

Contents

<i>Preface</i>	<i>page</i> xvii
<i>Notation</i>	xxi
<i>Abbreviations</i>	xxxix
<i>Algorithms</i>	xxxiv
PART I MATRIX ALGEBRA	1
1 Introduction to Matrix Algebra	3
1.1 Basic Concepts of Vectors and Matrices	3
1.1.1 Vectors and Matrices	3
1.1.2 Basic Vector Calculus	6
1.1.3 Basic Matrix Calculus	8
1.1.4 Linear Independence of Vectors	11
1.1.5 Matrix Functions	11
1.2 Elementary Row Operations and Applications	13
1.2.1 Elementary Row Operations	13
1.2.2 Gauss Elimination Methods	16
1.3 Sets, Vector Subspaces and Linear Mapping	20
1.3.1 Sets	20
1.3.2 Fields and Vector Spaces	22
1.3.3 Linear Mapping	24
1.4 Inner Products and Vector Norms	27
1.4.1 Inner Products of Vectors	27
1.4.2 Norms of Vectors	28
1.4.3 Similarity Comparison Between Vectors	32
1.4.4 Banach Space, Euclidean Space, Hilbert Space	35
1.4.5 Inner Products and Norms of Matrices	36
1.5 Random Vectors	40
1.5.1 Statistical Interpretation of Random Vectors	41
1.5.2 Gaussian Random Vectors	44

viii

1.6	Performance Indexes of Matrices	47
1.6.1	Quadratic Forms	47
1.6.2	Determinants	49
1.6.3	Matrix Eigenvalues	52
1.6.4	Matrix Trace	54
1.6.5	Matrix Rank	56
1.7	Inverse Matrices and Pseudo-Inverse Matrices	59
1.7.1	Definition and Properties of Inverse Matrices	59
1.7.2	Matrix Inversion Lemma	60
1.7.3	Inversion of Hermitian Matrices	61
1.7.4	Left and Right Pseudo-Inverse Matrices	63
1.8	Moore–Penrose Inverse Matrices	65
1.8.1	Definition and Properties	65
1.8.2	Computation of Moore–Penrose Inverse Matrix	69
1.9	Direct Sum and Hadamard Product	71
1.9.1	Direct Sum of Matrices	71
1.9.2	Hadamard Product	72
1.10	Kronecker Products and Khatri–Rao Product	75
1.10.1	Kronecker Products	75
1.10.2	Generalized Kronecker Products	77
1.10.3	Khatri–Rao Product	78
1.11	Vectorization and Matricization	79
1.11.1	Vectorization and Commutation Matrix	79
1.11.2	Matricization of a Vector	82
1.11.3	Properties of Vectorization Operator	83
1.12	Sparse Representations	84
1.12.1	Sparse Vectors and Sparse Representations	84
1.12.2	Sparse Representation of Face Recognition	86
	Exercises	87
2	Special Matrices	95
2.1	Hermitian Matrices	95
2.2	Idempotent Matrix	96
2.3	Permutation Matrix	98
2.3.1	Permutation Matrix and Exchange Matrix	98
2.3.2	Generalized Permutation Matrix	101
2.4	Orthogonal Matrix and Unitary Matrix	104
2.5	Band Matrix and Triangular Matrix	107
2.5.1	Band Matrix	107
2.5.2	Triangular Matrix	107
2.6	Summing Vector and Centering Matrix	109
2.6.1	Summing Vector	110

	2.6.2	Centering Matrix	111
2.7		Vandermonde Matrix and Fourier Matrix	112
	2.7.1	Vandermonde Matrix	112
	2.7.2	Fourier Matrix	114
	2.7.3	Index Vectors	116
	2.7.4	FFT Algorithm	118
2.8		Hadamard Matrix	121
2.9		Toeplitz Matrix	123
	2.9.1	Symmetric Toeplitz Matrix	123
	2.9.2	Discrete Cosine Transform of Toeplitz Matrix	125
		Exercises	126
3		Matrix Differential	129
	3.1	Jacobian Matrix and Gradient Matrix	129
	3.1.1	Jacobian Matrix	130
	3.1.2	Gradient Matrix	132
	3.1.3	Calculation of Partial Derivative and Gradient	133
	3.2	Real Matrix Differential	139
	3.2.1	Calculation of Real Matrix Differential	139
	3.2.2	Jacobian Matrix Identification	141
	3.2.3	Jacobian Matrix of Real Matrix Functions	147
	3.3	Real Hessian Matrix and Identification	150
	3.3.1	Real Hessian Matrix	150
	3.3.2	Real Hessian Matrix Identification	152
	3.4	Complex Gradient Matrices	157
	3.4.1	Holomorphic Function and Complex Partial Derivative	157
	3.4.2	Complex Matrix Differential	161
	3.4.3	Complex Gradient Matrix Identification	169
	3.5	Complex Hessian Matrices and Identification	174
	3.5.1	Complex Hessian Matrices	175
	3.5.2	Complex Hessian Matrix Identification	177
		Exercises	179
		PART II MATRIX ANALYSIS	181
4		Gradient Analysis and Optimization	183
	4.1	Real Gradient Analysis	183
	4.1.1	Stationary Points and Extreme Points	184
	4.1.2	Real Gradient Analysis of $f(\mathbf{x})$	186
	4.1.3	Real Gradient Analysis of $f(\mathbf{X})$	188
	4.2	Gradient Analysis of Complex Variable Function	190
	4.2.1	Extreme Point of Complex Variable Function	190

