Mixed Models

Theory and Applications with R

Second Edition

EUGENE DEMIDENKO
Dartmouth College

WILEY

Contents

Preface xvii

Preface to the Second Edition xix

R Software and Functions xx

Data Sets xxii

Open Problems in Mixed Models xxiii

1 Introduction: Why Mixed Models? 1

1.1 Mixed effects for clustered data 2
1.2 ANOVA, variance components, and the mixed model 4
1.3 Other special cases of the mixed effects model 6
1.4 Compromise between Bayesian and frequentist approaches 7
1.5 Penalized likelihood and mixed effects 9
1.6 Healthy Akaike information criterion 11
1.7 Penalized smoothing 13
1.8 Penalized polynomial fitting 16
1.9 Restraining parameters, or what to eat 18
1.10 Ill-posed problems, Tikhonov regularization, and mixed effects . . . 20
1.11 Computerized tomography and linear image reconstruction 23
1.12 GLMM for PET 26
1.13 Maple leaf shape analysis 29
1.14 DNA Western blot analysis 31
1.15 Where does the wind blow? 33
1.16 Software and books 36

1.17 Summary points 37

2 MLE for the LME Model **41**
2.1 Example: weight versus height 42
2.1.1 The first R script 43
2.2 The model and log-likelihood functions 45
2.2.1 The model 45
2.2.2 Log-likelihood functions 48
2.2.3 Dimension-reduction formulas 49
2.2.4 Profile log-likelihood functions 53
2.2.5 Dimension-reduction GLS estimate 55
2.2.6 Restricted maximum likelihood 56
2.2.7 Weight versus height (continued) 59
2.3 Balanced random-coefficient model 60
2.4 LME model with random intercepts 64
2.4.1 Balanced random-intercept model 67
2.4.2 How random effect affects the variance of MLE 71
2.5 Criterion for MLE existence 72
2.6 Criterion for the positive definiteness of matrix **D** 74
2.6.1 Example of an invalid LME model 75
2.7 Pre-estimation bounds for variance parameters 77
2.8 Maximization algorithms 79
2.9 Derivatives of the log-likelihood function 81
2.10 Newton–Raphson algorithm 82
2.11 Fisher scoring algorithm 85
2.11.1 Simplified FS algorithm 86
2.11.2 Empirical FS algorithm 86
2.11.3 Variance-profile FS algorithm 87
2.12 EM algorithm 88
2.12.1 Fixed-point algorithm 92
2.13 Starting point 93
2.13.1 FS starting point 93
2.13.2 FP starting point 94
2.14 Algorithms for restricted MLE 95
2.14.1 Fisher scoring algorithm 95
2.14.2 EM algorithm 96
2.15 Optimization on nonnegative definite matrices 96
2.15.1 How often can one hit the boundary? 97
2.15.2 Allow matrix **D** to be not nonnegative definite 98
2.15.3 Force matrix **D** to stay nonnegative definite 103
2.15.4 Matrix **D** reparameterization 104
2.15.5 Criteria for convergence 105
2.16 lmeFS and lme in R 107
2.17 Appendix: proof of the existence of MLE 111
2.18 Summary points 114

