

CONTENTS

List of Figures xvii

List of Tables xxi

Preface xxiv

PART I ORDINARY LEAST SQUARES 1

Chapter 1 The Least-Squares Linear Fit 3

- 1.1 Earnings and Attributes of Workers 3
- 1.2 Generalizing the Sample Average 7
- 1.3 OLS Regression 14
- 1.4 Overview 15
- 1.5 Exercises 16
- 1.6 Appendix: Data Collection 17

Chapter 2 The Geometry of Least Squares 19

- 2.1 Introductory Example 19
- 2.2 Ordinary Least Squares 21
- 2.3 Examples of OLS 25
- 2.4 Orthogonal Projection 28
 - 2.4.1 The Projection Theorem 31
 - 2.4.2 Orthogonal Projectors 31
- 2.5 Exact Multicollinearity 34
- 2.6 Mathematical Notes 36
 - 2.6.1 Gram–Schmidt Orthonormalization 36
 - 2.6.2 Properties of Orthogonal Projectors 37
 - 2.6.3 Proofs 38
- 2.7 Overview 40
- 2.8 Exercises 41
 - 2.8.1 Review 41
 - 2.8.2 Extensions 44

Chapter 3 Partitioned Fit 47

- 3.1 Introductory Example 47
- 3.2 Partitioned Fit 54
- 3.3 Projection 61
- 3.4 Projectors 66
- 3.5 Overview 68
- 3.6 Exercises 69
 - 3.6.1 Review 69
 - 3.6.2 Extensions 72

Chapter 4 Restricted Least Squares 74

- 4.1 Introduction 74
- 4.2 Linear Restrictions 77
- 4.3 Restricted Least Squares 79
- 4.4 Generalized Distance 84
 - 4.4.1 Translation 87
- 4.5 Generalized Projection 88
- 4.6 Overview 91
- 4.7 Exercises 92
 - 4.7.1 Review 92
 - 4.7.2 Extensions 94

Chapter 5 Overview of Ordinary Least Squares 97

- 5.1 Geometric Theory 97
- 5.2 Econometric Specifications 99
- 5.3 Econometric Method 100

PART II LINEAR REGRESSION 101

Chapter 6 Linear Unbiased Estimation 105

- 6.1 Experimental Example 105
- 6.2 First Moments 110
- 6.3 Conditional Means 113
- 6.4 Projection of Random Variables 114
- 6.5 Mathematical Notes 118
- 6.6 Methodological Notes 120
- 6.7 Overview 121
- 6.8 Exercises 122
 - 6.8.1 Review 122
 - 6.8.2 Extensions 123

Chapter 7 Variances and Covariances 125

- 7.1 Introduction 125
- 7.2 Second Moments 129

7.3	Spherical Distributions	131
7.4	The Variance Ellipse	133
7.5	Minimum MSE Linear Prediction	135
7.6	Mathematical Notes	140
	7.6.1 A Square Root of the Variance Matrix	140
	7.6.2 The Cauchy–Schwarz Inequality	143
	7.6.3 Linear Transformation of Variance Ellipses	144
	7.6.4 A Quadratic Decomposition	146
7.7	Overview	148
7.8	Exercises	149
	7.8.1 Review	149
	7.8.2 Extensions	152
Chapter 8	Variances and Covariances of Ordinary Least Squares	154
8.1	Experimental Example	154
8.2	Second-Moment Properties	157
8.3	Variance and Covariance Matrices	159
8.4	Estimation of the Variance Parameter	163
8.5	Methodological Note	167
8.6	Overview	168
8.7	Exercises	168
	8.7.1 Review	168
	8.7.2 Extensions	170
Chapter 9	Efficient Estimation	173
9.1	Introduction	173
9.2	Design and Precision	175
	9.2.1 Dispersion	177
	9.2.2 Sample Size	177
	9.2.3 Near Multicollinearity	178
	9.2.4 Forecast Variance	180
9.3	Restricted Estimation	182
9.4	The Gauss–Markov Theorem	186
	9.4.1 Geometry of the Gauss–Markov Theorem	187
	9.4.2 Proof of the Gauss–Markov Theorem	188
9.5	Mathematical Notes	189
9.6	Overview	190
9.7	Exercises	191
	9.7.1 Review	191
	9.7.2 Extensions	193
Chapter 10	Normal Distribution Theory	195
10.1	Introduction	195
10.2	Distribution Theory for OLS Estimators	198

