TABLE OF CONTENTS

PREFACE XIII

PART ONE: DISCRETE INVERSE PROBLEMS

1: The general discrete inverse problem 1

1.1: Model Space and Data space 2

1.1.1: The model space 2
1.1.2: The data space 7
1.1.3: The joint space $\boldsymbol{D} \times \boldsymbol{M}$ 8
1.1.4: Notations 9

1.2: States of Information 12

1.2.1: The mathematical concept of probability 12
1.2.2: The interpretation of a probability 16
1.2.3: The state of Perfect Knowledge 21
1.2.4: The state of Total Ignorance (or the reference state of information) 21
1.2.5: Shannon's measure of Information Content 28
1.2.6: The Combination of States of Information 29

1.3: Information obtained from physical theories (Solving the Forward Problem) 33

1.4: Information obtained from measurements, and a priori information 37

1.4.1: Results of the measurements 37
1.4.2: A priori information on model parameters 41
1.4.3: Joint prior information in the $\boldsymbol{D} \times \boldsymbol{M}$ space 51

1.5: Defining the solution of the Inverse Problem 52

1.5.1: Combination of experimental, a priori, and theoretical Information 52
1.5.2: Resolution of Inverse Problems 53

1.5.3: Some special cases 55

1.6: Using the solution of the Inverse Problem 61

1.6.1: Describing the a posteriori information in the model space 61
1.6.2: Analysis of error and resolution 63
1.6.3: Analytic solutions 64
1.6.4: Systematic exploration of the model space 65
1.6.5: Monte Carlo Methods 65
1.6.6: Computation of the maximum likelihood point 65

1.7: Special cases 67

1.7.1: The Gaussian Hypothesis (least-squares criterion). Case $\mathbf{d} = \mathbf{g}(\mathbf{m})$ 67
1.7.2: The Gaussian Hypothesis (least-squares criterion). Case $\mathbf{f}(\mathbf{d},\mathbf{m}) = \mathbf{0}$ 77
1.7.3: Generalized Gaussian (least-absolute values criterion, minimax criterion) 83

Box 1.1: Central Estimators and Estimators of Dispersion 18
Box 1.2: Generalized Gaussian 26
Box 1.3: The Log-normal probability density 48
Box 1.4: Solving Inverse Problems using the Bayesian paradigm . . . 58

Problems for Chapter 1 85

2: The trial and error method 163

Problems for Chapter 2 166

3: Monte Carlo methods 167

3.1: Search for the domain of admissible models 168

3.2: Nonlinear computation of variances and covariances 170

3.3: Simulated annealing 176

Box 3.1: The Monte Carlo method of numerical integration 174

4: **The least-squares (ℓ_2-norm) criterion** 187

4.1: Introducing least-squares 188

4.2: Methods of resolution (I) 191

4.2.1: Systematic exploration of the model space 191
4.2.2: Trial and error 191
4.2.3: Relaxation 191
4.2.4: Monte Carlo methods 192
4.2.5: The Gauss-Newton method 192

4.3: Analysis of error and resolution 196

4.3.1: Computation of the posterior covariance operator 196
4.3.2: Interpretation of the posterior covariance operator 198
4.3.3: The resolution operator 199
4.3.4: Eigenvector analysis 201
4.3.5: The posterior covariance operator in the data space. Importance of a datum 210
4.3.6: Are the residuals too large? 211

4.4: The mathematics of discrete least-squares 213

4.4.1: The least-squares definition of distance, norm and scalar product 213
4.4.2: Dual spaces 217
4.4.3: Gradient, Hessian, Steepest ascent direction, and Curvature of $S(\mathbf{m})$ 227
4.4.4: Introduction to gradient methods 230
4.4.5: The choice of the optimization method 239

4.5: Methods of resolution (II) 241

4.5.1: Preconditioned steepest descent 241
4.5.2: Preconditioned conjugate directions 242
4.5.3: Quasi-Newton method 243
4.5.4: Variable metric method 245

4.6: Particular formulas for linear problems 249

4.6.1: Preconditioned steepest descent 250

4.6.2: Preconditioned conjugate directions 251
4.6.3: Newton method 251
4.6.4: Variable metric method 252

4.7: Particular formulas for linearizable problems 252

4.7.1: Preconditioned steepest descent 254
4.7.2: Preconditioned conjugate directions 254
4.7.3: Newton method 255
4.8.4: Variable metric method 255

Box 4.1: The square root of a covariance matrix 205
Box 4.2: The Lanczos decomposition 207
Box 4.3: Distances, norms, ans scalar products 215
Box 4.4: The different meanings of the word "kernel" 223
Box 4.5: Transposed operator and adjoint operator 224
Box 4.6: The notation $O(\| m \|^r)$ 230
Box 4.7: Gradient and direction of steepest ascent 235
Box 4.8: What is ART? 248