3 Statistical Properties of the LME Model **117**
3.1 Introduction 117

3.2 Identifiability of the LME model 117
3.2.1 Linear regression with random coefficients 119
3.3 Information matrix for variance parameters 120
3.3.1 Efficiency of variance parameters for balanced data 129
3.4 Profile-likelihood confidence intervals 131
3.5 Statistical testing of the presence of random effects 133
3.6 Statistical properties of MLE 137
3.6.1 Small-sample properties 137
3.6.2 Large-sample properties 140
3.6.3 ML and RML are asymptotically equivalent 144
3.7 Estimation of random effects 145
3.7.1 Implementation in R 148
3.8 Hypothesis and membership testing 151
3.8.1 Membership test 152
3.9 Ignoring random effects 154
3.10 MINQUE for variance parameters 157
3.10.1 Example: linear regression 158
3.10.2 MINQUE for σ^2 160
3.10.3 MINQUE for $\mathbf{D}_*$ 162
3.10.4 Linear model with random intercepts 165
3.10.5 MINQUE for the balanced model 165
3.10.6 lmevarMINQUE function 166
3.11 Method of moments 166
3.11.1 lmevarMM function 171
3.12 Variance least squares estimator 171
3.12.1 Unbiased VLS estimator 173
3.12.2 Linear model with random intercepts 174
3.12.3 Balanced design 174
3.12.4 VLS as the first iteration of ML 175
3.12.5 lmevarUVLS function 175
3.13 Projection on $\mathbb{D}_+$ space 176
3.14 Comparison of the variance parameter estimation 176
3.14.1 lmesim function 179
3.15 Asymptotically efficient estimation for $\boldsymbol{\beta}$ 180
3.16 Summary points 181

4 Growth Curve Model and Generalizations 185
4.1 Linear growth curve model 185
4.1.1 Known matrix $\mathbf{D}$ 187
4.1.2 Maximum likelihood estimation 189
4.1.3 Method of moments for variance parameters 192
4.1.4 Two-stage estimation 196
4.1.5 Special growth curve models 196
4.1.6 Unbiasedness and efficient estimation for $\boldsymbol{\beta}$ 200
4.2 General linear growth curve model 201
4.2.1 Example: Calcium supplementation for bone gain 202
4.2.2 Variance parameters are known 204
4.2.3 Balanced model 207

4.2.4 Likelihood-based estimation . . . 208
4.2.5 MM estimator for variance parameters . . . 213
4.2.6 Two-stage estimator and asymptotic properties . . . 214
4.2.7 Analysis of misspecification . . . 215
4.3 Linear model with linear covariance structure . . . 219
4.3.1 Method of maximum likelihood . . . 220
4.3.2 Variance least squares . . . 222
4.3.3 Statistical properties . . . 223
4.3.4 LME model for longitudinal autocorrelated data . . . 224
4.3.5 Multidimensional LME model . . . 229
4.4 Robust linear mixed effects model . . . 233
4.4.1 Robust estimation of the location parameter with estimated σ and c . . . 235
4.4.2 Robust linear regression with estimated threshold . . . 238
4.4.3 Robust LME model . . . 239
4.4.4 Alternative robust functions . . . 239
4.4.5 Robust random effect model . . . 240
4.5 Appendix: derivation of the MM estimator . . . 241
4.6 Summary points . . . 242

5 Meta-analysis Model **245**
5.1 Simple meta-analysis model . . . 246
5.1.1 Estimation of random effects . . . 248
5.1.2 Maximum likelihood estimation . . . 248
5.1.3 Quadratic unbiased estimation for σ^2 . . . 253
5.1.4 Statistical inference . . . 260
5.1.5 Robust/median meta-analysis . . . 266
5.1.6 Random effect coefficient of determination . . . 271
5.2 Meta-analysis model with covariates . . . 273
5.2.1 Maximum likelihood estimation . . . 274
5.2.2 Quadratic unbiased estimation for σ^2 . . . 277
5.2.3 Hypothesis testing . . . 278
5.3 Multivariate meta-analysis model . . . 278
5.3.1 The model . . . 280
5.3.2 Maximum likelihood estimation . . . 283
5.3.3 Quadratic estimation of the heterogeneity matrix . . . 285
5.3.4 Test for homogeneity . . . 288
5.4 Summary points . . . 289

6 Nonlinear Marginal Model **291**
6.1 Fixed matrix of random effects . . . 292
6.1.1 Log-likelihood function . . . 293
6.1.2 nls function in R . . . 295
6.1.3 Computational issues of nonlinear least squares . . . 296
6.1.4 Distribution-free estimation . . . 297
6.1.5 Testing for the presence of random effects . . . 298
6.1.6 Asymptotic properties . . . 298
6.1.7 Example: log-Gompertz growth curve . . . 299

