
Contents

Preface to Third Edition	vii
Preface to Second Edition	ix
Preface to First Edition	xi
1 Introduction	1
1.1 Vectors	2
1.2 Arrays	3
1.3 Matrices	3
1.4 Representation of Data	6
1.5 What You Compute and What You Don't	7
1.6 R	8

Part I Linear Algebra

2 Vectors and Vector Spaces	17
2.1 Operations on Vectors	17
2.1.1 Linear Combinations and Linear Independence	18
2.1.2 Vector Spaces and Spaces of Vectors	20
2.1.3 Basis Sets for Vector Spaces	29
2.1.4 Inner Products	30
2.1.5 Norms	33
2.1.6 Normalized Vectors	38
2.1.7 Metrics and Distances	39
2.1.8 Orthogonal Vectors and Orthogonal Vector Spaces	41
2.1.9 The “One Vector”	42
2.2 Cartesian Coordinates and Geometrical Properties of Vectors	43
2.2.1 Cartesian Geometry	43
2.2.2 Projections	44

2.2.3	Angles between Vectors	45
2.2.4	Orthogonalization Transformations; Gram-Schmidt	46
2.2.5	Orthonormal Basis Sets	48
2.2.6	Approximation of Vectors	49
2.2.7	Flats, Affine Spaces, and Hyperplanes	51
2.2.8	Cones	51
2.2.9	Vector Cross Products in $\mathbb{R}^3$	54
2.3	Centered Vectors, and Variances and Covariances of Vectors	56
2.3.1	The Mean and Centered Vectors	57
2.3.2	The Standard Deviation, the Variance, and Scaled Vectors	58
2.3.3	Covariances and Correlations between Vectors	59
	Appendix: Vectors in $\mathbb{R}$	60
	Exercises	63
3	Basic Properties of Matrices	69
3.1	Basic Definitions and Notation	69
3.1.1	Matrix Shaping Operators	73
3.1.2	Partitioned Matrices and Submatrices	76
3.1.3	Matrix Addition	78
3.1.4	The Trace of a Square Matrix	80
3.1.5	The Determinant of a Square Matrix	81
3.1.6	Minors, Cofactors, and Adjugate Matrices	87
3.2	Multiplication of Matrices and Multiplication of Vectors and Matrices	91
3.2.1	Matrix Multiplication (Cayley)	91
3.2.2	Cayley Multiplication of Matrices with Special Patterns	95
3.2.3	Elementary Operations on Matrices	98
3.2.4	The Trace of a Cayley Product that Is Square	106
3.2.5	The Determinant of a Cayley Product of Square Matrices	106
3.2.6	Outer Products of Vectors	107
3.2.7	Bilinear and Quadratic Forms; Definiteness	108
3.2.8	Anisometric Spaces	110
3.2.9	The Hadamard Product	111
3.2.10	The Kronecker Product	112
3.2.11	The Inner Product of Matrices	115
3.3	Matrix Rank and the Inverse of a Matrix	117
3.3.1	The Rank of Partitioned Matrices, Products of Matrices, and Sums of Matrices	120
3.3.2	Full Rank Partitioning	122
3.3.3	Full Rank Matrices and Matrix Inverses	123
3.3.4	Full Rank Factorization	126
3.3.5	Multiplication by Full Rank Matrices	128
3.3.6	Nonfull Rank and Equivalent Matrices	130
3.3.7	Gramian Matrices: Products of the Form $A^T A$	132
3.3.8	A Lower Bound on the Rank of a Matrix Product	135
3.3.9	Determinants of Inverses	135
3.3.10	Inverses of Products and Sums of Nonsingular Matrices	136