	4.2.2	Complex Gradient Analysis	195
4.3		Convex Sets and Convex Function Identification	198
	4.3.1	Standard Constrained Optimization Problems	198
	4.3.2	Convex Sets and Convex Functions	200
	4.3.3	Convex Function Identification	203
4.4		Gradient Methods for Smooth Convex Optimization	205
	4.4.1	Gradient Method	205
	4.4.2	Conjugate Gradient Method	210
	4.4.3	Convergence Rates	215
4.5		Nesterov Optimal Gradient Method	217
	4.5.1	Lipschitz Continuous Function	217
	4.5.2	Nesterov Optimal Gradient Algorithms	220
4.6		Nonsmooth Convex Optimization	223
	4.6.1	Subgradient and Subdifferential	224
	4.6.2	Proximal Operator	228
	4.6.3	Proximal Gradient Method	232
4.7		Constrained Convex Optimization	237
	4.7.1	Lagrange Multiplier Method	237
	4.7.2	Penalty Function Method	238
	4.7.3	Augmented Lagrange Multiplier Method	240
	4.7.4	Lagrangian Dual Method	242
	4.7.5	Karush–Kuhn–Tucker Conditions	244
	4.7.6	Alternating Direction Method of Multipliers	248
4.8		Newton Methods	251
	4.8.1	Newton Method for Unconstrained Optimization	251
	4.8.2	Newton Method for Constrained Optimization	254
4.9		Original–Dual Interior-Point Method	260
	4.9.1	Original–Dual Problems	260
	4.9.2	First-Order Original–Dual Interior-Point Method	261
	4.9.3	Second-Order Original–Dual Interior-Point Method	263
		Exercises	268
5		Singular Value Analysis	271
	5.1	Numerical Stability and Condition Number	271
	5.2	Singular Value Decomposition (SVD)	274
	5.2.1	Singular Value Decomposition	274
	5.2.2	Properties of Singular Values	277
	5.2.3	Rank-Deficient Least Squares Solutions	280
	5.3	Product Singular Value Decomposition (PSVD)	283
	5.3.1	PSVD Problem	283
	5.3.2	Accurate Calculation of PSVD	284
	5.4	Applications of Singular Value Decomposition	285

	xi	
5.4.1	Static Systems	286
5.4.2	Image Compression	287
5.5	Generalized Singular Value Decomposition (GSVD)	289
5.5.1	Definition and Properties	289
5.5.2	Algorithms for GSVD	292
5.5.3	Two Application Examples of GSVD	294
5.6	Low-Rank–Sparse Matrix Decomposition	296
5.6.1	Matrix Decomposition Problems	297
5.6.2	Singular Value Thresholding	298
5.6.3	Robust Principal Component Analysis	300
5.7	Matrix Completion	302
5.7.1	Matrix Completion Problems	303
5.7.2	Matrix Completion Model and Incoherence	305
5.7.3	Singular Value Thresholding Algorithm	306
5.7.4	Fast and Accurate Matrix Completion	308
	Exercises	312
6	Solving Matrix Equations	315
6.1	Least Squares Method	316
6.1.1	Ordinary Least Squares Methods	316
6.1.2	Properties of Least Squares Solutions	317
6.1.3	Data Least Squares	320
6.2	Tikhonov Regularization and Gauss–Seidel Method	321
6.2.1	Tikhonov Regularization	321
6.2.2	Regularized Gauss–Seidel Method	324
6.3	Total Least Squares (TLS) Methods	328
6.3.1	TLS Problems	328
6.3.2	TLS Solution	329
6.3.3	Performances of TLS Solution	333
6.3.4	Generalized Total Least Squares	335
6.3.5	Total Least Squares Fitting	337
6.3.6	Total Maximum Likelihood Method	342
6.4	Constrained Total Least Squares	344
6.4.1	Constrained Total Least Squares Method	345
6.4.2	Harmonic Superresolution	347
6.4.3	Image Restoration	348
6.5	Subspace Method for Solving Blind Matrix Equations	350
6.6	Nonnegative Matrix Factorization: Optimization Theory	353
6.6.1	Nonnegative Matrices	353
6.6.2	Nonnegativity and Sparsity Constraints	355
6.6.3	Nonnegative Matrix Factorization Model	356
6.6.4	Divergences and Deformed Logarithm	361

