Functions of Matrices: Theory and Computation

Higham, Nicholas J.

2008

MIMS EPrint: **2008.39**

ISSN 1749-9097

Contents

List of Figures **xiii**

List of Tables **xv**

Preface **xvii**

1 Theory of Matrix Functions **1**

1.1 Introduction 1
1.2 Definitions of $f(A)$ 2
1.2.1 Jordan Canonical Form 2
1.2.2 Polynomial Interpolation 4
1.2.3 Cauchy Integral Theorem 7
1.2.4 Equivalence of Definitions 8
1.2.5 Example: Function of Identity Plus Rank-1 Matrix 8
1.2.6 Example: Function of Discrete Fourier Transform Matrix . . . 10
1.3 Properties 10
1.4 Nonprimary Matrix Functions 14
1.5 Existence of (Real) Matrix Square Roots and Logarithms 16
1.6 Classification of Matrix Square Roots and Logarithms 17
1.7 Principal Square Root and Logarithm 20
1.8 $f(AB)$ and $f(BA)$ 21
1.9 Miscellany 23
1.10 A Brief History of Matrix Functions 26
1.11 Notes and References 27
Problems 29

2 Applications **35**

2.1 Differential Equations 35
2.1.1 Exponential Integrators 36
2.2 Nuclear Magnetic Resonance 37
2.3 Markov Models 37
2.4 Control Theory 39
2.5 The Nonsymmetric Eigenvalue Problem 41
2.6 Orthogonalization and the Orthogonal Procrustes Problem 42
2.7 Theoretical Particle Physics 43
2.8 Other Matrix Functions 44
2.9 Nonlinear Matrix Equations 44
2.10 Geometric Mean 46
2.11 Pseudospectra 47
2.12 Algebras 47

2.13 Sensitivity Analysis 48
2.14 Other Applications 48
2.14.1 Boundary Value Problems 48
2.14.2 Semidefinite Programming 48
2.14.3 Matrix Sector Function 48
2.14.4 Matrix Disk Function 49
2.14.5 The Average Eye in Optics 50
2.14.6 Computer Graphics 50
2.14.7 Bregman Divergences 50
2.14.8 Structured Matrix Interpolation 50
2.14.9 The Lambert W Function and Delay Differential Equations . 51
2.15 Notes and References 51
Problems 52

3 Conditioning **55**
3.1 Condition Numbers 55
3.2 Properties of the Fréchet Derivative 57
3.3 Bounding the Condition Number 63
3.4 Computing or Estimating the Condition Number 64
3.5 Notes and References 69
Problems 70

4 Techniques for General Functions **71**
4.1 Matrix Powers 71
4.2 Polynomial Evaluation 72
4.3 Taylor Series 76
4.4 Rational Approximation 78
4.4.1 Best L_∞ Approximation 79
4.4.2 Padé Approximation 79
4.4.3 Evaluating Rational Functions 80
4.5 Diagonalization 81
4.6 Schur Decomposition and Triangular Matrices 84
4.7 Block Diagonalization 89
4.8 Interpolating Polynomial and Characteristic Polynomial 89
4.9 Matrix Iterations 91
4.9.1 Order of Convergence 91
4.9.2 Termination Criteria 92
4.9.3 Convergence 93
4.9.4 Numerical Stability 95
4.10 Preprocessing 99
4.11 Bounds for $\|f(A)\|$ 102
4.12 Notes and References 104
Problems 105

5 Matrix Sign Function **107**
5.1 Sensitivity and Conditioning 109
5.2 Schur Method 112
5.3 Newton's Method 113
5.4 The Padé Family of Iterations 115
5.5 Scaling the Newton Iteration 119

5.6 Terminating the Iterations 121
5.7 Numerical Stability of Sign Iterations 123
5.8 Numerical Experiments and Algorithm 125
5.9 Best L_∞ Approximation 128
5.10 Notes and References 129
Problems 131

6 Matrix Square Root 133
6.1 Sensitivity and Conditioning 133
6.2 Schur Method 135
6.3 Newton's Method and Its Variants 139
6.4 Stability and Limiting Accuracy 144
6.4.1 Newton Iteration 144
6.4.2 DB Iterations 145
6.4.3 CR Iteration 146
6.4.4 IN Iteration 146
6.4.5 Summary 147
6.5 Scaling the Newton Iteration 147
6.6 Numerical Experiments 148
6.7 Iterations via the Matrix Sign Function 152
6.8 Special Matrices 154
6.8.1 Binomial Iteration 154
6.8.2 Modified Newton Iterations 157
6.8.3 M-Matrices and H-Matrices 159
6.8.4 Hermitian Positive Definite Matrices 161
6.9 Computing Small-Normed Square Roots 162
6.10 Comparison of Methods 164
6.11 Involutory Matrices 165
6.12 Notes and References 166
Problems 168

