Contents

Part I Linear Algebra

1 Linear Equations 3
1.1 Row Reduction Methods 4
1.1.1 Gaussian Elimination 5
1.1.2 Pivoting 7
1.1.3 R Code: Gaussian Elimination and Backward Substitution 8
1.1.4 Gauss-Jordan Elimination 9
1.1.5 LU Decomposition 10
1.1.6 R Code: LU Decomposition 14
1.1.7 Cholesky Decomposition 15
1.1.8 R Code: Cholesky Decomposition of a Symmetric Matrix 18
1.2 Matrix Methods 19
1.2.1 Determinant 19
1.2.2 R Code: Determinant 21
1.2.3 Determinants and Linear Dependencies 21
1.2.4 R Code: Reduced Row Echelon Form and Linear Dependencies 22
1.2.5 Using the Determinant to Solve Linear Equations 23
1.2.6 R Code: Cramer's Rule 24
1.2.7 Matrix Inverse 24
1.2.8 R Code: Calculate Inverse Using Reduced Row Echelon Form 26
1.2.9 Norms, Errors, and the Condition Number of a Matrix 26
1.2.10 R Code: Condition Number and Norm Ratio 33

1.3 Iterative Methods . . . 34
1.3.1 Jacobi's Method . . . 34
1.3.2 Gauss-Seidel Method . . . 35
1.3.3 Convergence . . . 36
1.3.4 R Code: Gauss-Seidel . . . 37
1.4 Chapter Summary . . . 38
References . . . 38

2 Least Squares Estimation . . . 39
2.1 Line of Best Fit . . . 39
2.1.1 Deriving a Line of Best Fit . . . 39
2.1.2 Minimizing the Sum of Squared Differences . . . 41
2.1.3 Normal Equations . . . 42
2.1.4 Analytic Solution . . . 42
2.2 Solving the Normal Equations . . . 43
2.2.1 The QR Decomposition . . . 43
2.2.2 Advantages of an Orthonormal System . . . 44
2.2.3 Hat Matrix . . . 45
2.2.4 Coefficients . . . 47
2.2.5 Summary . . . 48
2.2.6 R Code: QR Solver . . . 49
2.3 Performing the QR Decomposition . . . 49
2.3.1 Gram-Schmidt Orthogonalization . . . 49
2.3.2 R Code: QR Decomposition; Gram-Schmidt Orthogonalization . . . 53
2.3.3 Givens Rotations . . . 54
2.3.4 R Code: QR Decomposition; Givens Rotations . . . 58
2.3.5 Householder Reflections . . . 58
2.3.6 R Code: QR Decomposition; Householder Reflectors . . . 61
2.3.7 Comparing the Decompositions . . . 61
2.3.8 R Code: QR Decomposition Comparison . . . 62
2.4 Linear Regression and its Assumptions . . . 62
2.4.1 Linearity . . . 63
2.4.2 Nature of the Variables . . . 64
2.4.3 Errors and their Distribution . . . 65
2.4.4 Regression Coefficients . . . 67
2.5 OLS Estimation and the Gauss-Markov Theorem . . . 67
2.5.1 Proving the OLS Estimates are Unbiased . . . 68
2.5.2 Proving the OLS Estimates are Efficient . . . 69
2.6 Maximum Likelihood Estimation . . . 71
2.6.1 Log Likelihood Function . . . 71
2.6.2 R Code: Maximum Likelihood Estimation . . . 74
2.7 Beyond OLS Estimation . . . 74
2.8 Chapter Summary . . . 75
References . . . 76