6.2 Varied matrix of random effects 305
6.2.1 Maximum likelihood estimation 305
6.2.2 Distribution-free variance parameter estimation 308
6.2.3 GEE and iteratively reweighted least squares 309
6.2.4 Example: logistic curve with random asymptote 310
6.3 Three types of nonlinear marginal models 316
6.3.1 Type I nonlinear marginal model 317
6.3.2 Type II nonlinear marginal model 319
6.3.3 Type III nonlinear marginal model 319
6.3.4 Asymptotic properties under distribution misspecification . . 320
6.4 Total generalized estimating equations approach 321
6.4.1 Robust feature of total GEE 323
6.4.2 Expected Newton–Raphson algorithm for total GEE 323
6.4.3 Total GEE for the mixed effects model 324
6.4.4 Total GEE for the LME model 324
6.4.5 Example (continued): log-Gompertz curve 325
6.4.6 Photodynamic tumor therapy 326
6.5 Summary points 328

7 Generalized Linear Mixed Models **331**
7.1 Regression models for binary data 332
7.1.1 Approximate relationship between logit and probit 336
7.1.2 Computation of the logistic-normal integral 338
7.1.3 Gauss-Hermite numerical quadrature for multidimensional integrals in R 350
7.1.4 Log-likelihood and its numerical properties 352
7.1.5 Unit step algorithm 353
7.2 Binary model with subject-specific intercept 355
7.2.1 Consequences of ignoring a random effect 357
7.2.2 ML logistic regression with a fixed subject-specific intercept 358
7.2.3 Conditional logistic regression 359
7.3 Logistic regression with random intercept 362
7.3.1 Maximum likelihood 362
7.3.2 Fixed sample likelihood approximation 368
7.3.3 Quadratic approximation 371
7.3.4 Laplace approximation to the likelihood 371
7.3.5 VARLINK estimation 374
7.3.6 Beta-binomial model 376
7.3.7 Statistical test of homogeneity 378
7.3.8 Asymptotic properties 381
7.4 Probit model with random intercept 382
7.4.1 Laplace and PQL approximations 382
7.4.2 VARLINK estimation 383
7.4.3 Heckman method for the probit model 383
7.4.4 Generalized estimating equations approach 384
7.4.5 Implementation in R 386
7.5 Poisson model with random intercept 386
7.5.1 Poisson regression for count data 387

7.5.2 Clustered count data 388
7.5.3 Fixed intercepts 389
7.5.4 Conditional Poisson regression 390
7.5.5 Negative binomial regression 391
7.5.6 Normally distributed intercepts 394
7.5.7 Exact GEE for any distribution 396
7.5.8 Exact GEE for balanced count data 397
7.5.9 Heckman method for the Poisson model 398
7.5.10 Tests for overdispersion 399
7.5.11 Implementation in R 400
7.6 Random intercept model: overview 401
7.7 Mixed models with multiple random effects 402
7.7.1 Multivariate Laplace approximation 403
7.7.2 Logistic regression 403
7.7.3 Probit regression 407
7.7.4 Poisson regression 408
7.7.5 Homogeneity tests 410
7.8 GLMM and simulation methods 412
7.8.1 General form of GLMM via the exponential family 412
7.8.2 Monte Carlo for ML 413
7.8.3 Fixed sample likelihood approach 413
7.9 GEE for clustered marginal GLM 416
7.9.1 Variance least squares 418
7.9.2 Limitations of the GEE approach 420
7.9.3 Marginal or conditional model? 422
7.9.4 Implementation in R 423
7.10 Criteria for MLE existence for a binary model 424
7.11 Summary points 429

8 Nonlinear Mixed Effects Model **433**
8.1 Introduction 433
8.2 The model 434
8.3 Example: height of girls and boys 437
8.4 Maximum likelihood estimation 439
8.5 Two-stage estimator 442
8.5.1 Maximum likelihood estimation 445
8.5.2 Method of moments 445
8.5.3 Disadvantage of two-stage estimation 446
8.5.4 Further discussion 446
8.5.5 Two-stage method in the presence of a common parameter . 447
8.6 First-order approximation 448
8.6.1 GEE and MLE 448
8.6.2 Method of moments and VLS 449
8.7 Lindstrom–Bates estimator 450
8.7.1 What if matrix **D** is not positive definite? 452
8.7.2 Relation to the two-stage estimator 452
8.7.3 Computational aspects of penalized least squares 453
8.7.4 Implementation in R: the function nlme 454

