

Likelihood Theory and Methods for Generalized Linear Mixed Models

by **Aishwarya Bhaskaran**

Thesis submitted in fulfilment of the requirements for
the degree of

Doctor of Philosophy in Mathematics

under the supervision of **Prof. Matt P. Wand** and
Dr. Joanna Wang

University of Technology Sydney
Faculty of Science

Submitted in September 2022

Certificate of original authorship

I, **Aishwarya Bhaskaran**, declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy in Mathematics, in the School of Mathematical and Physical Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by an Australian Government Research Training Program.

Signature:

Production Note:
Signature removed prior to publication.

Date: September 23, 2022

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Matt Wand. This journey was definitely a challenging one and I would like to thank you for all your patience, guidance and constant encouragement throughout the research process. I have indeed learnt a lot from your valuable insights and advice. Thank you so much for always looking out for my well-being as well. This milestone would not have been possible without your help and I am truly grateful.

Next, I would like to thank all my dear friends whom I have met throughout my life. Thank you all for constantly cheering me up, being such great listening ears and for always encouraging me to go the extra mile. I am really glad to have met each and every one of you.

Last but definitely not the least, I would like to give special thanks to my dear family. Thank you for always supporting me in my decisions and for offering me advice when I needed it. Thanks for also being so understanding especially these past three years and for always making sure that I am alright. I really appreciate everything that you all have done and will continue to do for me.

List of papers/publications

The following list of paper and publications awards relate to the work presented in this thesis:

- Jiang, J., Wand, M.P. and Bhaskaran, A. (2022), Usable and precise asymptotics for generalized linear mixed model analysis and design. *Journal of the Royal Statistical Society, Series B*, 84: 55-82. DOI: 10.1111/rssb.12473.
- Bhaskaran, A. and Wand, M.P.(2023), Dispersion parameter extension of precise generalized linear mixed model asymptotics. *Statistics and Probability Letters*, 193, Article 109691.

Notation

In this chapter, we introduce acronyms that are frequently used throughout this thesis.

Acronyms

Table 1: Table with acronyms used in the thesis with their meanings.

Acronym	Meaning
GLMM	Generalized linear mixed model
GVA	Gaussian variational approximation
MLE	Maximum likelihood estimation
TAP	Thouless-Anderson-Palmer

Contents

1	Introduction and Background	2
1.1	Introduction	2
1.2	Thesis Aim	3
1.3	Outline	3
1.4	Matrix Theory	4
1.4.1	Difference Between Two Matrix Inverses	4
1.4.2	Other Useful Matrix Identities	4
1.4.3	Block Matrix Inversion	5
1.4.4	The vec and vech Operators	6
1.4.4.1	The Commutation Matrix	7
1.4.4.2	The Duplication Matrix	7
1.4.5	Kronecker Products and Related Properties	8
1.4.6	Vector and Matrix Norms	8
1.4.6.1	Euclidean Norm	8
1.4.6.2	Frobenius Norm	8
1.4.6.3	Spectral Norm	9
1.4.7	Eigenvalue Bound Results	10
1.4.7.1	Matrix Identities from Harville (1977)	10
1.4.8	Vector Differential Calculus	11
1.5	Key Integral Results	12
1.5.1	Useful Integral Results	12
1.5.2	Integral Form of the Matrix Square Root	13
1.6	Key Expectation Results	13
1.6.1	Law of Total Expectation	13
1.6.2	Jensen's Inequality	13
1.6.3	Markov's Inequality	14
1.6.4	Cauchy-Schwarz Inequality	14
1.7	Exponential Families	14
1.7.1	One-Parameter Exponential Families	14
1.7.2	Two-Parameter Exponential Families	17
1.8	Generalized Linear Mixed Models	17
1.9	Maximum Likelihood for Generalized Linear Mixed Models	19
1.9.1	The Likelihood Function	19
1.9.2	Maximum Likelihood Estimation	20

