.6.2	Asymptotic Normality and \sqrt{n} Convergence Rate of the MLE	28
2.7	Complete–Incomplete Data Problems and the EM Algorithm	30
2.8	Hypothesis Testing and Model Complexity	38
2.8.1	Model-Order Estimation and the $d/2$ log Sample-Size Complexity	38
2.8.2	The Gaussian Case is Special	41
2.8.3	Model Complexity and the Gaussian Case	42
2.9	Building Probability Models via the Principle of Maximum Entropy	43
2.9.1	Principle of Maximum Entropy	44
2.9.2	Maximum Entropy Models	45
2.9.3	Conditional Distributions are Maximum Entropy	47
3	Probabilistic Directed Acyclic Graphs and Their Entropies	49
3.1	Directed Acyclic Graphs (DAGs)	49
3.2	Probabilities on Directed Acyclic Graphs (PDAGs)	51
3.3	Finite State Markov Chains	54
3.4	Multi-type Branching Processes	56
3.4.1	The Branching Matrix	59
3.4.2	The Moment-Generating Function	60
3.5	Extinction for Finite-State Markov Chains and Branching Processes	62
3.5.1	Extinction in Markov Chains	62
3.5.2	Extinction in Branching Processes	63
3.6	Entropies of Directed Acyclic Graphs	64
3.7	Combinatorics of Independent, Identically Distributed Strings via the Asymptotic Equipartition Theorem	65

3.8	Entropy and Combinatorics of Markov Chains	66
3.9	Entropies of Branching Processes	68
3.9.1	Tree Structure of Multi-Type Branching Processes	69
3.9.2	Entropies of Sub-Critical, Critical, and Super-Critical Processes	70
3.9.3	Typical Trees and the Equipartition Theorem	71
3.10	Formal Languages and Stochastic Grammars	74
3.11	DAGs for Natural Language Modelling	81
3.11.1	Markov Chains and m -Grams	81
3.11.2	Context-Free Models	82
3.11.3	Hierarchical Directed Acyclic Graph Model	84
3.12	EM Algorithms for Parameter Estimation in Hidden Markov Models	87
3.12.1	MAP Decoding of the Hidden State Sequence	88
3.12.2	ML Estimation of HMM parameters via EM Forward/Backward Algorithm	89
3.13	EM Algorithms for Parameter Estimation in Natural Language Models	92
3.13.1	EM Algorithm for Context-Free Chomsky Normal Form	93
3.13.2	General Context-Free Grammars and the Trellis Algorithm of Kupiec	94
4	Markov Random Fields on Undirected Graphs	95
4.1	Undirected Graphs	95
4.2	Markov Random Fields	96
4.3	Gibbs Random Fields	101
4.4	The Splitting Property of Gibbs Distributions	104
4.5	Bayesian Texture Segmentation: The log-Normalizer Problem	110
4.5.1	The Gibbs Partition Function Problem	110
4.6	Maximum-Entropy Texture Representation	112
4.6.1	Empirical Maximum Entropy Texture Coding	113
4.7	Stationary Gibbs Random Fields	116
4.7.1	The Dobrushin/Lanford/Ruelle Definition	116
4.7.2	Gibbs Distributions Exhibit Multiple Laws with the Same Interactions (Phase Transitions): The Ising Model at Low Temperature	117
4.8	1D Random Fields are Markov Chains	119
4.9	Markov Chains Have a Unique Gibbs Distribution	120
4.10	Entropy of Stationary Gibbs Fields	121
5	Gaussian Random Fields on Undirected Graphs	123
5.1	Gaussian Random Fields	123
5.2	Difference Operators and Adjoints	124
5.3	Gaussian Fields Induced via Difference Operators	126
5.4	Stationary Gaussian Processes on \mathbb{Z}^d and their Spectrum	133
5.5	Cyclo-Stationary Gaussian Processes and their Spectrum	134
5.6	The log-Determinant Covariance and the Asymptotic Normalizer	137
5.6.1	Asymptotics of the Gaussian processes and their Covariance	138
5.6.2	The Asymptotic Covariance and log-Normalizer	142
5.7	The Entropy Rates of the Stationary Process	142
5.7.1	Burg's Maximum Entropy Auto-regressive Processes on \mathbb{Z}^d	143
5.8	Generalized Auto-Regressive Image Modelling via Maximum-Likelihood Estimation	144