3 Linear Regression . . . 77
3.1 Simple Linear Regression . . . 77
3.1.1 Inspecting the Residuals . . . 79
3.1.2 Describing the Model's Fit to the Data . . . 80
3.1.3 Testing the Model's Fit to the Data . . . 80
3.1.4 Variance Estimates . . . 81
3.1.5 Tests of Significance . . . 82
3.1.6 Confidence Intervals . . . 83
3.1.7 R Code: Confidence Interval Simulation . . . 83
3.1.8 Confidence Regions . . . 83
3.1.9 Forecasting . . . 84
3.1.10 R Code: Simple Linear Regression . . . 87
3.1.11 R Code: Simple Linear Regression: Graphs . . . 88
3.2 Multiple Regression . . . 89
3.2.1 Regression Model . . . 90
3.2.2 Regression Coefficients . . . 92
3.2.3 Variance Estimates, Significance Tests, and Confidence Intervals . . . 94
3.2.4 Model Comparisons and Changes in R^2 . . . 95
3.2.5 Comparing Predictors . . . 97
3.2.6 Forecasting . . . 98
3.2.7 R Code: Multiple Regression . . . 99
3.3 Polynomials, Cross-Products, and Categorical Predictors . . . 99
3.3.1 Polynomial Regression . . . 100
3.3.2 R Code: Polynomial Regression . . . 105
3.3.3 Cross-Product Terms . . . 105
3.3.4 R Code: Cross-Product Terms and Simple Slopes . . . 109
3.3.5 Johnson-Neyman Procedure . . . 110
3.3.6 R Code: Johnson-Neyman Procedure . . . 111
3.3.7 Categorical Predictors . . . 111
3.3.8 R Code: Contrast Codes for Categorical Predictors . . . 113
3.3.9 Summary . . . 114
3.4 Chapter Summary . . . 114
References . . . 115

4 Eigen Decomposition . . . 117
4.1 Diagonalization . . . 117
4.1.1 Eigenvector Multiplication . . . 117
4.1.2 The Characteristic Equation . . . 119
4.1.3 R Code: Eigen Decomposition of a 2×2 Matrix with Real Eigenvalues . . . 121
4.1.4 Properties of a Diagonalized Matrix . . . 121
4.2 Eigenvalue Calculation . . . 122
4.2.1 Basic QR Algorithm . . . 122
4.2.2 R Code: QR Algorithm Using Gram-Schmidt Orthogonalization . . . 124

4.2.3 Improving the QR Algorithm . . . 124
4.2.4 R Code: Hessenberg Form . . . 126
4.2.5 R Code: Shifted QR Algorithm . . . 129
4.2.6 Francis (Implicitly-Shifted QR) Algorithm . . . 129
4.2.7 R Code: Francis Bulge Chasing Algorithm (Single-Shift) . . . 133
4.3 Eigenvector Calculation . . . 133
4.3.1 R Code: Eigenvector Calculation Using LU Decomposition . . . 135
4.4 Dynamical Systems . . . 135
4.4.1 Matrix Power . . . 135
4.4.2 R Code: Matrix Power Using Eigen Decomposition . . . 136
4.4.3 Power Method for Dominant Eigen Pair . . . 136
4.4.4 Population Ecology . . . 137
4.4.5 Predator-Prey Model . . . 139
4.4.6 Markov Chains . . . 141
4.4.7 R Code: Power Method and Applications . . . 144
4.5 Schur Decomposition . . . 144
4.5.1 Compute an Initial Eigenvector . . . 145
4.5.2 Create an Orthonormal Basis . . . 145
4.5.3 Rotate . . . 146
4.5.4 Deflate and Continue Iterating . . . 146
4.5.5 R Code: Schur Decomposition . . . 147
4.6 Chapter Summary . . . 148
Reference . . . 148

5 Singular Value Decomposition . . . 149
5.1 Introduction . . . 149
5.1.1 Illustration . . . 150
5.1.2 Geometric Interpretation . . . 151
5.1.3 R Code: Singular Value Decomposition . . . 154
5.1.4 Matrix Properties . . . 155
5.1.5 Pseudoinverse . . . 155
5.1.6 Solving Linear Equations . . . 155
5.1.7 R Code: Matrix Rank and Pseudoinverse . . . 156
5.2 Calculating the SVD . . . 157
5.2.1 Description . . . 157
5.2.2 Calculations . . . 157
5.2.3 R Code: One-Sided Jacobi Algorithm . . . 160
5.3 Data Reduction and Image Compression . . . 161
5.3.1 R Code: Image Compression Using Singular Value Decomposition . . . 162
5.4 Principal Components Analysis . . . 163
5.4.1 Mechanics . . . 163