10.3	Interval Estimators	200
10.3.1	Variance	200
10.3.2	Coefficient Vector with Known Variance	201
10.3.3	Coefficient Vector with Unknown Variance	203
10.3.4	Linear Functions of Coefficients	204
10.4	Efficiency of OLS	205
10.5	Basic Distribution Theory	205
10.5.1	The Multivariate Normal Distribution	206
10.5.2	The Chi-Square and F Distributions	210
10.5.3	Singular Variances and Generalized Inverses	211
10.5.4	Singular Multivariate Normal Distributions	214
10.6	Methodological Notes	215
10.7	Overview	216
10.8	Exercises	217
10.8.1	Review	217
10.8.2	Extensions	219
Chapter 11	Hypothesis Testing	222
11.1	Introduction	222
11.2	Hypothesis Testing	224
11.3	Statistical Power	228
11.3.1	Power Comparisons for Tests	229
11.3.2	t Statistics	229
11.3.3	Optimal Power and the F Test	231
11.4	Basic Distribution Theory	232
11.5	Methodological Notes	236
11.6	Overview	237
11.7	Exercises	238
11.7.1	Review	238
11.7.2	Extensions	239
Chapter 12	Overview of Linear Regression	240
12.1	Statistical Theory	240
12.2	Probability Distribution Theory	241
PART III	GENERALIZATIONS OF THE LINEAR MODEL	243
Chapter 13	Nonnormal Distribution Theory	245
13.1	Introduction	245
13.2	Nonnormal Parametric Distributions	246
13.2.1	The Student t Distribution	247
13.2.2	Laplace (Double Exponential) Distribution	249
13.2.3	Logistic Distribution	250
13.2.4	Power Exponential Distribution	250

13.3	LAD Estimation	251
	13.3.1 The Sample Median	252
	13.3.2 LAD Linear Regression	253
13.4	Asymptotic Distribution Theory	256
	13.4.1 Convergence in Distribution	259
	13.4.2 Law of Large Numbers	262
	13.4.3 Central Limit Theorem	265
	13.4.4 Sample Size	267
13.5	Mathematical Notes	270
	13.5.1 The Density of an Order Statistic	270
	13.5.2 Properties of LAD Fit	271
	13.5.3 Convergence Proofs	273
13.6	Overview	278
13.7	Exercises	279
	13.7.1 Review	279
	13.7.2 Extensions	281
Chapter 14 Maximum Likelihood Estimation 284		
14.1	Introduction	284
14.2	Probability Model Specification	285
14.3	The Likelihood Function	288
14.4	The Maximum Likelihood Estimator	293
14.5	Identification	295
14.6	The Score Function	300
14.7	The Information Matrix	302
14.8	The Cramér–Rao Lower Bound	305
14.9	Mathematical Notes	311
14.10	Overview	312
14.11	Exercises	314
	14.11.1 Review	314
	14.11.2 Extensions	315
Chapter 15 Maximum Likelihood Asymptotic Distribution Theory 318		
15.1	Introduction	318
15.2	Consistency	320
15.3	Asymptotic Normality	324
	15.3.1 Score	325
	15.3.2 Information	326
	15.3.3 Asymptotic Distribution	327
15.4	Variance Estimation	329
15.5	Efficiency	331
15.6	Linearized MLE	333
15.7	Restricted Estimation	334

15.8	Mathematical Notes	336
15.9	Overview	340
15.10	Exercises	342
	15.10.1 Review	342
	15.10.2 Extensions	343
Chapter 16 Maximum Likelihood Computation 347		
16.1	Introduction	347
16.2	Grid Search	349
16.3	Polynomial Approximation	349
16.4	Line Searches	351
	16.4.1 The Method of Steepest Ascent	353
	16.4.2 Quadratic Methods	355
	16.4.3 Quadratic Methods and the MLE	357
	16.4.4 LMLE	361
16.5	Convergence Criteria	362
16.6	Transformations of Parameters	364
16.7	Concentrating the Likelihood Function	368
16.8	The Gauss–Seidel Algorithm	371
16.9	Mathematical Notes	372
	16.9.1 Proofs	372
	16.9.2 Stochastic Order	374
	16.9.3 Uniqueness of the MLE	375
16.10	Overview	376
16.11	Exercises	377
	16.11.1 Review	377
	16.11.2 Extensions	379
Chapter 17 Maximum Likelihood Statistical Inference 380		
17.1	Introduction	380
17.2	The Classical Hypothesis Test Statistics	384
	17.2.1 The Wald Test	384
	17.2.2 The Score Test	385
	17.2.3 The Likelihood Ratio Test	388
	17.2.4 A Graphic Description of the Test Statistics	389
17.3	Asymptotic Distribution Theory	393
	17.3.1 The Likelihood Ratio Test	394
	17.3.2 The Score Test	395
	17.3.3 The Wald Test	395
	17.3.4 The $C(\alpha)$ Test	396
	17.3.5 Limiting Distribution	396
17.4	Parameter Transformations and Invariance	397
17.5	Power	402
	17.5.1 Local Power	403