1.9.3	Asymptotic Properties of Maximum Likelihood Estimators for Generalized Linear Mixed Models	21
1.10	Asymptotics	23
1.10.1	Convergence of Random Variables	23
1.10.1.1	Convergence in Probability	23
1.10.1.2	Convergence in Distribution	23
1.10.1.3	Continuous Mapping Theorem	24
1.10.1.4	Slutsky's Theorem	24
1.10.1.5	Cramér-Wold Device	24
1.10.2	Stochastic Order Notation	25
1.10.3	Other Tools for Working with Asymptotic Expansions	26
1.10.3.1	Inversion of Asymptotic Series	26
1.11	Frequentist Variational Approximations	27
1.11.1	Thouless-Anderson-Palmer Variational Approach	29
2	Preliminary Lemmas and Their Proofs	31
2.1	Lemma 1	31
2.2	Lemma 2	32
2.3	Lemma 3	33
2.4	Appendix	34
2.4.1	Proof of Lemma 1	34
2.4.2	Proof of Lemma 2	35
2.4.2.1	A Fundamental Inequality for the Spectral Norm of a Vectorised Matrix	35
2.4.2.2	Notational Definitions	35
2.4.2.3	Derivation of (2.7)	36
2.4.2.4	Expression for (2.6) with Lagrange Form of Remainder	37
2.4.2.5	Spectral Norm Bounding of (2.9)	38
2.4.2.6	Strategy for Proving (2.10)	38
2.4.2.7	Proof of Result (2.16)	39
2.4.2.8	Proof of Result (2.17)	45
2.4.2.9	Summary of Moment Assumptions	48
2.4.2.10	Succinct Expression for Moment Assumptions	49
2.4.2.11	A Sufficient Condition for the Moment Assumptions	50
2.4.3	Proof of Lemma 3	51
2.4.3.1	Matrix Extension of Results Concerning Integrals of Half-Cauchy Forms	51
2.4.3.2	Derivation of Integrand Expressions	52
2.4.3.3	Succinct Expressions for the Components in (2.53)	59
2.4.3.4	Simplification of Integrals	60
2.4.3.5	Explicit Expressions for (2.54)	60
2.4.3.6	Convergence in Probability Limits of the Functions in (2.55)	60
2.4.4	Multivariate Integral Limits for the Matrix Square Root Result	61
2.4.4.1	Overview of this Appendix	61
2.4.4.2	Computing Spectral Norms	61

2.4.4.3	Verifying Convergence in Probability Limits of the Functions in (2.55)	68
2.4.4.4	Conclusion for Multivariate Integral Limits for the Matrix Square Root Result	71
3	Usable Asymptotic Normality Results and Inference for Gaussian Response Linear Mixed Models	72
3.1	Model Description	73
3.2	Notation Required for Fisher Information Calculations	74
3.3	Asymptotic Normality Theorem	74
3.4	Appendix	75
3.4.1	Linear Mixed Models with Multivariate Fixed and Random Effects	75
3.4.2	Expression for Top Left Block of Fisher Information Matrix . . .	77
3.4.2.1	Top Left Block of (3.5)	77
3.4.2.2	Top Right Block of (3.5)	77
3.4.2.3	Bottom Left Block of (3.5)	78
3.4.2.4	Bottom Right Block of (3.5)	78
3.4.3	Expression for Bottom Right Block of Fisher Information Matrix	82
3.4.3.1	Top Left Block of (3.10)	82
3.4.3.2	Top Right Block of (3.10)	83
3.4.3.3	Bottom Left Block of (3.10)	84
3.4.3.4	Bottom Right Block of (3.10)	84
3.4.4	The Inverse of the Fisher Information Matrix	85
3.4.4.1	Expression for Top Left Block of Inverse Fisher Information Matrix	85
3.4.4.2	Expression for Bottom Right Block of Inverse Fisher Information Matrix	87
3.4.5	Derivation of the Final Asymptotic Normality Result for Gaussian Response Linear Mixed Models	90
4	Usable Asymptotic Normality Results and Inference for Generalized Linear Mixed Models	93
4.1	Model Description	94
4.2	Notation	95
4.3	Asymptotic Normality Theorem	96
4.4	Dispersion Parameter Extension	97
4.5	Appendix	98
4.5.1	Multivariate Extension of (2.6) of Tierney et al. (1989)	98
4.5.1.1	Overview	98
4.5.1.2	Multivariate Derivative Notation	99
4.5.1.3	Check of the Miyata (2004) Appendix A Result for the Univariate Case	99
4.5.1.4	The Multivariate Case	100
4.5.1.5	Final Expression for the Multivariate Extension of (2.6) of Tierney et al. (1989)	101
4.5.2	Proof of Theorem 12	101
4.5.2.1	Constructing the Fisher Information Matrix	101