5.4.2 R Code: Principal Components Analysis . . . 166
5.4.3 Total Least Squares . . . 167
5.4.4 R Code: Total Least Squares . . . 170
5.4.5 Dimension Reduction . . . 170
5.4.6 R Code: Principal Components Analysis of Cereal Data . . . 176
5.4.7 Data Construction . . . 176
5.4.8 R Code: Data Construction . . . 177
5.5 Collinearity . . . 177
5.5.1 Using the SVD to Detect Collinearity . . . 178
5.5.2 R Code: Collinearity Detection . . . 182
5.5.3 Principal Components Regression . . . 182
5.5.4 R Code: Principal Components Regression of (Fictitious) NFL Data . . . 183
5.6 Chapter Summary . . . 185
References . . . 186

Part II Bias and Efficiency

6 Generalized Least Squares Estimation . . . 189
6.1 Gauss–Markov Violations . . . 189
6.1.1 R Code: Simulations for Fig. 6.1 . . . 191
6.2 Generalized Least Squares . . . 192
6.2.1 R Code: OLS Estimation as GLS Estimation . . . 193
6.2.2 OLS and GLS . . . 195
6.2.3 R Code: Generalized Least Squares Estimation . . . 196
6.3 Heteroscedasticity and Feasible Weighted Least Squares . . . 196
6.3.1 Assessing Heteroscedasticity . . . 196
6.3.2 R Code: Breusch–Pagan Test of Heteroscedasticity . . . 198
6.3.3 Feasible Weighted Least Squares . . . 198
6.3.4 R Code: Feasible Weighted Least Squares . . . 199
6.3.5 Heteroscedasticity Consistent Covariance Matrix . . . 201
6.3.6 Summary . . . 202
6.3.7 R Code: Heteroscedasticity Consistent Covariance Matrix . . . 202
6.3.8 Confidence Interval Simulation . . . 203
6.3.9 R Code: Heteroscedasticity Confidence Interval Simulation . . . 204
6.4 Autocorrelated Errors . . . 204
6.4.1 Mathematical Representation . . . 205
6.4.2 Covariance Matrix . . . 206
6.4.3 Detecting Autocorrelations . . . 206
6.4.4 R Code: Detecting Autocorrelations . . . 208
6.4.5 Accommodating Autocorrelated Errors . . . 208
6.4.6 Feasible Generalized Least Squares . . . 210
6.4.7 R Code: Feasible Generalized Least Squares . . . 211

6.4.8 Autocorrelation Consistent Covariance Matrix 212
6.4.9 Summary 212
6.4.10 R Code: Autocorrelation Consistent Covariance Matrix 215
6.5 Chapter Summary 216
References 217

7 Robust Regression 219
7.1 Assessing Normality 220
7.1.1 Tests of Normality 221
7.1.2 R Code: Assessing the Normality of the Residuals 222
7.1.3 Influence and Normality 222
7.1.4 Leverage and Influence 223
7.1.5 Cook's D 224
7.1.6 Illustration 226
7.1.7 Handling Influential Observations 227
7.1.8 R Code: Cook's D 228
7.2 Robust Estimators and Influential Observations 228
7.2.1 Breakdown Point 229
7.2.2 Efficiency 229
7.2.3 R Code: Robust Regression Simulation 230
7.3 Resistant Estimation 231
7.3.1 Least Absolute Regression 231
7.3.2 R Code: Least Absolute Regression 232
7.3.3 Least Median of Squares 233
7.3.4 R Code: Least Median of Squares 235
7.4 M Estimation 236
7.4.1 Weighting Methods 236
7.4.2 Algorithm 237
7.4.3 R Code: Robust Regression with M Estimation 239
7.5 Bootstrapped Confidence Intervals 239
7.5.1 Case Resampling vs. Residual Resampling 240
7.5.2 Confidence Intervals 241
7.5.3 R Code: Bootstrapping with Robust Regression (M Estimation) 243
7.6 MM Estimation 244
7.6.1 S Estimation 244
7.6.2 R Code: S Estimation (Part 1) 246
7.6.3 R Code: MM Estimation (compact form with sub functions) 247
7.6.4 Application 248
7.6.5 R Code: Robust Regression of Star Data 250
7.7 Concluding Remarks 251
7.8 Chapter Summary 251
References 252