5.8.1	Anisotropic Textures	147

6	The Canonical Representations of General Pattern Theory	154
6.1	The Generators, Configurations, and Regularity of Patterns	154
6.2	The Generators of Formal Languages and Grammars	158
6.3	Graph Transformations	162
6.4	The Canonical Representation of Patterns: DAGs, MRFs, Gaussian Random Fields	166
6.4.1	Directed Acyclic Graphs	167
6.4.2	Markov Random Fields	169
6.4.3	Gaussian Random Fields: Generators induced via difference operators	170
7	Matrix Group Actions Transforming Patterns	174
7.1	Groups Transforming Configurations	174
7.1.1	Similarity Groups	174
7.1.2	Group Actions Defining Equivalence	175
7.1.3	Groups Actions on Generators and Deformable Templates	177
7.2	The Matrix Groups	177
7.2.1	Linear Matrix and Affine Groups of Transformation	177
7.2.2	Matrix groups acting on \mathbb{R}^d	179
7.3	Transformations Constructed from Products of Groups	181
7.4	Random Regularity on the Similarities	184
7.5	Curves as Submanifolds and the Frenet Frame	190
7.6	2D Surfaces in \mathbb{R}^3 and the Shape Operator	195
7.6.1	The Shape Operator	196
7.7	Fitting Quadratic Charts and Curvatures on Surfaces	198
7.7.1	Gaussian and Mean Curvature	198
7.7.2	Second Order Quadratic Charts	200
7.7.3	Isosurface Algorithm	201
7.8	Ridge Curves and Crest Lines	205
7.8.1	Definition of Sulcus, Gyrus, and Geodesic Curves on Triangulated Graphs	205
7.8.2	Dynamic Programming	207
7.9	Bijections and Smooth Mappings for Coordinatizing Manifolds via Local Coordinates	210
8	Manifolds, Active Models, and Deformable Templates	214
8.1	Manifolds as Generators, Tangent Spaces, and Vector Fields	214
8.1.1	Manifolds	214
8.1.2	Tangent Spaces	215
8.1.3	Vector Fields on M	217
8.1.4	Curves and the Tangent Space	218
8.2	Smooth Mappings, the Jacobian, and Diffeomorphisms	219
8.2.1	Smooth Mappings and the Jacobian	219
8.2.2	The Jacobian and Local Diffeomorphic Properties	221
8.3	Matrix Groups are Diffeomorphisms which are a Smooth Manifold	222
8.3.1	Diffeomorphisms	222
8.3.2	Matrix Group Actions are Diffeomorphisms on the Background Space	223
8.3.3	The Matrix Groups are Smooth Manifolds (Lie Groups)	224
8.4	Active Models and Deformable Templates as Immersions	226
8.4.1	Snakes and Active Contours	226
8.4.2	Deforming Closed Contours in the Plane	226
8.4.3	Normal Deformable Surfaces	227
8.5	Activating Shapes in Deformable Models	229
8.5.1	Likelihood of Shapes Partitioning Image	229
8.5.2	A General Calculus for Shape Activation	229

8.5.3	Active Closed Contours in \mathbb{R}^2	232
8.5.4	Active Unclosed Snakes and Roads	234
8.5.5	Normal Deformation of Circles and Spheres	236
8.5.6	Active Deformable Spheres	236
8.6	Level Set Active Contour Models	237
8.7	Gaussian Random Field Models for Active Shapes	240
9	Second Order and Gaussian Fields	244
9.1	Second Order Processes (SOP) and the Hilbert Space of Random Variables	244
9.1.1	Measurability, Separability, Continuity	244
9.1.2	Hilbert space of random variables	247
9.1.3	Covariance and Second Order Properties	249
9.1.4	Quadratic Mean Continuity and Integration	251
9.2	Orthogonal Process Representations on Bounded Domains	252
9.2.1	Compact Operators and Covariances	253
9.2.2	Orthogonal Representations for Random Processes and Fields	257
9.2.3	Stationary Periodic Processes and Fields on Bounded Domains	258
9.3	Gaussian Fields on the Continuum	262
9.4	Sobolev Spaces, Green's Functions, and Reproducing Kernel Hilbert Spaces	264
9.4.1	Reproducing Kernel Hilbert Spaces	265
9.4.2	Sobolev Normed Spaces	266
