

A First Course in Linear Model Theory

Nalini Ravishanker Dipak K. Dey

June 28, 2001

Contents

1	A Review of Vector and Matrix Algebra	7
1.1	Notation	7
1.2	Basic definitions and properties	9
	Exercises	35
2	Properties of Special Matrices	39
2.1	Partitioned matrices	39
2.2	Algorithms for matrix factorization	46
2.3	Symmetric and idempotent matrices	50
2.4	Nonnegative definite quadratic forms and matrices	56
2.5	Simultaneous diagonalization of matrices	62
2.6	Geometrical perspectives	64
2.7	Vector and matrix differentiation	69
2.8	Special operations on matrices	73
2.9	Linear optimization	75
	Exercises	76
3	Generalized Inverse and Solutions to Linear Systems	81
3.1	Generalized inverses	81
3.2	Solutions to linear systems	90
	Exercises	96
4	The General Linear Model	
	(in A First Course in Linear Model Theory	
	by Nalini Ravishanker and Dipak K. Dey	
	University of Connecticut)	99
4.1	Model definition and examples	99
4.2	The least squares approach	104
4.3	Estimable functions	121
4.4	Gauss-Markov theorem	126
4.5	Generalized least squares	129
4.6	Estimation subject to linear restrictions	134
	4.6.1 Method of Lagrangian multipliers	134
	4.6.2 Method of orthogonal projections	136

Exercises	138
5 Multivariate Normal and Related Distributions (in A First Course in Linear Model Theory by Nalini Ravishanker and Dipak K. Dey University of Connecticut)	141
5.1 Multivariate probability distributions	142
5.2 Multivariate normal distribution and properties	149
5.3 Some noncentral distributions	169
5.4 Distributions of quadratic forms	177
5.5 Alternatives to the multivariate normal distribution	184
5.5.1 Mixture of normal distributions	184
5.5.2 Spherical distributions	186
5.5.3 Elliptical distributions	188
Exercises	193
6 Sampling from the Multivariate Normal Distribution	199
6.1 Distribution of the sample mean and covariance matrix	199
6.2 Distributions related to correlation coefficients	204
6.3 Assessing the normality assumption	208
6.4 Transformations to approximate normality	212
6.4.1 Univariate transformations	213
6.4.2 Multivariate transformations	215
Exercises	216
7 Inference for the General Linear Model	217
7.1 Properties of least squares estimates	217
7.2 General linear hypotheses	221
7.2.1 Derivation of and motivation for the F -test	221
7.2.2 Power of the F -test	233
7.2.3 Testing independent and orthogonal contrasts	234
7.3 Confidence intervals and multiple comparisons	234
7.3.1 Joint and marginal confidence intervals	235
7.3.2 Simultaneous confidence intervals	237
7.3.3 Multiple comparison procedures	241
7.4 Restricted and reduced models	246
7.4.1 Lack of fit test	263
7.4.2 Non-testable hypotheses	265
7.5 Maximum likelihood estimation	265
7.5.1 Normal linear models	266
7.5.2 Elliptically contoured linear model	269
7.5.3 Model selection criteria	269
Exercises	269

8	Multiple Regression Models	275
8.1	Departures from model assumptions	275
8.1.1	Graphical procedures	276
8.1.2	Sequential and partial F -tests	279
8.1.3	Tests for heteroscedasticity	281
8.1.4	Tests for serial correlation	285
8.1.5	Measurement errors	288
8.2	Regression model selection	288
8.3	Orthogonal and collinear predictors	294
8.3.1	Orthogonality in regression	294
8.3.2	Multicollinearity	297
8.3.3	Ridge regression	300
8.3.4	Principal components regression	302
8.4	Prediction intervals and calibration	303
8.5	Regression diagnostics	307
8.5.1	Further properties of the projection matrix	308
8.5.2	Types of residuals	309
8.5.3	Outliers and high leverage observations	313
8.5.4	Diagnostic measures based on influence functions	314
8.6	Dummy variables in regression	322
8.7	Robust regression	324
8.7.1	Least absolute deviations (LAD) regression	325
8.7.2	M -regression	328
8.7.3	L -regression	329
8.8	Nonparametric regression methods	329
8.8.1	Neural networks regression	329
8.8.2	Projection pursuit regression	330
8.9	Regression based on wavelet methods	330
8.10	Additive models	330
	Exercises	332
9	Fixed Effects Linear Models	335
9.1	Checking model assumptions	335
9.2	Inference for unbalanced models	337
9.2.1	One-way cell means model	338
9.2.2	Two-factor models	339
9.2.3	Cross-classified models	340
9.3	High order ANOVA models	340
9.4	Analysis of covariance	340
9.5	Nonparametric procedures	344
9.5.1	Kruskal-Wallis procedure	344
9.5.2	Friedman's test procedure	346
	Exercises	346

10 Random-Effects and Mixed-Effects Models	347
10.1 One-way random-effects model	347
10.1.1 ANOVA method	349
10.1.2 Maximum likelihood estimation	354
10.1.3 Restricted maximum likelihood (REML) estimation	355
10.2 Higher-order random-effects models	356
10.2.1 Two-way nested classification	356
10.2.2 Two-way crossed classification	357
10.3 Mixed-effects linear models	359
10.3.1 Estimation procedures	361
10.3.2 Two-factor models	361
10.3.3 Higher-order models	362
Exercises	362
11 Special Topics	363
11.1 Bayesian linear models	363
11.1.1 Posterior distribution	367
11.1.2 Predictive distribution	367
11.2 Dynamic linear model and the Kalman filter	367
11.2.1 Smoothing equations	369
11.2.2 Parameter estimation	369
11.3 Longitudinal models	369
11.3.1 Estimation approaches	370
11.3.2 Uniform and exponential correlation models	371
11.4 Generalized linear models	371
11.4.1 Components of GLIM	371
11.4.2 Estimation and model fitting	371
Exercises	371
A Review of some univariate distributions	373
A Solutions to Selected Exercises	379