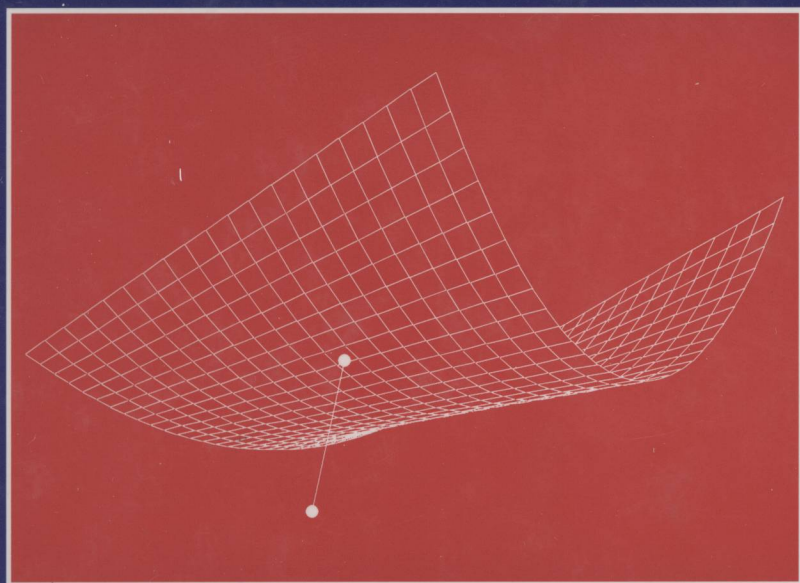


Mixed Models

Theory and Applications



Eugene Demidenko

WWW.
LINK AVAILABLE

0212.1
D379

Mixed Models

Theory and Applications

EUGENE DEMIDENKO
Dartmouth College



E200601287

 **WILEY-
INTERSCIENCE**

A JOHN WILEY & SONS, INC., PUBLICATION

Copyright © 2004 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department within the U.S. at 877-762-2974, outside the U.S. at 317-572-3993 or fax 317-572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic format.

Library of Congress Cataloging-in-Publication Data Is Available

ISBN 0-471-60161-6

Printed in the United States of America.

10 9 8 7 6 5 4 3 2

Mixed Models

WILEY SERIES IN PROBABILITY AND STATISTICS

Established by WALTER A. SHEWHART and SAMUEL S. WILKS

Editors: *David J. Balding, Noel A. C. Cressie, Nicholas I. Fisher,
Iain M. Johnstone, J. B. Kadane, Geert Molenberghs, Louise M. Ryan,
David W. Scott, Adrian F. M. Smith, Jozef L. Teugels*
Editors Emeriti: *Vic Barnett, J. Stuart Hunter, David G. Kendall*

A complete list of the titles in this series appears at the end of this volume.

To my family

Preface

Technological advances change the world, and statistics is no exception. The cornerstone of classical statistics is the notion of *sample*. Today, data are richer: We may have repeated measurements with thousands of clusters; data may come in the form of shapes or images. This book is about statistical analysis of data that constitute a *sample of samples*. In the first ten chapters we discuss statistical models when data come in traditional form as a sequence of numbers. Chapter 11 deals with a sample (ensemble) of shapes and in Chapter 12 we discuss how to analyze an ensemble of images.

We take the statistical model-based approach to analyze data. Then the method of analysis is a derivative. Although the method sometimes comes first, the model-based approach has obvious advantages: assumptions are clearly formulated, and properties of several methods can be studied and compared. For example, least squares is a method of fitting, but its pros and cons can be fully understood only when a statistical model is put forward to describe how observations are obtained. Then least squares is deduced, for example, from maximum likelihood.

Statistical treatment is carried out under a unifying mixed effects approach. This approach becomes fruitful not only to analyze complex clustered data (a sample of samples) but also as a statistical model for penalization and a common ground for the Bayesian and frequentist camps.

Use of the mixed modeling technique in shape and image analysis is exciting and promising. Much work remains to reveal the full power of this statistical approach to these nontraditional statistical data.

The book is divided into three parts. The first eight chapters cover the theory of mixed models: the linear mixed effects (LME) model, the generalized linear mixed model (GLMM), and the nonlinear mixed effects (NLME) model. In Chapter 9 we discuss methods of model diagnostics and influential analysis. The last three chapters are devoted to applications: tumor regrowth, shape, and image. Major results and points of discussion in each chapter are written in lay language and are collected in *Summary points* sections so that the reader can get a quick chapter overview.