9.4.3	Relation to Green's Functions	267
9.4.4	Gradient and Laplacian Induced Green's Kernels	267
9.5	Gaussian Processes Induced via Linear Differential Operators	271
9.6	Gaussian Fields in the Unit Cube	274
9.6.1	Maximum Likelihood Estimation of the Fields: Generalized ARMA Modelling	278
9.6.2	Small Deformation Vector Fields Models in the Plane and Cube	280
9.7	Discrete Lattices and Reachability of Cyclo-Stationary Spectra	283
9.8	Stationary Processes on the Sphere	285
9.8.1	Laplacian Operator Induced Gaussian Fields on the Sphere	289
9.9	Gaussian Random Fields on an Arbitrary Smooth Surface	293
9.9.1	Laplace-Beltrami Operator with Neumann Boundary Conditions	293
9.9.2	Smoothing an Arbitrary Function on Manifolds by Orthonormal Bases of the Laplace-Beltrami Operator	297
9.10	Sample Path Properties and Continuity	299
9.11	Gaussian Random Fields as Prior Distributions in Point Process Image Reconstruction	303
9.11.1	The Need for Regularization in Image Reconstruction	304
9.11.2	Smoothness and Gaussian Priors	304
9.11.3	Good's Roughness as a Gaussian Prior	305
9.11.4	Exponential Spline Smoothing via Good's Roughness	306
9.12	Non-Compact Operators and Orthogonal Representations	309
9.12.1	Cramer Decomposition for Stationary Processes	311
9.12.2	Orthogonal Scale Representation	312
10	Metrics Spaces for the Matrix Groups	316
10.1	Riemannian Manifolds as Metric Spaces	316
10.1.1	Metric Spaces and Smooth Manifolds	316
10.1.2	Riemannian Manifold, Geodesic Metric, and Minimum Energy	317
10.2	Vector Spaces as Metric Spaces	319
10.3	Coordinate Frames on the Matrix Groups and the Exponential Map	320

10.3.1	Left and Right Group Action	320
10.3.2	The Coordinate Frames	321
10.3.3	Local Optimization via Directional Derivatives and the Exponential Map	323
10.4	Metric Space Structure for the Linear Matrix Groups	324
10.4.1	Geodesics in the Matrix Groups	324
10.5	Conservation of Momentum and Geodesic Evolution of the Matrix Groups via the Tangent at the Identity	326
10.6	Metrics in the Matrix Groups	327
10.7	Viewing the Matrix Groups in Extrinsic Euclidean Coordinates	329
10.7.1	The Frobenius Metric	329
10.7.2	Comparing intrinsic and extrinsic metrics in $SO(2,3)$	330
11	Metrics Spaces for the Infinite Dimensional Diffeomorphisms	332
11.1	Lagrangian and Eulerian Generation of Diffeomorphisms	332
11.1.1	On Conditions for Generating Flows of Diffeomorphisms	333
11.1.2	Modeling via Differential Operators and the Reproducing Kernel Hilbert Space	335
11.2	The Metric on the Space of Diffeomorphisms	336
11.3	Momentum Conservation for Geodesics	338
11.4	Conservation of Momentum for Diffeomorphism Splines Specified on Sparse Landmark Points	340
11.4.1	An ODE for Diffeomorphic Landmark Mapping	343
12	Metrics on Photometric and Geometric Deformable Templates	346
12.1	Metrics on Dense Deformable Templates: Geometric Groups Acting on Images	346
12.1.1	Group Actions on the Images	346
12.1.2	Invariant Metric Distances	347
12.2	The Diffeomorphism Metric for the Image Orbit	349
12.3	Normal Momentum Motion for Geodesic Connection Via Inexact Matching	350
12.4	Normal Momentum Motion for Temporal Sequences	354
12.5	Metric Distances Between Orbits Defined Through Invariance of the Metric	356
12.6	Finite Dimensional Landmarked Shape Spaces	357
12.6.1	The Euclidean Metric	357
12.6.2	Kendall's Similitude Invariant Distance	359
12.7	The Diffeomorphism Metric and Diffeomorphism Splines on Landmark Shapes	361
12.7.1	Small Deformation Splines	361
12.8	The Deformable Template: Orbits of Photometric and Geometric Variation	365