8.8 Likelihood approximations . . . 456
8.8.1 Linear approximation of the likelihood at zero . . . 456
8.8.2 Laplace and PQL approximations . . . 457
8.9 One-parameter exponential model . . . 459
8.9.1 Maximum likelihood estimator . . . 459
8.9.2 First-order approximation . . . 460
8.9.3 Two-stage estimator . . . 461
8.9.4 Lindstrom–Bates estimator . . . 463
8.10 Asymptotic equivalence of the TS and LB estimators . . . 466
8.11 Bias-corrected two-stage estimator . . . 468
8.12 Distribution misspecification . . . 470
8.13 Partially nonlinear marginal mixed model . . . 473
8.14 Fixed sample likelihood approach . . . 474
8.14.1 Example: one-parameter exponential model . . . 475
8.15 Estimation of random effects and hypothesis testing . . . 476
8.15.1 Estimation of the random effects . . . 476
8.15.2 Hypothesis testing for the NLME model . . . 477
8.16 Example (continued) . . . 478
8.17 Practical recommendations . . . 480
8.18 Appendix: Proof of theorem on equivalence . . . 481
8.19 Summary points . . . 484

9 Diagnostics and Influence Analysis 487
9.1 Introduction . . . 487
9.2 Influence analysis for linear regression . . . 488
9.3 The idea of infinitesimal influence . . . 491
9.3.1 Data influence . . . 491
9.3.2 Model influence . . . 492
9.4 Linear regression model . . . 493
9.4.1 Influence of the dependent variable . . . 494
9.4.2 Influence of the continuous explanatory variable . . . 495
9.4.3 Influence of the binary explanatory variable . . . 497
9.4.4 Influence on the predicted value . . . 497
9.4.5 Case or group deletion . . . 498
9.4.6 R code . . . 500
9.4.7 Influence on regression characteristics . . . 501
9.4.8 Example 1: Women's body fat . . . 503
9.4.9 Example 2: gypsy moth study . . . 507
9.5 Nonlinear regression model . . . 510
9.5.1 Influence of the dependent variable on the LSE . . . 510
9.5.2 Influence of the explanatory variable on the LSE . . . 510
9.5.3 Influence on the predicted value . . . 511
9.5.4 Influence of case deletion . . . 511
9.5.5 Example 3: logistic growth curve model . . . 512
9.6 Logistic regression for binary outcome . . . 515
9.6.1 Influence of the covariate on the MLE . . . 516
9.6.2 Influence on the predicted probability . . . 516
9.6.3 Influence of the case deletion on the MLE . . . 517

9.6.4 Sensitivity to misclassification 517
9.6.5 Example: Finney data 522
9.7 Influence of correlation structure 524
9.8 Influence of measurement error 525
9.9 Influence analysis for the LME model 528
9.9.1 Example: Weight versus height 532
9.10 Appendix: MLE derivative with respect to σ^2 534
9.11 Summary points 535

10 Tumor Regrowth Curves **539**
10.1 Survival curves 541
10.2 Double-exponential regrowth curve 543
10.2.1 Time to regrowth, T_R 546
10.2.2 Time to reach specific tumor volume, T_* 547
10.2.3 Doubling time, T_D 547
10.2.4 Statistical model for regrowth 548
10.2.5 Variance estimation for tumor regrowth outcomes 549
10.2.6 Starting values 550
10.2.7 Example: chemotherapy treatment comparison 551
10.3 Exponential growth with fixed regrowth time 557
10.3.1 Statistical hypothesis testing 558
10.3.2 Synergistic or supra-additive effect 558
10.3.3 Example: combination of treatments 559
10.4 General regrowth curve 563
10.5 Double-exponential transient regrowth curve 564
10.5.1 Example: treatment of cellular spheroids 570
10.6 Gompertz transient regrowth curve 571
10.6.1 Example: tumor treated in mice 572
10.7 Summary points 574