I look forward to hearing from readers and invite them to visit the book web site at

<http://www.dartmouth.edu/~eugened>

where some additional information with data and images is presented.

I would like to thank the many people I worked with on various projects that have led up to this book. First, I would like to mention my long-term collaboration with Therese Stukel and Tor Tosteson and thank them for their support. I am grateful to Harold Swartz and Jack Hoopes for the exposure to biological problems, and to the team led by Keith Paulsen, including Alex Hartov, Paul Meaney, and Brian Pogue, all from Dartmouth, who introduced me to the world of image reconstruction. Many thanks to John Baron, Margaret Karagas, and Mark Israel for creating a friendly scientific atmosphere. I am grateful to Ed Vonesh for discussion and his helpful comments.

Finally, thanks to the *Scientific Workplace*, a WYSIWYG version of the $\text{\LaTeX}$ typesetting system (<http://www.mackichan.com>)—it is hard to imagine writing this book without this software.

Eugene Demidenko

Hanover, New Hampshire
Dartmouth College
January 2004

Contents

Preface

xvii

1	Introduction: Why Mixed Models?	1
1.1	Mixed effects for clustered data	2
1.2	ANOVA, variance components, and the mixed model	5
1.3	Other special cases of the mixed effects model	7
1.4	A compromise between Bayesian and frequentist approaches	8
1.5	Penalized likelihood and mixed effects	11
1.6	Healthy Akaike information criterion	13
1.7	Penalized smoothing	15
1.8	Penalized polynomial fitting	18
1.9	Restraining parameters, or what to eat	19
1.10	Ill-posed problems, Tikhonov regularization, and mixed effects	22
1.11	Computerized tomography and linear image reconstruction	26
1.12	GLMM for PET	30
1.13	Maple shape leaf analysis	33
1.14	DNA Western blot analysis	35
1.15	Where does the wind blow?	38
1.16	Commercial software and books	40
1.17	Summary points	42
2	MLE for LME Model	45
2.1	Example: Height versus weight	46
2.2	The model and log-likelihood functions	48

2.2.1	The model	48
2.2.2	Log-likelihood functions	51
2.2.3	Dimension-reduction formulas	52
2.2.4	Profile log-likelihood functions	56
2.2.5	Restricted maximum likelihood	58
2.3	Balanced random-coefficient model	61
2.4	LME model with random intercepts	66
2.4.1	Balanced random-intercept model	69
2.4.2	How random effect affects the variance of MLE	72
2.5	Criterion for the MLE existence	74
2.6	Criterion for positive definiteness of matrix $\mathbf{D}$	76
2.6.1	Example of an invalid LME model	77
2.7	Preestimation bounds for variance parameters	78
2.8	Maximization algorithms	80
2.9	Derivatives of the log-likelihood function	82
2.10	Newton–Raphson algorithm	83
2.11	Fisher scoring algorithm	86
2.11.1	Simplified FS algorithm	87
2.11.2	Empirical FS algorithm	88
2.11.3	Variance-profile FS algorithm	88
2.11.4	Full-profile FS algorithm	89
2.12	EM algorithm	90
2.12.1	Fixed-point algorithm	94
2.13	Starting point	96
2.13.1	FS starting point	96
2.13.2	FP starting points	97
2.14	Algorithms for restricted MLE	98
2.14.1	Fisher scoring algorithm	98
2.14.2	EM algorithm	98
2.15	Optimization on nonnegative definite matrices	99
2.15.1	How often can one hit the boundary?	100
2.15.2	Allow matrix $\mathbf{D}$ to be negative definite	101
2.15.3	Force matrix $\mathbf{D}$ to stay nonnegative definite	106
2.15.4	Matrix $\mathbf{D}$ reparameterization	108
2.15.5	Criteria for convergence	108
2.16	Appendix: Proof of the MLE existence	111
2.17	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	121
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	134
3.6	Statistical properties of MLE	139
3.6.1	Small-sample properties	139
3.6.2	Large-sample properties	141
3.6.3	ML and RML are asymptotically equivalent	146
3.7	Estimation of random effects	147
3.8	Hypothesis and membership testing	150
3.8.1	Membership test	151
3.9	Ignoring random effects	154
3.10	MINQUE for variance parameters	157
3.10.1	Linear regression	157
3.10.2	MINQUE for σ^2	159
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	166
3.11	Method of moments	166
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	176
3.13	Projection on $\mathbb{D}_+$ space	176
3.14	Comparison of the variance parameter estimation	176
3.15	Asymptotically efficient estimation for β	178
3.16	Summary points	180
4	Growth Curve Model and Generalizations	183
4.1	Linear growth curve model	183
4.1.1	Known matrix $\mathbf{D}$	185
4.1.2	Maximum likelihood estimation	187
4.1.3	Method of moments for variance parameters	191
4.1.4	Two-stage estimation	196
4.1.5	Special growth curve models	196
4.1.6	Unbiasedness and efficient estimation for β	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	205
4.2.3	Balanced model	207
4.2.4	Likelihood-based estimation	208
4.2.5	MM estimator for variance parameters	214
4.2.6	Two-stage estimator and asymptotic properties	215
4.2.7	Analysis of misspecification	215
4.3	Linear model with linear covariance structure	220
4.3.1	Method of maximum likelihood	221
4.3.2	Variance least squares	223