12.8.1	Metric Spaces for Photometric Variability	365
12.8.2	The Metrics Induced via Photometric and Geometric Flow	366
12.9	The Euler Equations for Photometric and Geometric Variation	369
12.10	Metrics between Orbits of the Special Euclidean Group	373
12.11	The Matrix Groups (Euclidean and Affine Motions)	374
12.11.1	Computing the Affine Motions	376
13	Estimation Bounds for Automated Object Recognition	378
13.1	The Communications Model for Image Transmission	378
13.1.1	The Source Model: Objects Under Matrix Group Actions	379
13.1.2	The Sensing Models: Projective Transformations in Noise	379
13.1.3	The Likelihood and Posterior	379

13.2	Conditional Mean Minimum Risk Estimation	381
13.2.1	Metrics (Risk) on the Matrix Groups	381
13.2.2	Conditional Mean Minimum Risk Estimators	382
13.2.3	Computation of the HSE for $SE(2,3)$	384
13.2.4	Discrete integration on $SO(3)$	385
13.3	MMSE Estimators for Projective Imagery Models	385
13.3.1	3D to 2D Projections in Gaussian Noise	385
13.3.2	3D to 2D Synthetic Aperture Radar Imaging	389
13.3.3	3D to 2D LADAR Imaging	392
13.3.4	3D to 2D Poisson Projection Model	393
13.3.5	3D to 1D Projections	395
13.3.6	3D(2D) to 3D(2D) Medical Imaging Registration	397
13.4	Parameter Estimation and Fisher Information	398
13.5	Bayesian Fusion of Information	402
13.6	Asymptotic Consistency of Inference and Symmetry Groups	405
13.6.1	Consistency	405
13.6.2	Symmetry Groups and Sensor Symmetry	406
13.7	Hypothesis Testing and Asymptotic Error-Exponents	407
13.7.1	Analytical Representations of the Error Probabilities and the Bayesian Information Criterion	408
13.7.2	m-ary Multiple Hypotheses	412
14	Estimation on Metric Spaces with Photometric Variation	414
14.1	The Deformable Template: Orbits of Signature and Geometric Variation	414
14.1.1	The Robust Deformable Templates	414
14.1.2	The Metric Space of the Robust Deformable Template	415
14.2	Empirical Covariance of Photometric Variability via Principle Components	416
14.2.1	Signatures as a Gaussian Random Field Constructed from Principle Components	417
14.2.2	Algorithm for Empirical Construction of Bases	418
14.3	Estimation of Parameters on the Conditionally Gaussian Random Field Models	422
14.4	Estimation of Pose by Integrating Out EigenSignatures	424
14.4.1	Bayes Integration	427
14.5	Multiple Modality Signature Registration	429
14.6	Models for Clutter: The Transported Generator Model	431
14.6.1	Characteristic Functions and Cumulants	432
14.7	Robust Deformable Templates for Natural Clutter	438
14.7.1	The Euclidean Metric	439
14.7.2	Metric Space Norms for Clutter	439
14.7.3	Computational Scheme	442
14.7.4	Empirical Construction of the Metric from Rendered Images	444
14.8	Target detection/identification in EO imagery	445
15	Information Bounds for Automated Object Recognition	447
15.1	Mutual Information for Sensor Systems	447
15.1.1	Quantifying Multiple-Sensor Information Gain Via Mutual Information	447
15.1.2	Quantifying Information Loss with Model Uncertainty	449
15.1.3	Asymptotic Approximation of Information Measures	452

15.2	Rate-Distortion Theory	456
15.2.1	The Rate-Distortion Problem	456
15.3	The Blahut Algorithm	457
15.4	The Remote Rate Distortion Problem	459
15.4.1	Blahut Algorithm extended	460
15.5	Output Symbol Distribution	465
16	Computational Anatomy: Shape, Growth and Atrophy Comparison via Diffeomorphisms	468
16.1	Computational Anatomy	468
16.1.1	Diffeomorphic Study of Anatomical Submanifolds	469
16.2	The Anatomical Source Model of CA	470