4.5.2.2	Expression for Conditional Density Function	102
4.5.2.3	Introduction of Useful Notation and its Properties . . .	103
4.5.2.4	Computing an Asymptotic Approximation for the First Entry in (4.7)	105
4.5.2.5	Computing an Asymptotic Approximation for the Sec- ond Entry in (4.7)	109
4.5.2.6	Computing an Asymptotic Approximation for the Third Entry in (4.7)	112
4.5.2.7	The Quadratic Conditional Expectations of the Scores .	115
4.5.2.8	Treating the Leading Term of the (2,2)-Entry of the Fisher Information Matrix	122
4.5.2.9	The Fisher Information Matrix	123
4.5.2.10	The Inverse of the Fisher Information Matrix	123
4.5.2.11	Derivation of the Final Asymptotic Normality Result for Generalized Response Linear Mixed Models	126
4.5.3	The Reciprocal Dispersion Parameter Fisher Information Block for Gamma Responses	128
4.5.3.1	The Conditional Density Function	128
4.5.3.2	The Score of the Reciprocal Dispersion Parameter . . .	131
4.5.3.3	Computing the Fisher Information Block for the Recip- rocal Dispersion Parameter	132
4.5.3.4	Asymptotic Normality and Variance Results for the Max- imum Likelihood Estimator of the Reciprocal Dispersion Parameter	145
4.5.3.5	Asymptotic Normality and Variance Results for the Max- imum Likelihood Estimator of the Dispersion Parameter	145
5	Consequences and Applications of Asymptotic Normality Results	146
5.1	Asymptotically Valid Inference	147
5.1.1	Construction of Asymptotically Valid Confidence Intervals	147
5.1.2	Simulation Study	148
5.2	Approximate Optimal Design	155
5.2.1	Background and Model Description	155
5.2.2	Approximate Locally D-Optimal Design Determination	156
5.2.3	Illustration of Theorem 13	158
5.3	Appendix	159
5.3.1	Model Description	159
5.3.2	Asymptotic Assumption for Support Point Sample Sizes	161
5.3.3	Useful Notation	161
5.3.4	Key Moment Results	162
5.3.5	The Fisher Information Matrix	164
5.3.6	The Asymptotic D-Optimality Criterion	164
5.3.7	Alternative Final Asymptotic D-optimality Criterion	166
5.3.8	Special Distribution Cases	167
6	Thouless-Anderson-Palmer Enhancement of Generalized Linear Mixed Models	169