4.3.3	Statistical properties	224
4.3.4	LME model for longitudinal autocorrelated data . .	225
4.3.5	Multidimensional LME model	230
4.4	Robust linear mixed effects model	235
4.4.1	Robust estimation of the location parameter with es- timated σ and c	237
4.4.2	Robust linear regression with estimated threshold . .	241
4.4.3	Robust LME model	241
4.4.4	Alternative robust functions	242
4.4.5	Robust random effect model	242
4.5	Appendix: Derivation of the MM estimator	243
4.6	Summary points	244
5	Meta-analysis Model	247
5.1	Simple meta-analysis model	248
5.1.1	Maximum likelihood estimation	250
5.1.2	Quadratic unbiased estimation for σ^2	255
5.1.3	Statistical inference	262
5.1.4	Robust/median meta-analysis	268
5.2	Meta-analysis model with covariates	272
5.2.1	Maximum likelihood estimation	273
5.2.2	Quadratic unbiased estimation for σ^2	276
5.2.3	Hypothesis testing	277
5.3	Multivariate meta-analysis model	278
5.3.1	The model	281
5.3.2	Maximum likelihood estimation	282
5.3.3	Quadratic estimation of the heterogeneity matrix . .	285
5.3.4	Test for homogeneity	289
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	Computational issues of nonlinear least squares . . .	295
6.1.3	Distribution-free estimation	296
6.1.4	Testing for the presence of random effects	297
6.1.5	Asymptotic properties	298
6.1.6	Example: Log-Gompertz growth curve	298
6.2	Varied matrix of random effects	304
6.2.1	Maximum likelihood estimation	304
6.2.2	Distribution-free variance parameter estimation . . .	307
6.2.3	GEE and iteratively reweighted least squares	307
6.2.4	Example: Logistic curve with random asymptote . .	309
6.3	Three types of nonlinear marginal models	315
6.3.1	Type I nonlinear marginal model	317

6.3.2	Type II nonlinear marginal model	318
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	320
6.4.1	A robust feature of total GEE	323
6.4.2	Expected Newton–Raphson algorithm for total GEE	323
6.4.3	Total GEE for mixed effects model	324
6.4.4	Total GEE for the LME model	324
6.4.5	Example (continued): Log-Gompertz curve	326
6.5	Summary points	327
7	Generalized Linear Mixed Models	329
7.1	Regression models for binary data	330
7.1.1	Approximate relationship between logit and probit	334
7.1.2	Computation of the logistic-normal integral	337
7.1.3	Log-likelihood and its numerical properties	349
7.1.4	Unit step algorithm	350
7.2	Binary model with subject-specific intercept	351
7.2.1	Consequences of ignoring a random effect	354
7.2.2	ML logistic regression with a fixed subject-specific intercept	355
7.2.3	Conditional logistic regression	356
7.3	Logistic regression with random intercept	358
7.3.1	Maximum likelihood	359
7.3.2	Fixed sample likelihood approximation	365
7.3.3	Quadratic approximation	367
7.3.4	Laplace approximation to the likelihood	368
7.3.5	VARLINK estimation	371
7.3.6	Beta-binomial model	372
7.3.7	Statistical test of homogeneity	374
7.3.8	Asymptotic properties	378
7.4	Probit model with random intercept	378
7.4.1	Laplace and PQL approximations	379
7.4.2	VARLINK estimation	380
7.4.3	Heckman method for probit model	380
7.4.4	Generalized estimating equations approach	380
7.5	Poisson model with random intercept	382
7.5.1	Poisson regression for count data	383
7.5.2	Clustered count data	384
7.5.3	Fixed intercepts	385
7.5.4	Conditional Poisson regression	386
7.5.5	Negative binomial regression	387
7.5.6	Normally distributed intercepts	392
7.5.7	Exact GEE for any distribution	393

