

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

GENERAL ALGORITHMS BASED ON LEAST SQUARES CALCULATIONS
FOR MAXIMUM LIKELIHOOD ESTIMATION
IN MULTIPARAMETER MODELS

A thesis presented in partial fulfilment of the requirements
for the degree of Ph.D. in Statistics at Massey University.

by
W. Douglas Stirling

1985

ABSTRACT

This thesis develops algorithms for maximum likelihood estimation that can be implemented using a sequence of weighted least squares computations and examines their properties.

Standard least squares algorithms are first described and their execution times, storage requirements and accuracies are compared. The Givens QR algorithm uses less storage than other algorithms of comparable accuracy and in good implementations is virtually as fast as them if there are several explanatory variables. A version that can be used for constrained least squares is described; it is used for least squares calculations in the remainder of the thesis.

In many maximum likelihood problems, the likelihood can be written as a sum of functions called log-likelihood components and these often depend on the unknown parameters only through one or two quantities called systematic parts. For these models, a class of algorithms called NRL algorithms approaches the maximum likelihood estimate with a sequence of least squares calculations. For many common models, the Newton-Raphson algorithm and Fisher's scoring technique are particular NRL algorithms. Implementation of NRL algorithms is described in detail and the relative merits of the various NRL algorithms are discussed. If the NR algorithm is in the class, it converges best near the maximum likelihood estimate, but other NRL algorithms may perform better in the first few iterations. Several examples are analysed to illustrate the various possible methods.

When the maximum likelihood estimates of some parameters can be written as explicit functions of the rest, the convergence of the NR and NRL algorithms can often be improved by adjusting these parameters between iterations. The relationship of this technique to elimination of these parameters from the likelihood is investigated. In several types of model, including nonlinear least squares, adjustment can be

performed without slowing the NRL iterations. A related more general method is also described for improving NRL iterations when some parameters are linear and some are nonlinear in the systematic parts.

Another general algorithm called the EM algorithm is described. It can be applied to several types of model for which the NRL algorithm cannot be used. In some models, it can also be implemented using a sequence of least squares calculations, but for applications where both EM and NRL algorithms can be used, the latter usually converge faster.

Finally, in two appendices, Fortran subroutines that can be used to implement the algorithms in the thesis are described and listed.

ACKNOWLEDGEMENTS

I would like to thank Dr R J Brook for supervising this thesis. I am also grateful to the Department of Mathematics and Statistics for allowing me time to complete this thesis and to Massey University for making available computer facilities for typing it.

I finally wish to acknowledge the support my wife Sue has given me and to thank her and my children David and George for tolerating the evening and weekends I have spent on the thesis in the last few years.

TABLE OF CONTENTS

Chapter 1. Introduction

| | |
|-------------------------------------|---|
| 1.1 Purpose and Scope of the Thesis | 1 |
| 1.2 Structure of the Thesis | 3 |

Chapter 2. Least Squares Algorithms

| | |
|--|----|
| 2.1 The Normal Linear Model | 5 |
| 2.2 Algorithms Based on the Normal Equations | 7 |
| 2.3 QR Algorithms | 9 |
| 2.4 A Numerical Comparison of the Accuracy of the Algorithms | 17 |
| 2.5 Linear Constraints | 25 |
| 2.6 Dependencies Between Explanatory Variables | 27 |
| 2.7 A Set of Fortran Subroutines for Least Squares | 29 |

Chapter 3. Models with a Single Systematic Part

| | |
|--|----|
| 3.1 Generalized Linear Models | 31 |
| 3.2 Some Models Outside the Class of Generalized Linear Models | 35 |
| 3.3 General Optimization Algorithms for Maximum Likelihood | 44 |
| 3.4 Implementation of Unmodified NRL Algorithms | 52 |
| 3.5 Asymptotic Performance of NRL Algorithms | 60 |
| 3.6 The FS algorithm and its Relationship to NRL | 74 |
| 3.7 Implementation of Iteratively Reweighted Least Squares Algorithms in the Initial Iterations | 77 |
| 3.8 Constraints | 84 |
| 3.9 Variances and Tests | 86 |

Chapter 4. Models with Two or More Systematic Parts

| | | |
|-----|---|-----|
| 4.1 | Fitting Models with Iteratively Reweighted Least Squares | 90 |
| 4.2 | Variances and Tests | 99 |
| 4.3 | Examples : Normal Models with Variance a Function of Mean | 102 |
| 4.4 | Examples : Normal Models with Variance a Function of Explanatory Variables | 113 |
| 4.5 | Examples : Negative Binomial Models | 122 |
| 4.6 | Examples : Robust Estimation in Linear Models | 128 |

Chapter 5. Elimination and Adjustment of Parameters

| | | |
|-----|--|-----|
| 5.1 | General Description of Elimination and Adjustment | 138 |
| 5.2 | Elimination and Adjustment Applied to the NR, NRL and FS Algorithms | 140 |
| 5.3 | Nonlinear Least Squares | 145 |
| 5.4 | A Nonlinear Least Squares Example | 150 |
| 5.5 | Other Applications of Elimination and Adjustment | 154 |
| 5.6 | Systematic Parts with Linear and Nonlinear Parameters | 159 |
| 5.7 | Concluding Remarks about Elimination and Adjustment | 162 |

Chapter 6. The EM Algorithm

| | | |
|-----|---|-----|
| 6.1 | General Description of the EM Algorithm | 164 |
| 6.2 | Examples : Missing Data | 170 |
| 6.3 | Examples : Mixtures and Totals | 174 |
| 6.4 | Examples : Variance Components | 176 |
| 6.5 | Examples : Hyperparameter Estimation | 179 |
| 6.6 | Examples : Robust Estimation in Linear Models | 181 |
| 6.7 | Concluding Remarks About the EM Algorithm | 187 |

Chapter 7. Conclusion 188

Appendices

| | | |
|---|---|-----|
| A | Fortran Subroutines for Model Fitting -- Parameters | 192 |
| B | Fortran Subroutines for Model Fitting -- Code | 202 |

| | | |
|--|-------------------|-----|
| | <u>References</u> | 215 |
|--|-------------------|-----|