11 Statistical Analysis of Shape **577**
11.1 Introduction 577
11.2 Statistical analysis of random triangles 579
11.3 Face recognition 582
11.4 Scale-irrelevant shape model 583
11.4.1 Random effects scale-irrelevant shape model 585
11.4.2 Scale-irrelevant shape model on the log scale 586
11.4.3 Fixed or random size? 587
11.5 Gorilla vertebrae analysis 587
11.6 Procrustes estimation of the mean shape 589
11.6.1 Polygon estimation 592
11.6.2 Generalized Procrustes model 592
11.6.3 Random effects shape model 593
11.6.4 Random or fixed (Procrustes) effects model? 594
11.6.5 Maple leaf analysis 594
11.7 Fourier descriptor analysis 596
11.7.1 Analysis of a star shape 596
11.7.2 Random Fourier descriptor analysis 602

11.7.3 Potato project 604
11.8 Summary points 605

12 Statistical Image Analysis **607**
12.1 Introduction 607
12.1.1 What is a digital image? 608
12.1.2 Image arithmetic 609
12.1.3 Ensemble and repeated measurements 609
12.1.4 Image and spatial statistics 610
12.1.5 Structured and unstructured images 610
12.2 Testing for uniform lighting 610
12.2.1 Estimating light direction and position 612
12.3 Kolmogorov–Smirnov image comparison 614
12.3.1 Kolmogorov–Smirnov test for image comparison 614
12.3.2 Example: histological analysis of cancer treatment 615
12.4 Multinomial statistical model for images 618
12.4.1 Multinomial image comparison 620
12.5 Image entropy 621
12.5.1 Reduction of a gray image to binary 623
12.5.2 Entropy of a gray image and histogram equalization 623
12.6 Ensemble of unstructured images 625
12.6.1 Fixed-shift model 626
12.6.2 Random-shift model 628
12.6.3 Mixed model for gray images 631
12.6.4 Two-stage estimation 633
12.6.5 Schizophrenia MRI analysis 635
12.7 Image alignment and registration 638
12.7.1 Affine image registration 641
12.7.2 Weighted sum of squares 642
12.7.3 Nonlinear transformations 643
12.7.4 Random registration 643
12.7.5 Linear image interpolation 644
12.7.6 Computational aspects 645
12.7.7 Derivative-free algorithm for image registration 646
12.7.8 Example: clock alignment 647
12.8 Ensemble of structured images 650
12.8.1 Fixed affine transformations 650
12.8.2 Random affine transformations 651
12.9 Modeling spatial correlation 652
12.9.1 Toeplitz correlation structure 654
12.9.2 Simultaneous estimation of variance and transform parameters 656
12.10 Summary points 658

13 Appendix: Useful Facts and Formulas **661**
13.1 Basic facts of asymptotic theory . . . 661
13.1.1 Central Limit Theorem . . . 661
13.1.2 Generalized Slutsky theorem . . . 662
13.1.3 Pseudo-maximum likelihood . . . 664
13.1.4 Estimating equations approach and the sandwich formula . . 665
13.1.5 Generalized estimating equations approach . . . 667
13.2 Some formulas of matrix algebra . . . 668
13.2.1 Some matrix identities . . . 668
13.2.2 Formulas for generalized matrix inverse . . . 668
13.2.3 *Vec* and vech functions; duplication matrix . . . 669
13.2.4 Matrix differentiation . . . 670
13.3 Basic facts of optimization theory . . . 672
13.3.1 Criteria for unimodality . . . 673
13.3.2 Criteria for global optimum . . . 674
13.3.3 Criteria for minimum existence . . . 674
13.3.4 Optimization algorithms in statistics . . . 675
13.3.5 Necessary condition for optimization and criteria for convergence . . . 678

References **681**

Index **711**