UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Science

Conjugate Generalized Linear Mixed Models with Applications

by

Jarod Yan Liang Lee

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Sydney, Australia

2017

Certificate of Original Authorship

I certify that the work in this thesis has not been previously submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

I also certify that this thesis has been written by me. Any help that I have received in my research and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

Date:

Acknowledgements

The successful completion of this research work is not only the result of my own effort, but also a series of contribution from many others. My deepest gratitude goes to my supervisors Professor Louise Ryan, Professor Peter Green and Professor James Brown for always being encouraging, supportive and inspirational. Their great insight into statistical methodology and applied statistics have profoundly influenced the way I approach statistical research. Without them, I could not have finished my PhD.

I am grateful to the Australian Bureau of Statistics for providing funding that allowed me to present my work at various conferences. Also, thanks to Professor Michael Martin and Professor Alan Welsh who provided advice and encouragement during my undergraduate years at the Australian National University, where I first became interested in statistics.

Last but certainly not least I would like to thank my family and my partner for all of their support, patience and understanding. I consider myself very fortunate to have such a great family.

Jarod Lee

Sydney, Australia, 2017

ABSTRACT

Conjugate Generalized Linear Mixed Models with Applications

by

Jarod Yan Liang Lee

This thesis focuses on the development of conjugate generalized linear mixed models (CGLMMs), which is a computationally efficient modelling framework for longitudinal and multilevel data where the likelihood can be expressed in closed-form. We focus on the scenario where the random effects are mapped uniquely onto the grouping structure and are independent between groups. Compared with conventional inference methods for generalized linear mixed models (GLMMs), CGLMMs allow the parameters to be estimated directly without the need for computational intensive numerical approximation methods. The proposed framework has important implications in terms of distributed computing, privacy preservation in large-scale administrative databases and discrete choice models, which we illustrate using several real data. Altogether, CGLMMs prove to be a credible inference framework and a good alternative to GLMMs, especially when dealing with a large amount of data and/or privacy is of concern.

Contents