3.3.11	Inverses of Matrices with Special Forms	138
3.3.12	Determining the Rank of a Matrix	139
3.4	The Schur Complement in Partitioned Square Matrices	139
3.4.1	Inverses of Partitioned Matrices	140
3.4.2	Determinants of Partitioned Matrices	140
3.5	Linear Systems of Equations	141
3.5.1	Solutions of Linear Systems	141
3.5.2	Null Space: The Orthogonal Complement	144
3.5.3	Orthonormal Completion	145
3.6	Generalized Inverses	145
3.6.1	Immediate Properties of Generalized Inverses	146
3.6.2	Special Generalized Inverses	146
3.6.3	The Moore-Penrose Inverse	146
3.6.4	Generalized Inverses of Products and Sums of Matrices	148
3.6.5	Generalized Inverses of Partitioned Matrices	149
3.7	Orthogonality	149
3.8	Eigenanalysis; Canonical Factorizations	152
3.8.1	Basic Properties of Eigenvalues and Eigenvectors	153
3.8.2	The Characteristic Polynomial	156
3.8.3	The Spectrum	159
3.8.4	Similarity Transformations	164
3.8.5	Schur Factorization	166
3.8.6	Similar Canonical Factorization; Diagonalizable Matrices	167
3.8.7	Properties of Diagonalizable Matrices	170
3.8.8	Eigenanalysis of Symmetric Matrices	172
3.8.9	Generalized Eigenvalues and Eigenvectors	177
3.8.10	Singular Values and the Singular Value Decomposition (SVD)	178
3.9	Positive Definite and Nonnegative Definite Matrices	181
3.9.1	Eigenvalues of Positive and Nonnegative Definite Matrices	182
3.9.2	Inverse of Positive Definite Matrices	182
3.9.3	Diagonalization of Positive Definite Matrices	182
3.9.4	Square Roots of Positive and Nonnegative Definite Matrices	183
3.10	Matrix Norms	183
3.10.1	Matrix Norms Induced from Vector Norms	184
3.10.2	The Frobenius Norm — The “Usual” Norm	187
3.10.3	Other Matrix Norms	189
3.10.4	Matrix Norm Inequalities	189
3.10.5	The Spectral Radius	190
3.10.6	Convergence of a Matrix Power Series	191
3.11	Approximation of Matrices	194
	Appendix: Matrices in R	197
	Exercises	203

4	Matrix Transformations and Factorizations	211
4.1	Linear Geometric Transformations	213
4.1.1	Transformations by Orthogonal Matrices	215
4.1.2	Rotations	215
4.1.3	Reflections	217
4.1.4	Translations; Homogeneous Coordinates	218
4.2	Householder Transformations (Reflections)	219
4.3	Givens Transformations (Rotations)	222
4.4	Factorization of Matrices	226
4.5	LU and LDU Factorizations	226
4.6	QR Factorization	232
4.6.1	Householder Reflections to Form the QR Factorization	237
4.6.2	Givens Rotations to Form the QR Factorization	238
4.6.3	Gram-Schmidt Transformations to Form the QR Factorization	238
4.7	Factorizations of Nonnegative Definite Matrices	238
4.7.1	Square Roots	239
4.7.2	Cholesky Factorization	240
4.7.3	Factorizations of a Gramian Matrix	242
4.8	Approximate Matrix Factorization	243
4.8.1	Nonnegative Matrix Factorization	244
4.8.2	Incomplete Factorizations	245
	Appendix: R Functions for Matrix Computations	245
	Exercises	247
5	Solution of Linear Systems	253
5.1	Condition of Matrices	254
5.1.1	Condition Number	255
5.1.2	Improving the Condition Number	260
5.1.3	Numerical Accuracy	260
5.2	Direct Methods for Consistent Systems	262
5.2.1	Gaussian Elimination and Matrix Factorizations	262
5.2.2	Choice of Direct Method	266
5.3	Iterative Methods for Consistent Systems	267
5.3.1	The Gauss-Seidel Method with Successive Overrelaxation	267
5.3.2	Conjugate Gradient Methods for Symmetric Positive Definite Systems	269
5.3.3	Multigrid Methods	273
5.4	Iterative Refinement	273
5.5	Updating a Solution to a Consistent System	274
5.6	Overdetermined Systems; Least Squares	276
5.6.1	Least Squares Solution of an Overdetermined System	278
5.6.2	Least Squares with a Full Rank Coefficient Matrix	280
5.6.3	Least Squares with a Coefficient Matrix Not of Full Rank	281
5.6.4	Weighted Least Squares	282
5.6.5	Updating a Least Squares Solution of an Overdetermined System	283
5.7	Other Solutions of Overdetermined Systems	284