6.7	Nonnegative Matrix Factorization: Optimization Algorithms	366
6.7.1	Multiplication Algorithms	366
6.7.2	Nesterov Optimal Gradient Algorithm	372
6.7.3	Alternating Nonnegative Least Squares	374
6.7.4	Quasi-Newton Method	377
6.7.5	Sparse Nonnegative Matrix Factorization	378
6.8	Sparse Matrix Equation Solving: Optimization Theory	381
6.8.1	ℓ_1 -Norm Minimization	381
6.8.2	Lasso and Robust Linear Regression	384
6.8.3	Mutual Coherence and RIP Conditions	387
6.8.4	Relation to Tikhonov Regularization	389
6.8.5	Gradient Analysis of ℓ_1 -Norm Minimization	390
6.9	Sparse Matrix Equation Solving: Optimization Algorithms	391
6.9.1	Basis Pursuit Algorithms	391
6.9.2	First-Order Augmented Lagrangian Algorithm	394
6.9.3	Barzilai–Borwein Gradient Projection Algorithm	394
6.9.4	ADMM Algorithms for Lasso Problems	397
6.9.5	LARS Algorithms for Lasso Problems	398
6.9.6	Covariance Graphical Lasso Method	400
6.9.7	Homotopy Algorithm	402
6.9.8	Bregman Iteration Algorithms	403
	Exercises	409
7	Eigenanalysis	413
7.1	Eigenvalue Problem and Characteristic Equation	413
7.1.1	Eigenvalue Problem	413
7.1.2	Characteristic Polynomial	415
7.2	Eigenvalues and Eigenvectors	416
7.2.1	Eigenvalues	416
7.2.2	Eigenvectors	419
7.3	Similarity Reduction	422
7.3.1	Similarity Transformation of Matrices	423
7.3.2	Similarity Reduction of Matrices	426
7.3.3	Similarity Reduction of Matrix Polynomials	430
7.4	Polynomial Matrices and Balanced Reduction	434
7.4.1	Smith Normal Forms	434
7.4.2	Invariant Factor Method	437
7.4.3	Conversion of Jordan Form and Smith Form	441
7.4.4	Finding Smith Blocks from Jordan Blocks	442
7.4.5	Finding Jordan Blocks from Smith Blocks	443
7.5	Cayley–Hamilton Theorem with Applications	446
7.5.1	Cayley–Hamilton Theorem	446

	xiii	
7.5.2	Computation of Inverse Matrices	448
7.5.3	Computation of Matrix Powers	450
7.5.4	Calculation of Matrix Exponential Functions	452
7.6	Application Examples of Eigenvalue Decomposition	455
7.6.1	Pisarenko Harmonic Decomposition	455
7.6.2	Discrete Karhunen–Loeve Transformation	458
7.6.3	Principal Component Analysis	461
7.7	Generalized Eigenvalue Decomposition (GEVD)	463
7.7.1	Generalized Eigenvalue Decomposition	463
7.7.2	Total Least Squares Method for GEVD	467
7.7.3	Application of GEVD: ESPRIT	468
7.7.4	Similarity Transformation in GEVD	471
7.8	Rayleigh Quotient	474
7.8.1	Definition and Properties of Rayleigh Quotient	474
7.8.2	Rayleigh Quotient Iteration	475
7.8.3	Algorithms for Rayleigh Quotient	476
7.9	Generalized Rayleigh Quotient	478
7.9.1	Definition and Properties	478
7.9.2	Effectiveness of Class Discrimination	480
7.9.3	Robust Beamforming	482
7.10	Quadratic Eigenvalue Problems	484
7.10.1	Description of Quadratic Eigenvalue Problems	484
7.10.2	Solving Quadratic Eigenvalue Problems	486
7.10.3	Application Examples	490
7.11	Joint Diagonalization	495
7.11.1	Joint Diagonalization Problems	495
7.11.2	Orthogonal Approximate Joint Diagonalization	497
7.11.3	Nonorthogonal Approximate Joint Diagonalization	500
	Exercises	503
8	Subspace Analysis and Tracking	511
8.1	General Theory of Subspaces	511
8.1.1	Bases of Subspaces	511
8.1.2	Disjoint Subspaces and Orthogonal Complement	513
8.2	Column Space, Row Space and Null Space	516
8.2.1	Definitions and Properties	516
8.2.2	Subspace Basis Construction	520
8.2.3	SVD-Based Orthonormal Basis Construction	522
8.2.4	Basis Construction of Subspaces Intersection	525
8.3	Subspace Methods	526
8.3.1	Signal Subspace and Noise Subspace	526
8.3.2	Multiple Signal Classification (MUSIC)	529