- 17.5.2 Neyman–Pearson Lemma 406
- 17.6 Interval Estimation 408
- 17.7 Overview 409
- 17.8 Exercises 410
 - 17.8.1 Review 410
 - 17.8.2 Extensions 412

Chapter 18 Heteroskedasticity 416

- 18.1 Heteroskedasticity in Wages 417
- 18.2 Heteroskedasticity and OLS 421
- 18.3 Testing for Heteroskedasticity 423
 - 18.3.1 The Goldfeld–Quandt F Test 424
 - 18.3.2 The Breusch–Pagan Score Test 424
- 18.4 Adjustments to OLS 427
- 18.5 Heteroskedasticity and WLS/GLS 429
 - 18.5.1 Maximum Likelihood 433
 - 18.5.2 FGLS 435
 - 18.5.3 Adaptive Estimation 440
- 18.6 Methodological Notes 442
- 18.7 Mathematical Notes 443
 - 18.7.1 Score and Information 444
 - 18.7.2 The Maximum Likelihood Estimator 444
 - 18.7.3 The Breusch–Pagan Score Test 446
 - 18.7.4 Regularity 447
 - 18.7.5 Asymptotic Theory for Heteroskedasticity 448
- 18.8 Overview 451
- 18.9 Exercises 452
 - 18.9.1 Review 452
 - 18.9.2 Extensions 453

Chapter 19 Serial Correlation 455

- 19.1 The Phillips Curve 455
- 19.2 The Basic Autoregressive Model 458
 - 19.2.1 The Autocorrelation Function 458
 - 19.2.2 The Log-Likelihood Function 460
- 19.3 Autocorrelation and OLS 462
- 19.4 Testing for Autocorrelation 464
 - 19.4.1 Breusch–Godfrey Score Test 464
 - 19.4.2 The Durbin–Watson Test 466
- 19.5 Variance Estimation for OLS 466
- 19.6 Serial Correlation and GLS 468
 - 19.6.1 Maximum Likelihood Estimation 469
 - 19.6.2 FGLS 471
- 19.7 Prediction 471

19.8	Methodological Notes	472
19.9	Mathematical Notes	475
	19.9.1 Score and Information	475
	19.9.2 Breusch–Godfrey Score Test	476
	19.9.3 Asymptotic Distribution Theory	477
	19.9.4 OLS versus GLS	480
19.10	Overview	481
19.11	Exercises	482
	19.11.1 Review	482
	19.11.2 Extensions	483
Chapter 20	Instrumental Variables Estimation	486
20.1	The Phillips Curve Revisited	487
20.2	Latent Variable Models	491
20.3	Omitted Explanatory Variables	493
20.4	Consistent Estimation	499
20.5	Two-Stage Least Squares	502
20.6	Two-Step Variance Estimation	505
20.7	Efficiency	509
	20.7.1 Simultaneous Equations	509
	20.7.2 Dynamic Regression	511
	20.7.3 IV and GLS	512
20.8	Issues in Small Samples	514
20.9	Methodological Notes	515
20.10	Mathematical Notes	516
	20.10.1 Covariance Stationarity	517
	20.10.2 Dynamic Regression Log-Likelihood	518
	20.10.3 Hatanaka’s Estimator	518
	20.10.4 Two-Step Estimation	519
	20.10.5 Optimal Instruments	520
20.11	Overview	521
20.12	Exercises	522
	20.12.1 Review	522
	20.12.2 Extensions	526
Chapter 21	The Generalized Method of Moments	531
21.1	A Random Walk	532
21.2	Definition of GMM	536
	21.2.1 Turning Moments into Estimators	537
	21.2.2 Nonlinear Least Squares	539
	21.2.3 Two-Stage Least Squares	541
21.3	Identification	542
21.4	Distribution Theory	545
	21.4.1 Proof of Consistency	546