5.7.1	Solutions that Minimize Other Norms of the Residuals	284
5.7.2	Regularized Solutions	287
5.7.3	Other Restrictions on the Solutions	289
5.7.4	Minimizing Orthogonal Distances	290
	Appendix: R Functions for Solving Linear Systems	293
	Exercises	294
6	Evaluation of Eigenvalues and Eigenvectors	297
6.1	General Computational Methods	298
6.1.1	Numerical Condition of an Eigenvalue Problem	298
6.1.2	Eigenvalues from Eigenvectors and Vice Versa	299
6.1.3	Deflation	300
6.1.4	Preconditioning	302
6.1.5	Shifting	302
6.2	Power Method	303
6.3	Jacobi Method	305
6.4	QR Method	308
6.5	Krylov Methods	310
6.6	Generalized Eigenvalues	311
6.7	Singular Value Decomposition	312
	Exercises	314
7	Real Analysis with Vectors and Matrices	317
7.1	Basics of Differentiation	319
7.2	Types of Differentiation	322
7.2.1	Differentiation with Respect to a Scalar	322
7.2.2	Differentiation with Respect to a Vector	323
7.2.3	Differentiation with Respect to a Matrix	328
7.3	Integration	331
7.3.1	Multidimensional Integrals and Integrals Involving Vectors and Matrices	331
7.3.2	Change of Variables; The Jacobian	332
7.3.3	Integration Combined with Other Operations	333
7.4	Multivariate Probability Theory	333
7.4.1	Random Variables and Probability Distributions	334
7.4.2	Distributions of Transformations of Random Variables	338
7.4.3	The Multivariate Normal Distribution	339
7.4.4	Distributions Derived from the Multivariate Normal	343
7.4.5	Chi-Squared Distributions	344
7.4.6	Wishart Distributions	347
7.5	Multivariate Random Number Generation	348
	Appendix: R for Working with Probability Distributions and for Simulating Random Data	350
	Exercises	353

Part II Applications in Statistics and Data Science

8	Special Matrices and Operations Useful in Modeling and Data Science	359
8.1	Data Matrices and Association Matrices	360
8.1.1	Flat Files	360
8.1.2	Graphs and Other Data Structures	361
8.1.3	Term-by-Document Matrices	368
8.1.4	Probability Distribution Models	369
8.1.5	Derived Association Matrices	369
8.2	Symmetric Matrices and Other Unitarily Diagonalizable Matrices	370
8.2.1	Some Important Properties of Symmetric Matrices	370
8.2.2	Approximation of Symmetric Matrices and an Important Inequality	371
8.2.3	Normal Matrices	374
8.3	Nonnegative Definite Matrices; Cholesky Factorization	376
8.4	Positive Definite Matrices	378
8.5	Idempotent and Projection Matrices	382
8.5.1	Idempotent Matrices	383
8.5.2	Projection Matrices: Symmetric Idempotent Matrices	388
8.6	Special Matrices Occurring in Data Analysis	389
8.6.1	Gramian Matrices	389
8.6.2	Projection and Smoothing Matrices	391
8.6.3	Centered Matrices and Variance-Covariance Matrices	395
8.6.4	The Generalized Variance	398
8.6.5	Similarity Matrices	400
8.6.6	Dissimilarity Matrices	401
8.7	Nonnegative and Positive Matrices	401
8.7.1	Properties of Square Positive Matrices	403
8.7.2	Irreducible Square Nonnegative Matrices	404
8.7.3	Stochastic Matrices	409
8.7.4	Leslie Matrices	410
8.8	Other Matrices with Special Structures	410
8.8.1	Helmert Matrices	410
8.8.2	Vandermonde Matrices	411
8.8.3	Hadamard Matrices and Orthogonal Arrays	412
8.8.4	Toeplitz Matrices	414
8.8.5	Circulant Matrices	415
8.8.6	Fourier Matrices and the Discrete Fourier Transform	416
8.8.7	Hankel Matrices	420
8.8.8	Cauchy Matrices	421
8.8.9	Matrices Useful in Graph Theory	421
8.8.10	Z-Matrices and M-Matrices	425
	Exercises	425