xiv

	8.3.3	Subspace Whitening	531
8.4		Grassmann Manifold and Stiefel Manifold	532
	8.4.1	Equivalent Subspaces	532
	8.4.2	Grassmann Manifold	533
	8.4.3	Stiefel Manifold	535
8.5		Projection Approximation Subspace Tracking (PAST)	536
	8.5.1	Basic PAST Theory	537
	8.5.2	PAST Algorithms	540
8.6		Fast Subspace Decomposition	542
	8.6.1	Rayleigh–Ritz Approximation	542
	8.6.2	Fast Subspace Decomposition Algorithm	544
		Exercises	546
9		Projection Analysis	551
	9.1	Projection and Orthogonal Projection	551
		9.1.1 Projection Theorem	552
		9.1.2 Mean Square Estimation	554
	9.2	Projectors and Projection Matrices	556
		9.2.1 Projector and Orthogonal Projector	556
		9.2.2 Projection Matrices	558
		9.2.3 Derivatives of Projection Matrix	561
	9.3	Updating of Projection Matrices	562
		9.3.1 Updating Formulas for Projection Matrices	562
		9.3.2 Prediction Filters	564
		9.3.3 Updating of Lattice Adaptive Filter	567
	9.4	Oblique Projector of Full Column Rank Matrix	570
		9.4.1 Definition and Properties of Oblique Projectors	571
		9.4.2 Geometric Interpretation of Oblique Projectors	575
		9.4.3 Recursion of Oblique Projectors	578
	9.5	Oblique Projector of Full Row Rank Matrices	579
		9.5.1 Definition and Properties	579
		9.5.2 Calculation of Oblique Projection	581
		9.5.3 Applications of Oblique Projectors	583
		Exercises	585
		PART III HIGHER-ORDER MATRIX ANALYSIS	587
10		Tensor Analysis	589
	10.1	Tensors and their Presentation	589
		10.1.1 Tensors	589
		10.1.2 Tensor Representation	592
	10.2	Vectorization and Matricization of Tensors	597

10.2.1	Vectorization and Horizontal Unfolding	597
10.2.2	Longitudinal Unfolding of Tensors	601
10.3	Basic Algebraic Operations of Tensors	606
10.3.1	Inner Product, Norm and Outer Product	606
10.3.2	Mode- n Product of Tensors	608
10.3.3	Rank of Tensor	612
10.4	Tucker Decomposition of Tensors	614
10.4.1	Tucker Decomposition (Higher-Order SVD)	615
10.4.2	Third-Order SVD	617
10.4.3	Alternating Least Squares Algorithms	621
10.5	Parallel Factor Decomposition of Tensors	625
10.5.1	Bilinear Model	625
10.5.2	Parallel Factor Analysis	627
10.5.3	Uniqueness Condition	635
10.5.4	Alternating Least Squares Algorithm	637
10.6	Applications of Low-Rank Tensor Decomposition	641
10.6.1	Multimodal Data Fusion	642
10.6.2	Fusion of Multimodal Brain Images	644
10.6.3	Process Monitoring	646
10.6.4	Note on Other Applications	648
10.7	Tensor Eigenvalue Decomposition	649
10.7.1	Tensor–Vector Products	649
10.7.2	Determinants and Eigenvalues of Tensors	651
10.7.3	Generalized Tensor Eigenvalues Problems	656
10.7.4	Orthogonal Decomposition of Symmetric Tensors	658
10.8	Preprocessing and Postprocessing	659
10.8.1	Centering and Scaling of Multi-Way Data	660
10.8.2	Compression of Data Array	661
10.9	Nonnegative Tensor Decomposition Algorithms	664
10.9.1	Multiplication Algorithm	664
10.9.2	ALS Algorithms	667
10.10	Tensor Completion	670
10.10.1	Simultaneous Tensor Decomposition and Completion	671
10.10.2	Smooth PARAFAC Tensor Completion	674
10.11	Software	676
	Exercises	678
	<i>References</i>	681
	<i>Index</i>	708