16.2.1	Group Actions for the Anatomical Source Model	472
16.2.2	The Data Channel Model	473
16.3	Normal Momentum Motion for Large Deformation Metric Mapping (LDDMM) for Growth and Atrophy	474
16.4	Christensen Non-Geodesic Mapping Algorithm	478
16.5	Extrinsic Mapping of Surface and Volume Submanifolds	480
16.5.1	Diffeomorphic Mapping of the Face	481
16.5.2	Diffeomorphic Mapping of Brain Submanifolds	481
16.5.3	Extrinsic Mapping of Subvolumes for Automated Segmentation	481
16.5.4	Metric Mapping of Cortical Atlases	483
16.6	Heart Mapping and Diffusion Tensor Magnetic Resonance Imaging	484
16.7	Vector Fields for Growth	488
16.7.1	Growth from Landmarked Shape Spaces	488
17	Computational Anatomy: Hypothesis Testing on Disease	494
17.1	Statistics Analysis for Shape Spaces	494
17.2	Gaussian Random Fields	495
17.2.1	Empirical Estimation of Random Variables	496
17.3	Shape Representation of the Anatomical Orbit Under Large Deformation Diffeomorphisms	496
17.3.1	Principal Component Selection of the Basis from Empirical Observations	497
17.4	The Momentum of Landmarked Shape Spaces	498
17.4.1	Geodesic evolution equations for landmarks	498
17.4.2	Small Deformation PCA Versus Large Deformation PCA	499
17.5	The Small Deformation Setting	502
17.6	Small Deformation Gaussian Fields on Surface Submanifolds	502
17.7	Disease Testing of Automorphic Pathology	503
17.7.1	Hypothesis Testing on Disease in the Small Noise Limit	503
17.7.2	Statistical Testing	505
17.8	Distribution Free Testing	510
17.9	Heteromorphic Tumors	511
18	Markov Processes and Random Sampling	514
18.1	Markov Jump Processes	514
18.1.1	Jump Processes	515
18.2	Random Sampling and Stochastic Inference	516
18.2.1	Stationary or Invariant Measures	517
18.2.2	Generator for Markov Jump Processes	519
18.2.3	Jump Process Simulation	520
18.2.4	Metropolis–Hastings Algorithm	521

18.3	Diffusion Processes for Simulation	523
18.3.1	Generators of 1D Diffusions	525
18.3.2	Diffusions and SDEs for Sampling	527
18.4	Jump-Diffusion Inference on Countable Unions of Spaces	528
18.4.1	The Basic Problem	529
19	Jump Diffusion Inference in Complex Scenes	532
19.1	Recognition of Ground Vehicles	533
19.1.1	CAD Models and the Parameter Space	533
19.1.2	The FLIR Sensor Model	534
19.2	Jump Diffusion for Sampling the Target Recognition Posterior	536
19.2.1	The Posterior distribution	536
19.2.2	The Jump Diffusion Algorithms	536
19.2.3	Jumps via Gibbs' Sampling	539
19.2.4	Jumps via Metropolis–Hastings Acceptance/Rejection	541
19.3	Experimental Results for FLIR and LADAR	543
19.3.1	Detection and Removal of Objects	543
19.3.2	Identification	543
19.3.3	Pose and Identification	544
19.3.4	Identification and recognition via High Resolution Radar (HRR)	546
19.3.5	The Dynamics of Pose Estimation via the Jump–Diffusion Process	546
19.3.6	LADAR Recognition	548
19.4	Powerful Prior Dynamics for Airplane Tracking	549
19.4.1	The Euler-Equations Inducing the Prior on Airplane Dynamics	550
19.4.2	Detection of Airframes	552
19.4.3	Pruning via the Prior distribution	552
19.5	Deformable Organelles: Mitochondria and Membranes	553
19.5.1	The Parameter Space for Contour Models	553
19.5.2	Stationary Gaussian Contour Model	554
19.5.3	The Electron Micrograph Data Model: Conditional Gaussian Random Fields	555
19.6	Jump–Diffusion for Mitochondria	556
19.6.1	The jump parameters	557
19.6.2	Computing gradients for the drifts	557
19.6.3	Jump Diffusion for Mitochondria Detection and Deformation	558
19.6.4	Pseudolikelihood for Deformation	560
	References	563
	Index	581