9	Selected Applications in Statistics	429
9.1	Linear Models	430
9.1.1	Fitting the Model	433
9.1.2	Least Squares Fit of Full Rank Models	435
9.1.3	Least Squares Fits of Non-Full Rank Models	436
9.1.4	Computing the Solution	439
9.1.5	Properties of a Least Squares Fit	441
9.1.6	Linear Least Squares Subject to Linear Equality Constraints	443
9.1.7	Weighted Least Squares	443
9.1.8	Updating Linear Regression Statistics	444
9.1.9	Linear Smoothing	447
9.1.10	Multivariate Linear Models	448
9.2	Statistical Inference in Linear Models	449
9.2.1	The Probability Distribution of ϵ	450
9.2.2	Estimability	453
9.2.3	The Gauss-Markov Theorem	455
9.2.4	Testing Linear Hypotheses	456
9.2.5	Estimating the Variance of the Errors	457
9.2.6	Full-Rank Models; Tests and Confidence Intervals of Individual Coefficients	457
9.2.7	Statistical Inference for Multivariate Linear Models	457
9.3	Principal Components	459
9.3.1	Principal Components of a Random Vector	460
9.3.2	Principal Components of Data	461
9.4	Condition of Models and Data	463
9.4.1	Ill-Conditioning in Statistical Applications	464
9.4.2	Variable Selection	464
9.4.3	Principal Components Regression	466
9.4.4	Shrinkage Estimation	466
9.4.5	Statistical Inference about the Rank of a Matrix	469
9.4.6	Incomplete Data	473
9.5	Stochastic Processes	476
9.5.1	Markov Chains	476
9.5.2	Markovian Population Models	479
9.5.3	Autoregressive Processes	481
9.6	Optimization of Scalar-Valued Functions	483
9.6.1	Stationary Points of Functions	485
9.6.2	Newton's Method	485
9.6.3	Least Squares	487
9.6.4	Maximum Likelihood	491
9.6.5	Optimization of Functions with Constraints	493
9.6.6	Optimization without Differentiation	498
	Appendix: R for Applications in Statistics	499
	Exercises	506

Part III Numerical Methods and Software

10 Numerical Methods	517
10.1 Digital Representation of Numeric Data	522
10.1.1 The Fixed-Point Number System	523
10.1.2 The Floating-Point Model for Real Numbers	525
10.1.3 Language Constructs for Representing Numeric Data	532
10.1.4 Other Variations in the Representation of Data; Portability of Data	537
10.2 Computer Operations on Numeric Data	539
10.2.1 Fixed-Point Operations	540
10.2.2 Floating-Point Operations	541
10.2.3 Language Constructs for Operations on Numeric Data	546
10.2.4 Software Methods for Extending the Precision	548
10.2.5 Exact Computations	550
10.3 Numerical Algorithms and Analysis	551
10.3.1 Error in Numerical Computations	552
10.3.2 Efficiency	560
10.3.3 Iterations and Convergence	565
10.3.4 Computations without Storing Data	568
10.3.5 Other Computational Techniques	569
Appendix: Numerical Computations in R	572
Exercises	573
11 Numerical Linear Algebra	579
11.1 Computer Storage of Vectors and Matrices	579
11.2 General Computational Considerations for Vectors and Matrices	581
11.2.1 Relative Magnitudes of Operands	581
11.2.2 Iterative Methods	583
11.2.3 Assessing Computational Errors	584
11.3 Multiplication of Vectors and Matrices	585
11.4 Other Matrix Computations	589
Exercises	592
12 Software for Numerical Linear Algebra	595
12.1 General Considerations	596
12.1.1 Software Design	598
12.1.2 Software Development, Maintenance, and Testing	607
12.1.3 Reproducible Research	610
12.2 Software Libraries	611
12.2.1 BLAS	612
12.2.2 Level 2 and Level 3 BLAS, LAPACK, and Related Libraries	614
12.2.3 Libraries for High Performance Computing	615
12.2.4 The IMSL Libraries	619
12.3 General Purpose Languages and Programming Systems	621

12.3.1 Programming Considerations	623
12.3.2 Modern Fortran	625
12.3.3 C and C++	627
12.3.4 Python	628
12.3.5 MATLAB and Octave	628
12.3.6 R; Additional Comments	630
Exercises	633

Appendices and Back Matter

A Notation and Definitions	639
A.1 General Notation	639
A.2 Computer Number Systems	641
A.3 General Mathematical Functions and Operators	642
A.4 Linear Spaces and Matrices	645
A.5 Models and Data	652
Bibliography	653
Index	665