8 Model Selection and Biased Estimation 253
8.1 Prediction Error and Model Complexity 253
8.1.1 Prediction Errors and the Bias-Variance Tradeoff 254
8.1.2 Cross-Validation 255
8.1.3 Information Criteria Measures and Model Selection 257
8.1.4 R Code: Cross Validation and Information Criteria Measures 258
8.2 Subset Selection Methods 259
8.2.1 Stepwise Regression 259
8.2.2 R Code: Fictitious Data Predicting College Performance 261
8.2.3 R Code: Stepwise Regression 265
8.2.4 Best Subset Regression 266
8.2.5 R Code: Sweep Operator 268
8.2.6 R Code: Sweep Operator for Best Subset Regression 270
8.2.7 Branch and Bound Algorithm 271
8.2.8 R Code: Branch and Bound (Compact Form) 273
8.2.9 Comparing the Models 274
8.2.10 R Code: Model Comparison 274
8.3 Shrinkage Estimators and Regularized Regression 276
8.3.1 Ridge Regression 276
8.3.2 R Code: Ridge Regression: Augmented Matrix Method 279
8.3.3 R Code: Ridge Regression 281
8.3.4 Lasso 281
8.3.5 R Code: LASSO 284
8.4 Comparing the Methods 285
8.5 Chapter Summary 286
References 287

9 Cubic Splines and Additive Models 289
9.1 Piecewise Polynomials and Regression Splines 289
9.1.1 Truncated Power Basis 291
9.1.2 Natural Cubic Spline 293
9.1.3 R Code: Truncated Power Series and Natural Cubic Spline Bases 294
9.1.4 B Spline Basis 294
9.1.5 R Code: B-Spline Basis 298
9.1.6 Bias-Variance Trade-Off 299
9.2 Penalized Smoothing Splines 299
9.2.1 Reinsch Form 300

9.2.2 R Code: Penalized Smoothing Spline: Reinsch Form . . . 302
9.2.3 P-Splines . . . 302
9.2.4 Statistical Inference and Confidence Intervals . . . 305
9.2.5 Comparing Penalized Smoothing Splines and Regression Splines . . . 307
9.2.6 R Code: P-Spline . . . 309
9.3 Additive Models . . . 309
9.3.1 Fitting an Additive Model . . . 310
9.3.2 Backfitting . . . 311
9.3.3 Partial Slopes . . . 312
9.3.4 Model Selection and Inference . . . 313
9.3.5 R Code: Additive Model: Backfitting . . . 314
9.3.6 Penalized Least Squares . . . 315
9.3.7 R Code: Additive Model: Penalized Least Squares . . . 317
9.4 Chapter Summary . . . 318
References . . . 319

Part III Nonlinear Models

10 Nonlinear Regression and Optimization . . . 323
10.1 Comparing Linear and Nonlinear Models . . . 323
10.1.1 Model Representation . . . 323
10.1.2 Partial Derivatives . . . 325
10.1.3 Parameter Estimation . . . 325
10.1.4 Standard Errors, Parameter Interpretation, and Degrees of Freedom . . . 325
10.1.5 Variety . . . 326
10.2 Root Finding Algorithms . . . 326
10.2.1 Newton's Method . . . 328
10.2.2 Secant Method . . . 329
10.2.3 R Code: Root-Finding Algorithms . . . 330
10.3 Optimization . . . 331
10.3.1 Exponential Growth Model . . . 331
10.3.2 Newton-Raphson . . . 333
10.3.3 R Code: Newton-Raphson . . . 335
10.3.4 Fisher's Method of Scoring . . . 335
10.3.5 R Code: Fisher's Method of Scoring . . . 337
10.3.6 Gauss-Newton . . . 338
10.3.7 R Code: Gauss-Newton . . . 340
10.3.8 BFGS Algorithm . . . 340
10.3.9 R Code: BFGS . . . 343
10.3.10 Nelder-Mead . . . 344
10.3.11 R Code: Nelder-Mead (Compact Form) . . . 348
10.3.12 Summary . . . 349