6.1	Model Description	170
6.2	The Gaussian Variational Approximate Log-Likelihood	170
6.3	Overview of Thouless-Anderson-Palmer Enhancement	172
6.4	The Thouless-Anderson-Palmer Approximate Negative Log-Likelihood	173
6.5	Thouless-Anderson-Palmer Enhancement for Poisson Generalized Linear Mixed Models	175
6.5.1	The Gaussian Variational Approximate Log-Likelihood for Simu- lation Set-Up	175
6.5.2	The Thouless-Anderson-Palmer Negative Approximate Log-Likelihood for Simulation Set-Up	176
6.5.3	Optimisation Issues	176
6.5.3.1	A Simplified Version of the Optimisation Problem	177
6.5.3.2	Simplified Simulation Study	178
6.5.3.3	Results and Conclusion	179
6.5.4	Simulation Study	181
6.6	Appendix	185
6.6.1	Proof of Result 2	185
6.6.1.1	Main Quantity in Onsager's Correction Term	185
6.6.1.2	An Explicit Expression for the First Term in (6.10)	185
6.6.1.3	An Explicit Expression for the Second Term in (6.10)	189
6.6.1.4	An Explicit Expression for the Third Term in (6.10)	190
6.6.1.5	The Resultant Expression for the Main Quantity in the Onsager's Correction Term	190
6.6.2	Expressing the Main Quantity in the Onsager's Correction Term Using Integral Families	192
6.6.3	Proof of Result 3	193
6.6.3.1	Simplifications in Poisson Case	193
7	Extensions to Noncanonical Link Generalized Linear Mixed Models	195
7.1	Asymptotic Normality Results Involving Noncanonical Links	196
7.1.1	Model Description	196
7.1.2	Notation	197
7.1.3	Asymptotic Normality Theorem	198
7.2	Thouless-Anderson-Palmer Approach Involving Noncanonical Links	199
7.2.1	Model Description	199
7.2.2	The Gaussian Variational Approximate Log-Likelihood	200
7.2.3	Overview of Thouless-Anderson-Palmer Enhancement	200
7.3	Appendix	202
7.3.1	Constructing the Fisher Information Matrix	202
7.3.2	Expression for Conditional Density Function	202
7.3.3	Deriving Expressions for the Expectation and Variance of the Response Variable	203
7.3.4	Introduction of Useful Notation and Its Properties	205
7.3.5	Key Conditional Moment Results	205
7.3.6	Computing an Asymptotic Approximation for the First Entry in (7.9)	207

7.3.6.1	The First Term of the First Score	210
7.3.6.2	The Other Terms of the First Score	210
7.3.6.3	Overall Leading Term Expression for the First Score . .	210
7.3.7	Computing an Asymptotic Approximation for the Second Entry in (7.9)	210
7.3.7.1	The First Term of the Second Score	211
7.3.7.2	The Other Terms of the Second Score	211
7.3.7.3	Overall Leading Term Expression for the Second Score	211
7.3.8	Computing an Asymptotic Approximation for the Third Entry in (7.9)	211
7.3.8.1	The First Term of the Third Score	212
7.3.8.2	The Other Terms of the Third Score	212
7.3.8.3	Overall Leading Term Expression for the Third Score .	212
7.3.9	The Quadratic Conditional Expectations of the Scores	212
7.3.9.1	The Conditional Expectation of the Square of the First Score	213
7.3.9.2	The Conditional Expectation of the Square of the Second Score	213
7.3.9.3	The Conditional Expectation of the Square of the Third Score	215
7.3.10	The Fisher Information Matrix	216
7.3.11	The Inverse of the Fisher Information Matrix	218
7.3.12	Final Asymptotic Normality Result	218
8	Discussion and Conclusion	219
9	References	223

Abstract

Generalized linear mixed models are an essential group of models for analysing many present-day complex data sets, especially those that contain non-normal and correlated response data. Despite the large volume of research concerning this group of models, there is very little theory concerning the statistical properties of maximum likelihood estimators for generalized linear mixed models. Existing theoretical results available for the asymptotic variance-covariance matrix for such estimators contain limits and expectations over the response distribution, hence such results are not in ready-to-use forms when carrying out tasks such as constructing studentized confidence intervals or optimal design determination. In this thesis, we derive precise asymptotic results for likelihood-based generalized linear mixed model analysis. The novel asymptotic normality results are derived for both cases involving either a canonical or noncanonical link function. In our approach, we derive the exact leading term behaviour of the Fisher information matrix when both the number of groups and number of observations within each group diverge. This leads to asymptotic normality results with explicit and simple studentizable forms. The implications of these results in optimal design theory is also explored, leading to simpler and more direct determination of approximate locally D-optimal designs. Towards the end of this thesis, a Thouless-Anderson-Palmer approach is introduced for modern statistical inference for generalized linear mixed models. Such methods have proven to provide accurate approximations to problems arising in machine learning contexts. However, statistical applications such as generalized linear mixed model analysis have not been investigated. Thus, we derive results for implementing the Thouless-Anderson-Palmer frequentist variational approach to generalized linear mixed models and analyse the accuracy of its variational estimates.