Appendix 4.1: Tensor notations for Discrete Inverse Problems 256
Appendix 4.2: More notations 262
Appendix 4.3: Rate of convergence of an iterative sequence . . 264
Appendix 4.4: Iterative resolution of linear systems 267
Appendix 4.5: The Cholesky decomposition 269
Appendix 4.6: The Lagrange multipliers 271
Appendix 4.7: Gradient methods of minimization in normed linear spaces 277
Appendix 4.8: The spherical harmonics: basis of functional spaces 287

Problems for Chapter 4 292

5: **The least-absolute values (ℓ_1-norm) criterion and the minimax (ℓ_∞-norm) criterion** 303

5.1: The ℓ_p norm 304

5.1.1: Definition of the weighted ℓ_p norm 304
5.1.2: The dual of an ℓ_p space 305
5.1.3: The ℓ_p-norm criterion for the resolution of inverse problems 314

5.2: The ℓ_1-norm criterion for the resolution of inverse problems 316

5.2.1: Linear problems 318
5.2.2: Nonlinear problems 322

5.3: The ℓ_1-norm criterion and the method of steepest descent 325

5.3.1: Gradient and direction of steepest ascent 325
5.3.2: The algorithm 326

5.4: The ℓ_1-norm criterion and the linear programming techniques . . 328

5.4.1: The FIFO method 329
5.4.2: The FIFO algorithm 334
5.4.3: Application to nonlinear problems 335

5.5: Robust inversion using the Cauchy probability density 342

5.6: The ℓ_∞-norm criterion for the resolution of inverse problems . . 347

5.7: The ℓ_∞ norm and the method of steepest descent 349

5.7.1: Gradient and direction of steepest ascent 350
5.7.2: The algorithm 350

5.8: The ℓ_∞-norm criterion and the linear programming techniques . 351

Box 5.1: Gradient and direction of steepest ascent in ℓ_p norm spaces 310
Box 5.2: The multidimensional exponential probability 322
Box 5.3: The simplex method of linear programming 336
Box 5.4: ℓ_1-norm minimization using linear programming 339
Box 5.5: The Cauchy probability density 345
Box 5.6: Dual problems in linear programming 353
Box 5.7: Slack variables and linear programming 355

Appendix 5.1: The Newton step for the minimization of a ℓ_p norm 357
Appendix 5.2: Gradient methods for ℓ_p norms 360

Problems for Chapter 5 372

PART TWO: GENERAL INVERSE PROBLEMS

6: The general problem 391

6.1: Random processes in functional spaces 392

6.1.1: Description of a random function 392
6.1.2: Computing probabilities 395
6.1.3: Does a covariance function give a convenient description of a random function? 401
6.1.4: General random processes 402

6.2: Solution of general random processes 404

Box 6.1: The covariance operator as a scalar product over the dual space 396
Box 6.2: Cylindrical measures 398
Box 6.3: The Bayesian viewpoint of Backus (1970) 405

7: The least-squares criterion in functional spaces 409

7.1: Covariance operators and working spaces 410

7.1.1: Representation of the dual of a linear space 410
7.1.2: Definition of a covariance operator 411
7.1.3: Some examples of covariance functions and associated random realizations 413
7.1.4: The power spectrum of a random realization equals the spectral amplitude of the correlation function 421
7.1.5: Weighting operators, scalar products, and "least-squares" norms 425
7.1.6: Relationship between the "least-squares" norm and the L_2 norm 429

7.2: Derivative operators and transposed operators in functional spaces 430

7.2.1: The derivative operator 430
7.2.2: The transposed operator 442

7.3: General least-squares inversion 456

7.3.1: Linear problems 456
7.3.2: Nonlinear problems 457
7.3.3: Linearizable problems 458

7.4: Methods of resolution 459

7.5: Example: The inverse problem of X-ray tomography 465

7.5.1: Iterative solution 467
7.5.2: Newton's noniterative solution 473

7.6: Example: The inversion of acoustic waveforms 478

7.7: Analysis of error and resolution 488

7.7.1: Analysis of error 489
7.7.2: Analysis of resolution 491
7.7.3: Practical computation of the resolving kernel and the posterior covariance function 494

7.8: Do it formally 495

Box 7.1: Transposed and adjoint of a differential operator 447
Box 7.2: The method of Backus and Gilbert 461
Box 7.3: The first Born and Rytov approximations 484

Appendix 7.1: Spaces and operators: basic definitions and properties 507
Appendix 7.2: Usual functional spaces 525

Problems for Chapter 7 529

REFERENCES AND REFERENCES FOR GENERAL READING . . . 585

INDEX 601