Contents

1 Conjugate Direction Methods for Quadratic Problems 1
1.1 Introduction 1
1.2 Conjugate Direction Methods 4
1.3 Conjugate Gradient Error Algorithms 13
1.4 Conjugate Gradient Residual Algorithms 21
1.5 The Method of Shortest Residuals 26
1.6 Rate of Convergence 28
1.7 Relations with a Direct Solver for a Set of Linear Equations 38
1.8 Limited Memory Quasi-Newton Algorithms 45
1.9 Reduced-Hessian Quasi-Newton Algorithms 53
1.10 Limited Memory Reduced-Hessian Quasi-Newton Algorithms 56
1.11 Conjugate Non-Gradient Algorithm 59
1.12 Notes 61

2 Conjugate Gradient Methods for Nonconvex Problems 63
2.1 Introduction 63
2.2 Line Search Methods 65
2.3 General Convergence Results 70
2.4 Moré–Thuente Step-Length Selection Algorithm 76
2.5 Hager–Zhang Step-Length Selection Algorithm 80
2.6 Nonlinear Conjugate Gradient Algorithms 84
2.7 Global Convergence of the Fletcher–Reeves Algorithm 87
2.8 Other Versions of the Fletcher–Reeves Algorithm 96
2.9 Global Convergence of the Polak–Ribière Algorithm 98
2.10 Hestenes–Stiefel Versions of the Standard Conjugate Gradient Algorithm 101
2.11 Notes 106

3 Memoryless Quasi-Newton Methods 109
3.1 Introduction 109

3.2 Conjugate Gradient Algorithm as the Memoryless Quasi-Newton Method . . . 112
3.3 Beale's Restart Rule . . . 116
3.4 Convergence of the Shanno Algorithm . . . 118
3.5 Notes . . . 127

4 Preconditioned Conjugate Gradient Algorithms . . . 133
4.1 Introduction . . . 133
4.2 Rates of Convergence . . . 134
4.3 Conjugate Gradient Algorithm and the Newton Method . . . 139
4.4 Generic Preconditioned Conjugate Gradient Algorithm . . . 145
4.5 Quasi-Newton-like Variable Storage Conjugate Gradient Algorithm . . . 148
4.6 Scaling the Identity . . . 155
4.7 Notes . . . 158

5 Limited Memory Quasi-Newton Algorithms . . . 159
5.1 Introduction . . . 159
5.2 Global Convergence of the Limited Memory BFGS Algorithm . . . 163
5.3 Compact Representation of the BFGS Approximation of the Inverse Hessian Matrix . . . 169
5.4 The Compact Representation of the BFGS Approximation of the Hessian Matrix . . . 175
5.5 Numerical Experiments . . . 178
5.6 Notes . . . 186

6 The Method of Shortest Residuals and Nondifferentiable Optimization . . . 191
6.1 Introduction . . . 191
6.2 The Method of Conjugate Subgradient by Lemaréchal and Wolfe . . . 192
6.3 Conjugate Subgradient Algorithm for Problems with Semismooth Functions . . . 204
6.4 Subgradient Algorithms with Finite Storage . . . 210
6.5 Notes . . . 215

7 The Method of Shortest Residuals for Differentiable Problems . . . 217
7.1 Introduction . . . 217
7.2 General Algorithm . . . 219
7.3 Versions of the Method of Shortest Residuals . . . 226
7.4 Polak–Ribière Version of the Method of Shortest Residuals . . . 230
7.5 The Method of Shortest Residuals by Dai and Yuan . . . 233
7.6 Global Convergence of the Lemaréchal–Wolfe Algorithm Without Restarts . . . 241
7.7 A Counter-Example . . . 244
7.8 Numerical Experiments: First Comparisons . . . 250

7.9 Numerical Experiments: Comparison to the Memoryless Quasi-Newton Method . . . 256
7.10 Numerical Experiments: Comparison to the Limited Memory Quasi-Newton Method . . . 257
7.11 Numerical Experiments: Comparison to the Hager–Zhang Algorithm . . . 264
7.12 Notes . . . 267

8 The Preconditioned Shortest Residuals Algorithm . . . 279
8.1 Introduction . . . 279
8.2 General Preconditioned Method of Shortest Residuals . . . 281
8.3 Globally Convergent Preconditioned Conjugate Gradient Algorithm . . . 286
8.4 Scaling Matrices . . . 288
8.5 Conjugate Gradient Algorithm with the BFGS Scaling Matrices . . . 291
8.6 Numerical Experiments . . . 293
8.7 Notes . . . 297

9 Optimization on a Polyhedron . . . 299
9.1 Introduction . . . 299
9.2 Exposed Sets . . . 308
9.3 The Identification of Active Constraints . . . 310
9.4 Gradient Projection Method . . . 313
9.5 On the Stopping Criterion . . . 319
9.6 Notes . . . 324

10 Conjugate Gradient Algorithms for Problems with Box Constraints . 327
10.1 Introduction . . . 327
10.2 General Convergence Theory . . . 328
10.3 The Globally Convergent Polak–Ribière Version of Algorithm 10.1 . . . 342
10.4 Convergence Analysis of the Fletcher–Reeves Version of Algorithm 10.1 . . . 350
10.5 Numerical Experiments . . . 357
10.6 Notes . . . 359

11 Preconditioned Conjugate Gradient Algorithms for Problems with Box Constraints . . . 371
11.1 Introduction . . . 371
11.2 Active Constraints Identification by Solving Quadratic Problem . . . 371
11.3 The Preconditioned Shortest Residuals Algorithm: Outline . . . 376
11.4 The Preconditioned Shortest Residuals Algorithm: Global Convergence . . . 384
11.5 The Limited Memory Quasi-Newton Method for Problems with Box Constraints . . . 386
11.6 Numerical Algebra . . . 387

11.7 Algorithm 11.1 and Algorithm 11.2 Applied to Problems with Strongly Convex Functions . . . 389
11.8 Numerical Experiments . . . 390
11.9 Notes . . . 391

12 Preconditioned Conjugate Gradient Based Reduced-Hessian Methods . . . 399
12.1 Introduction . . . 399
12.2 BFGS Update with Column Scaling . . . 400
12.3 The Preconditioned Method of Shortest Residuals with Column Scaling . . . 409
12.4 Reduced-Hessian Quasi-Newton Algorithms . . . 412
12.5 Limited Memory Reduced-Hessian Quasi-Newton Algorithms . . . 420
12.6 Preconditioned Conjugate Gradient Based Reduced-Hessian Algorithm . . . 424
12.7 Notes . . . 428

A Elements of Topology and Analysis . . . 429
A.1 The Topology of the Euclidean Space . . . 429
A.2 Continuity and Convexity . . . 432
A.3 Derivatives . . . 435

B Elements of Linear Algebra . . . 441
B.1 Vector and Matrix Norms . . . 441
B.2 Spectral Decomposition . . . 442
B.3 Determinant and Trace . . . 447

C Elements of Numerical Linear Algebra . . . 449
C.1 Condition Number and Linear Equations . . . 449
C.2 The LU and Cholesky Factorizations . . . 450
C.3 The QR Factorization . . . 452
C.4 Householder QR . . . 454
C.5 Givens QR . . . 456
C.6 Gram–Schmidt QR . . . 459

References . . . 463

Index . . . 473