21.4.2	Proof of Asymptotic Normality	547
21.4.3	Variance Matrix Estimation	548
21.4.4	Efficiency	550
21.5	Methodological Notes	555
21.6	Mathematical Notes	555
21.6.1	Uniform Convergence of the GMM Criterion Function	555
21.6.2	Nonsingularity of the GMM Hessian	556
21.6.3	GMM Efficiency	557
21.7	Overview	557
21.8	Exercises	558
21.8.1	Review	558
21.8.2	Extensions	561
21.9	Appendix: Data Collection	562
Chapter 22	Generalized Method of Moments Hypothesis Tests	564
22.1	Tests of Parameter Restrictions	565
22.1.1	Wald Test	566
22.1.2	Gradient Test	567
22.1.3	Distance Difference Test	567
22.1.4	Minimum Chi-Square Test	568
22.1.5	Special Identities	569
22.1.6	Generalizing Likelihood-Based Diagnostics	570
22.2	Tests of Moment Restrictions	572
22.2.1	Overidentifying Restrictions Tests	576
22.3	Hausman Specification Tests	578
22.4	Equivalence among Test Statistics	585
22.4.1	A Trinity of GMM Test Statistics	586
22.4.2	Minimum Chi-Square	587
22.5	Statistical Power	590
22.6	Sequential Testing	592
22.7	Minimum Distance Estimation	594
22.8	Mathematical Notes	597
22.9	Methodological Notes	601
22.10	Overview	601
22.11	Exercises	602
22.11.1	Review	602
22.11.2	Extensions	605
Chapter 23	Overview	608
PART IV	LATENT VARIABLE MODELS	613

Chapter 24 Panel Data Models 615

24.1	Introduction	615
------	--------------	-----

24.2	Fixed Individual Effects	616
24.3	Random Individual Effects	618
24.4	Fixed versus Random Effects	622
24.5	Generalizations	623
	24.5.1 Individual-Specific Explanatory Variables	624
	24.5.2 Time-Specific Effects	625
	24.5.3 Dynamic Models	626
24.6	Specification Tests	628
24.7	Linear Projection	630
	24.7.1 Identification and OLS	631
	24.7.2 Efficient Estimation	632
	24.7.3 Diagnostic Tests	634
24.8	Additional Moment Restrictions	635
	24.8.1 Identification	635
	24.8.2 Estimation	636
24.9	Mathematical Notes	638
24.10	Overview	640
24.11	Exercises	641
	24.11.1 Review	641
	24.11.2 Extensions	643
Chapter 25 Autoregressive Moving-Average Time Series Models 645		
25.1	Introduction	645
25.2	Autoregressive Processes	649
	25.2.1 Stationarity	651
	25.2.2 Restricted Estimation	656
	25.2.3 Sequential Testing for Order	657
25.3	Moving-Average Processes	658
	25.3.1 Identification	660
	25.3.2 Kalman Filter	663
	25.3.3 Estimation	667
	25.3.4 Testing Serial Correlation	670
25.4	ARMA Processes	673
	25.4.1 Identification and Invertibility	675
	25.4.2 Kalman Filter and Estimation	679
	25.4.3 Hypothesis Tests	679
25.5	Wold Decomposition	680
	25.5.1 Linearly Deterministic Processes	681
	25.5.2 Wold Decomposition Theorem	682
25.6	Methodological Notes	684
25.7	Mathematical Notes	685
	25.7.1 Yule-Walker Equations	685
	25.7.2 Kalman Filter	686
	25.7.3 MA(q) Identification	689