7.5.8	Exact GEE for balanced count data	394
7.5.9	Heckman method for the Poisson model	396
7.5.10	Tests for overdispersion	397
7.6	Random intercept model: overview	398
7.7	Mixed models with multiple random effects	399
7.7.1	Multivariate Laplace approximation	400
7.7.2	Logistic regression	400
7.7.3	Probit regression	405
7.7.4	Poisson regression	406
7.8	Homogeneity tests	408
7.9	GLMM and simulation methods	409
7.9.1	General form of GLMM via the exponential family	409
7.9.2	Monte Carlo for ML	410
7.9.3	Fixed sample likelihood approach	411
7.10	GEE for clustered marginal GLM	414
7.10.1	Variance least squares	416
7.10.2	Limitations of the GEE approach	418
7.10.3	Marginal or conditional model?	420
7.11	Criteria for MLE existence for binary model	422
7.12	Summary points	427
8	Nonlinear Mixed Effects Model	431
8.1	Introduction	431
8.2	The model	432
8.3	Example: Height of girls and boys	435
8.4	Maximum likelihood estimation	438
8.5	Two-stage estimator	441
8.5.1	Maximum likelihood estimation	442
8.5.2	Method of moments	443
8.5.3	Drawback of two-stage estimation	444
8.5.4	Further discussion	444
8.5.5	Two-stage method in the presence of a common parameter	445
8.6	First-order approximation	445
8.6.1	GEE and MLE	446
8.6.2	Method of moments and VLS	446
8.7	Lindstrom–Bates estimator	447
8.7.1	What if matrix $\mathbf{D}$ is not positive definite?	449
8.7.2	Relation to the two-stage estimator	450
8.7.3	Computational aspects of penalized least squares	451
8.8	Likelihood approximations	452
8.8.1	Likelihood approximation at zero	452
8.8.2	Laplace and PQL approximations	453
8.9	One-parameter exponential model	455
8.9.1	Maximum likelihood estimator	455

8.9.2	First-order approximation	456
8.9.3	Two-stage estimator	458
8.9.4	Lindstrom–Bates estimator	460
8.10	Asymptotic equivalence of the TS and LB estimators	462
8.11	Bias-corrected two-stage estimator	463
8.12	Distribution misspecification	465
8.13	Partially nonlinear marginal mixed model	468
8.14	Fixed sample likelihood approach	470
8.15	Estimation of random effects and hypothesis testing	471
8.15.1	Estimation of the random effects	471
8.15.2	Hypothesis testing for NLME model	472
8.16	Example (continued)	473
8.17	Practical recommendations	474
8.18	Appendix: Proof of theorem on equivalence	475
8.19	Summary points	479
9	Diagnostics and Influence Analysis	481
9.1	Introduction	481
9.2	Influence analysis for linear regression	483
9.3	The idea of infinitesimal influence	485
9.3.1	Data influence	486
9.3.2	Model influence	486
9.4	Linear regression model	487
9.4.1	Influence of the dependent variable	489
9.4.2	Influence of the continuous explanatory variable	489
9.4.3	Influence of the binary explanatory variable	491
9.4.4	Influence on the predicted value	492
9.4.5	Case or group deletion	493
9.4.6	Influence on regression characteristics	495
9.4.7	Example 1: Women’s body fat	496
9.4.8	Example 2: Gypsy moth study	501
9.5	Nonlinear regression model	503
9.5.1	Influence of the dependent variable on the LSE	504
9.5.2	Influence of the explanatory variable on the LSE	504
9.5.3	Influence on the predicted value	504
9.5.4	Influence of case deletion	505
9.5.5	Example 3: Logistic growth curve model	505
9.6	Logistic regression	508
9.6.1	Influence of the covariate on the MLE	509
9.6.2	Influence on the predicted probability	509
9.6.3	Influence of the case deletion on the MLE	510
9.6.4	Sensitivity to misclassification	511
9.6.5	Example: Finney data	516
9.7	Influence of correlation structure	518
9.8	Influence of measurement error	518