7 Matrix pth Root 173
7.1 Theory 173
7.2 Schur Method 175
7.3 Newton's Method 177
7.4 Inverse Newton Method 181
7.5 Schur–Newton Algorithm 184
7.6 Matrix Sign Method 186
7.7 Notes and References 187
Problems 189

8 The Polar Decomposition 193
8.1 Approximation Properties 197
8.2 Sensitivity and Conditioning 199
8.3 Newton's Method 202
8.4 Obtaining Iterations via the Matrix Sign Function 202
8.5 The Padé Family of Methods 203
8.6 Scaling the Newton Iteration 205
8.7 Terminating the Iterations 207
8.8 Numerical Stability and Choice of H 209

8.9 Algorithm 210
8.10 Notes and References 213
Problems 216

9 Schur–Parlett Algorithm 221
9.1 Evaluating Functions of the Atomic Blocks 221
9.2 Evaluating the Upper Triangular Part of $f(T)$ 225
9.3 Reordering and Blocking the Schur Form 226
9.4 Schur–Parlett Algorithm for $f(A)$ 228
9.5 Preprocessing 230
9.6 Notes and References 231
Problems 231

10 Matrix Exponential 233
10.1 Basic Properties 233
10.2 Conditioning 238
10.3 Scaling and Squaring Method 241
10.4 Schur Algorithms 250
10.4.1 Newton Divided Difference Interpolation 250
10.4.2 Schur–Fréchet Algorithm 251
10.4.3 Schur–Parlett Algorithm 251
10.5 Numerical Experiment 252
10.6 Evaluating the Fréchet Derivative and Its Norm 253
10.6.1 Quadrature 254
10.6.2 The Kronecker Formulae 256
10.6.3 Computing and Estimating the Norm 258
10.7 Miscellany 259
10.7.1 Hermitian Matrices and Best L_∞ Approximation 259
10.7.2 Essentially Nonnegative Matrices 260
10.7.3 Preprocessing 261
10.7.4 The ψ Functions 261
10.8 Notes and References 262
Problems 265

11 Matrix Logarithm 269
11.1 Basic Properties 269
11.2 Conditioning 272
11.3 Series Expansions 273
11.4 Padé Approximation 274
11.5 Inverse Scaling and Squaring Method 275
11.5.1 Schur Decomposition: Triangular Matrices 276
11.5.2 Full Matrices 278
11.6 Schur Algorithms 279
11.6.1 Schur–Fréchet Algorithm 279
11.6.2 Schur–Parlett Algorithm 279
11.7 Numerical Experiment 280
11.8 Evaluating the Fréchet Derivative 281
11.9 Notes and References 283
Problems 284

12 Matrix Cosine and Sine **287**
12.1 Basic Properties 287
12.2 Conditioning 289
12.3 Padé Approximation of Cosine 290
12.4 Double Angle Algorithm for Cosine 290
12.5 Numerical Experiment 295
12.6 Double Angle Algorithm for Sine and Cosine 296
12.6.1 Preprocessing 299
12.7 Notes and References 299
Problems 300

13 Function of Matrix Times Vector: $f(A)b$ **301**
13.1 Representation via Polynomial Interpolation 301
13.2 Krylov Subspace Methods 302
13.2.1 The Arnoldi Process 302
13.2.2 Arnoldi Approximation of $f(A)b$ 304
13.2.3 Lanczos Biorthogonalization 306
13.3 Quadrature 306
13.3.1 On the Real Line 306
13.3.2 Contour Integration 307
13.4 Differential Equations 308
13.5 Other Methods 309
13.6 Notes and References 309
Problems 310

14 Miscellany **313**
14.1 Structured Matrices 313
14.1.1 Algebras and Groups 313
14.1.2 Monotone Functions 315
14.1.3 Other Structures 315
14.1.4 Data Sparse Representations 316
14.1.5 Computing Structured $f(A)$ for Structured A 316
14.2 Exponential Decay of Functions of Banded Matrices 317
14.3 Approximating Entries of Matrix Functions 318

A Notation **319**

B Background: Definitions and Useful Facts **321**
B.1 Basic Notation 321
B.2 Eigenvalues and Jordan Canonical Form 321
B.3 Invariant Subspaces 323
B.4 Special Classes of Matrices 323
B.5 Matrix Factorizations and Decompositions 324
B.6 Pseudoinverse and Orthogonality 325
B.6.1 Pseudoinverse 325
B.6.2 Projector and Orthogonal Projector 326
B.6.3 Partial Isometry 326
B.7 Norms 326
B.8 Matrix Sequences and Series 328
B.9 Perturbation Expansions for Matrix Inverse 328

B.10 Sherman–Morrison–Woodbury Formula 329
B.11 Nonnegative Matrices 329
B.12 Positive (Semi)definite Ordering 330
B.13 Kronecker Product and Sum 331
B.14 Sylvester Equation 331
B.15 Floating Point Arithmetic 331
B.16 Divided Differences 332
Problems 334

C Operation Counts **335**

D Matrix Function Toolbox **339**

E Solutions to Problems **343**

Bibliography **379**

Index **415**