25.7.4	Score Test Equivalence	690
25.8	Overview	691
25.9	Exercises	693
25.9.1	Review	693
25.9.2	Extensions	696
Chapter 26	Simultaneous Equations	697
26.1	Introduction	697
26.2	Seemingly Unrelated Regressions	698
26.2.1	Estimation of a Cost Function	699
26.2.2	Assumptions	700
26.2.3	OLS versus GLS	701
26.2.4	Feasible GLS Estimation	704
26.2.5	Maximum Likelihood Estimation	705
26.3	Simultaneous Equations	706
26.3.1	Definitions	709
26.3.2	Assumptions	710
26.4	Identification	711
26.4.1	Equation Identification	714
26.4.2	System Identification	717
26.5	Estimation	718
26.5.1	Limited Information	719
26.5.2	Full Information	721
26.5.3	Maximum Likelihood	723
26.6	Hypothesis Tests	727
26.7	Mathematical Notes	730
26.7.1	Score Functions	730
26.7.2	Information Matrix	732
26.8	Overview	734
26.9	Exercises	736
26.9.1	Review	736
26.9.2	Extensions	742
Chapter 27	Discrete Dependent Variables	747
27.1	Bernoulli Dependent Variables	748
27.1.1	Bernoulli Regression	748
27.1.2	Estimation	750
27.1.3	A Latent Variable Interpretation	755
27.2	Additional Univariate Models	757
27.2.1	Ordered Data	758
27.2.2	Count Data	761
27.3	Multivariate Models	764
27.3.1	Multiple Choice	764
27.3.2	Rank-Ordered Multiple Choice	770

27.4	Latent Variables and Computation	771
27.4.1	Score Functions	772
27.4.2	Hessian and Information Functions	773
27.4.3	EM Algorithm	774
27.4.4	Simulation	775
27.5	Methodological Notes	777
27.6	Mathematical Notes	778
27.6.1	Katz Family of Distributions	779
27.6.2	Logit Probabilities	780
27.6.3	EM Algorithm	782
27.6.4	Simulation	784
27.7	Overview	784
27.8	Exercises	786
27.8.1	Review	786
27.8.2	Extensions	788
Chapter 28	Censored and Truncated Variables	791
28.1	Labor Supply	791
28.2	Mixed Probability Functions	794
28.3	Censored Moments	798
28.4	Estimation	800
28.5	Prediction and Truncated Means	802
28.6	Truncated Regression	803
28.7	Nonrandom Sample Selection	806
28.7.1	Log-Likelihood	807
28.7.2	Moments	808
28.7.3	Estimation	809
28.8	Specification of Distribution	810
28.8.1	Censored Regression	810
28.8.2	Truncated Regression	814
28.9	Mathematical Notes	817
28.9.1	Integrals	817
28.9.2	Censored Moments	819
28.9.3	Nonrandom Sample Selection	822
28.10	Overview	824
28.11	Exercises	824
28.11.1	Review	824
28.11.2	Extensions	827
Chapter 29	Overview	831
PART V	APPENDICES	833
Appendix A	Abbreviations and Acronyms	835
Appendix B	Notation	837

B.1	Limits	837
B.2	Sets	838
B.3	Functions	838
B.4	Linear Vector Spaces	838
B.5	Matrices	839
B.6	Random Variables	840
B.7	Optima and Roots	840
Appendix C	Linear Algebra and Matrix Theory	841
C.1	Linear Vector Spaces	841
C.2	Linear Transformations	847
C.3	Inner Products and Orthogonality	851
C.4	Normed Linear Vector Spaces	855
C.5	Determinants	856
	C.5.1 Volume of a Parallelogram	857
	C.5.2 Determinant of a Matrix	860
	C.5.3 The Cofactor Expansion	862
C.6	Eigenvalues and Eigenvectors	865
Appendix D	Probability	867
D.1	Fundamental Concepts	867
D.2	Random Variables	868
	D.2.1 Mathematical Notes	875
D.3	Joint and Conditional Probability	879
	D.3.1 Mathematical Notes	883
D.4	Special Distributions	884
D.5	Limiting Approximations	891
	D.5.1 A Sequence of Densities	893
	D.5.2 Sequences of Moments	895
	D.5.3 Sequences of c.f.s	896
	D.5.4 Mathematical Notes	898
Appendix E	Classical Statistics	902
E.1	Sampling	902
E.2	Classical Statistical Inference	904
	E.2.1 Estimation	904
	E.2.2 Hypothesis Tests	906
	E.2.3 Estimation Methods	907
	E.2.4 Asymptotic Distribution Theory	913
Appendix F	Noncentral Distributions	916
Appendix G	Multivariate Differentiation	922
G.1	Basic Notation	922
G.2	Vectorization and Kronecker Products	924

	G.2.1 Kronecker Products	925
	G.3 Derivative Vectors	926
	G.4 Derivative Matrices	927
	G.5 The Normal Log-Likelihood Function	928
Appendix H	Characteristic Functions	931
	Bibliography	935
	Index	945