10.4 Missing Observations . . . 349
10.4.1 Classifying Missing Data . . . 349
10.4.2 Maximum Likelihood Estimation and the Expectation-Maximization Algorithm . . . 350
10.4.3 Bivariate Example . . . 351
10.4.4 Multivariate Illustration . . . 354
10.4.5 R Code: EM Algorithm for Multivariate Normal with Missing Data . . . 356
10.4.6 Multiple Regression with Missing Observations . . . 357
10.4.7 R Code: EM Regression with Bootstrapped Standard Errors . . . 358
10.5 Chapter Summary . . . 359
References . . . 359

11 Generalized Linear Models . . . 361
11.1 Generalized Linear Models . . . 361
11.1.1 Log Likelihood Functions . . . 362
11.1.2 Components of a Generalized Linear Model . . . 364
11.1.3 Iteratively Reweighted Least Squares Estimation . . . 365
11.1.4 Canonical Link . . . 367
11.1.5 R Code: IRLS Estimation for GLM with Canonical Links . . . 368
11.2 Poisson Distribution . . . 369
11.2.1 Estimation . . . 370
11.2.2 Deviance and Goodness of Fit . . . 371
11.2.3 Goodness of Fit . . . 372
11.2.4 Regression Coefficients and Fitted Values . . . 373
11.2.5 Standard Errors, Tests of Significance, and Confidence Intervals . . . 374
11.2.6 R Code: GLM Fit . . . 376
11.2.7 R-Code: GLM: Profile Likelihood . . . 377
11.2.8 Diagnostics . . . 378
11.2.9 Overdispersion and Quasi-Likelihood Estimation . . . 379
11.2.10 R-Code: GLM Residuals . . . 381
11.3 Binomial Distribution . . . 381
11.3.1 Overview . . . 381
11.3.2 GLM with a Binomial Distribution . . . 382
11.3.3 Goodness of Fit . . . 382
11.3.4 Interpreting the Fitted Values and Regression Coefficients . . . 384
11.3.5 Standard Errors and Confidence Intervals . . . 386
11.3.6 Extensions . . . 386
11.3.7 R Code: GLM: Binomial Distribution with Logit Link . . . 387

11.4 Gamma Distribution 387
11.4.1 Properties of a Gamma Distribution 387
11.4.2 R Code: Gamma Distribution Maximum Likelihood Estimation 391
11.4.3 Gamma GLM with Canonical Link 392
11.4.4 R Code: GLM: Gamma Distribution 394
11.4.5 Gamma GLM with Non Canonical Links 395
11.4.6 R Code: GLM: Gamma Distribution 396
11.5 Chapter Summary 397
References 398

12 Survival Analysis 399
12.1 Overview 399
12.1.1 Censoring 399
12.1.2 Statistical Functions 400
12.1.3 Statistical Models 402
12.2 Nonparameteric Model 403
12.2.1 Kaplan-Meier Estimator 404
12.2.2 Standard Errors 405
12.2.3 Confidence Intervals 405
12.2.4 Median Survival 406
12.2.5 Hazard Rate 407
12.2.6 R Code: Kaplan-Meier Estimator (Log-Log Confidence Intervals) 408
12.2.7 Log-Rank Test 409
12.2.8 R Code: Log Rank Test 412
12.2.9 R Code: Log Rank Test cont. 413
12.3 Semiparametric Model 413
12.3.1 Hazard Function 413
12.3.2 Preliminary Example Without Ties 414
12.3.3 Interpreting the Hazard Ratio 415
12.3.4 Partial Likelihood Function 415
12.3.5 Goodness of Fit 416
12.3.6 Diagnostics 419
12.3.7 R Code: Cox Regression–No Ties/Single Predictor 421
12.3.8 Handling Ties 421
12.3.9 R Code: Cox Regression (Compact Form) 426
12.3.10 Residuals When Ties are Present 427
12.3.11 Interpretation 428
12.3.12 R Code: Cox Regression Residuals (Compact Form) 429
12.4 Parametric Model 430
12.4.1 Properties of a Weibull Distribution 430
12.4.2 Assessing the Appropriateness of a Weibull Distribution 432