9.9	Influence analysis for the LME model	522
9.9.1	Example: Weight versus height	526
9.10	Appendix: MLE derivative with respect to σ^2	527
9.11	Summary points	529
10	Tumor Regrowth Curves	531
10.1	Survival curves	534
10.2	Double-exponential regrowth curve	536
10.2.1	Time to regrowth, T_R	539
10.2.2	Time to reach specific tumor volume, T_*	539
10.2.3	Doubling time, T_D	539
10.2.4	Statistical model for regrowth	540
10.2.5	Variance estimation for tumor regrowth outcomes	542
10.2.6	Starting values	542
10.2.7	Example: Chemotherapy treatment comparison	543
10.3	Exponential growth with fixed regrowth time	547
10.3.1	Statistical hypothesis testing	548
10.3.2	Synergistic or supra-additive effect	548
10.3.3	Example: Combination of treatments	549
10.4	General regrowth curve	553
10.5	Double-exponential transient regrowth curve	554
10.5.1	Example: Treatment of cellular spheroids	561
10.6	Gompertz transient regrowth curve	562
10.6.1	Example: Tumor treated in mice	563
10.7	Summary points	565
11	Statistical Analysis of Shape	567
11.1	Introduction	567
11.2	Statistical analysis of random triangles	569
11.3	Face recognition	572
11.4	Scale-irrelevant shape model	574
11.4.1	Random effects scale-irrelevant shape model	575
11.4.2	Scale-irrelevant shape model on the log scale	576
11.4.3	Fixed or random size?	577
11.5	Gorilla vertebrae analysis	578
11.6	Procrustes estimation of the mean shape	579
11.6.1	Polygon estimation	582
11.6.2	Generalized Procrustes model	583
11.6.3	Random effects shape model	583
11.6.4	Random or fixed (Procrustes) effects model?	585
11.7	Fourier descriptor analysis	585
11.7.1	Analysis of star shape	586
11.7.2	Random Fourier descriptor analysis	592
11.8	Summary points	593

12 Statistical Image Analysis	595
12.1 Introduction	595
12.1.1 What is a digital image?	596
12.1.2 Image arithmetic	597
12.1.3 Ensemble and repeated measurements	598
12.1.4 Image and spatial statistics	598
12.1.5 Structured and unstructured images	598
12.2 Testing for uniform lighting	599
12.2.1 Estimating light direction and position	600
12.3 Kolmogorov–Smirnov image comparison	601
12.3.1 Kolmogorov–Smirnov test for image comparison	602
12.3.2 Example: Histological analysis of cancer treatment	602
12.4 Multinomial statistical model for images	603
12.4.1 Multinomial image comparison	605
12.5 Image entropy	606
12.5.1 Reduction of a gray image to binary	607
12.5.2 Entropy of a gray image and histogram equalization	609
12.6 Ensemble of unstructured images	610
12.6.1 Fixed-shift model	612
12.6.2 Random-shift model	613
12.6.3 Mixed model for gray images	616
12.6.4 Two-stage estimation	618
12.6.5 Schizophrenia MRI analysis	620
12.7 Image alignment and registration	622
12.7.1 Affine image registration	624
12.7.2 Weighted sum of squares	625
12.7.3 Nonlinear transformations	625
12.7.4 Random registration	626
12.7.5 Linear image interpolation	627
12.7.6 Computational aspects	628
12.7.7 Derivative-free algorithm for image registration	629
12.7.8 Example: Clock alignment	631
12.8 Ensemble of structured images	631
12.8.1 Fixed affine transformations	631
12.8.2 Random affine transformations	632
12.9 Modeling spatial correlation	633
12.9.1 Toeplitz correlation structure	635
12.9.2 Simultaneous estimation of variance and transform parameters	639
12.10 Summary points	639