12.4.3 R Code: Weibull MLE 433
12.4.4 Weibull Regression 433
12.4.5 Interpreting the AFT Coefficients 434
12.4.6 Model Fit and Standard Errors 437
12.4.7 R Code: Weibull Regression Single Predictor 438
12.4.8 Conversions to Proportional Hazard 439
12.4.9 R Code: Weibull AFT and PH Conversion 440
12.4.10 Diagnostics 440
12.4.11 Diagnostic Plots 441
12.4.12 R Code: Weibull AFT Diagnostics 445
12.5 For Further Reading 446
12.6 Chapter Summary 446
References 447

13 Time Series Analysis 449
13.1 Overview 449
13.1.1 Dynamics of a Time Series 450
13.1.2 Stationarity and Differencing 452
13.1.3 From ARIMA to ARMA 456
13.1.4 R Code: Stationarity and Differencing 456
13.2 Autocorrelations 457
13.2.1 Autocorrelation Function 457
13.2.2 Partial Autocorrelations 458
13.2.3 R Code: Autocorrelation and Partial Autocorrelation Function 459
13.3 Moving Averages and Autoregressive Processes 459
13.3.1 Moving Averages 460
13.3.2 Autoregressive Processes 461
13.3.3 Wold Representation 462
13.3.4 ARMA(p,q) 463
13.3.5 Simulations and Filters 465
13.3.6 R Code: ARMA Simulations and Wold Representation 467
13.4 Model Identification 468
13.4.1 Moving Averages 468
13.4.2 Autoregressive Processes 471
13.4.3 Mixed Models 473
13.4.4 R Code: Stationarity, Invertibility, and Redundancy 477
13.5 Model Estimation 477
13.5.1 State-Space Representation 478
13.5.2 Kalman Filter 479
13.5.3 R Code: State Space Representation and Kalman Filter 482
13.5.4 R Code: ARMA Estimation 483

13.6 Model Adequacy 484
13.6.1 Test Residuals for Dependencies 484
13.6.2 R Code: Box-Pierce & Ljung-Box Tests of Model Fit 484
13.6.3 Models Comparisons 485
13.6.4 R Code: Tests of Model Fit 485
13.7 Forecasting 486
13.7.1 Generating Predicted Values 486
13.7.2 R Code: Forecasting (Mean–Centered) Observations 487
13.8 Integregation 488
13.8.1 Model Identification 488
13.8.2 Model Estimation 488
13.8.3 R Code: Dyads 492
13.9 Chapter Summary 493
References 494

14 Mixed-Effects Models 495
14.1 Overview 495
14.1.1 Mixed–Effects Models in Matrix Form 496
14.1.2 Estimation and Prediction 498
14.1.3 Mixed Model Equations 498
14.1.4 Variance Components Estimation 499
14.1.5 Modeling the Random Terms 499
14.1.6 R Code: Preliminary Functions for Mixed Modeling 501
14.2 Understanding Mixed–Effects Models 502
14.2.1 Creating a Simulation 502
14.2.2 Plotting the Data 503
14.2.3 Testing Hierarchical Models 503
14.2.4 Intraclass Correlation 505
14.2.5 Fixed Effects and Random Effects 507
14.2.6 Interpreting the Random Coefficients 507
14.2.7 Standard Errors 510
14.2.8 Adding a Random Slope 511
14.2.9 Diagnostics 512
14.2.10 Summary 512
14.2.11 R Code: Mixed Model Simulations 513
14.3 Estimation 514
14.3.1 R Code: Mixed Model Fit 515
14.3.2 Illustration 516
14.3.3 R Code: Mixed Model Estimation 517
14.4 Repeated Measures and Growth Curve Models 518
14.4.1 Growth Curve Model 518
14.4.2 Reshaping the Data 519

14.4.3 Slopes and Intercepts 519
14.4.4 Model Comparisons 520
14.4.5 R Code: Growth Curve Analysis 524
14.5 Chapter Summary 525
References 526