

University of Wollongong - Research Online

Thesis Collection

Title: Curvature measures for generalized linear models

Author: Bernard A Ellem

Year: 1999

Repository DOI:

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Research Online is the open access repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

1999

Curvature measures for generalized linear models

Bernard A. Ellem
University of Wollongong

Follow this and additional works at: <https://ro.uow.edu.au/theses>

University of Wollongong

Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following: This work is copyright. Apart from any use permitted under the Copyright Act 1968, no part of this work may be reproduced by any process, nor may any other exclusive right be exercised, without the permission of the author. Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material.

Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

Unless otherwise indicated, the views expressed in this thesis are those of the author and do not necessarily represent the views of the University of Wollongong.

Recommended Citation

Ellem, Bernard A., Curvature measures for generalized linear models, Doctor of Philosophy thesis, School of Mathematics and Applied Statistics, University of Wollongong, 1999. <https://ro.uow.edu.au/theses/2045>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

CURVATURE MEASURES FOR GENERALIZED LINEAR MODELS

A thesis submitted in fulfilment of the
requirements for the award of the degree

DOCTOR of PHILOSOPHY

from

UNIVERSITY of WOLLONGONG

by

BERNARD A. ELLEM, B.Sc, M.Ec *NE*

SCHOOL of MATHEMATICS

and

APPLIED STATISTICS

1999

Declaration

In accordance with the regulations of the University of Wollongong, I hereby state that the work described here is my original work, except where due references are made, and has not been submitted for a degree in any university or institution.

Bernard A. Ellem

For *Cherie*.

Exegi monumentum aere perennius
regalique situ pyramidum altius . . .

Q. Horatii Flacci

Carminum, Liber III, Carmen XXX

Acknowledgments

This research has been made possible by the interest, involvement and generosity of others.

I would like to thank my wife Cherie for her constant and unflagging support during the extended period of study which often meant I was absent at inconvenient times. This work would have been impossible without her full support.

My supervisor, Professor David Griffiths, deserves special mention due to his patient guidance from the initial program of reading through to the final write-up where his critical review was invaluable. His support throughout the various stages of the research was very much appreciated, and his suggestions at all stages have proved to be very worthwhile.

I commend the University of Wollongong on having the vision to provide the mechanism for part-time study at the doctoral level at a time when it was not always available elsewhere.

The Staff of the School of Mathematics and Applied Statistics at the University of Wollongong are to be congratulated for their continued encouragement of my endeavours at all stages of the program.

By providing an environment for research and the means to continue my study program, Charles Sturt University through the Faculty of Science and Agriculture have aided the completion of this doctoral thesis.

I would also like to record the support given by NSW Agriculture (Biometrics Section) during the course of this investigation.

Finally, my thanks go to Karen and Peter Hiscocks for their generosity over the long stretch of time that this study occupied.

Contents

Abstract	x
1 Introduction	1
1.1 Background	1
1.2 Rationale for the study	2
1.3 Role of Curvature Measures in Nonlinear Regression	3
1.4 Curvature Measures	6
1.4.1 Linear Model	8
1.4.2 Non-linear Model	10
1.4.3 Higher Dimensions	12
1.4.4 Practical Considerations	17
1.5 Generalized Linear Models	20
1.5.1 Leverage	21
1.6 Exponential Families	24
1.7 Curved Exponential Families	25
1.8 Tensor Notation	28
1.8.1 Indexing	29
1.8.2 Summation Convention	29
1.8.3 Tensor Laws	30
1.8.4 Coordinate Free Methods	31
1.9 The Generalization	32

2	Differential Geometric Approach	35
2.1	Preliminaries	35
2.1.1	Likelihood	36
2.1.2	Regularity Conditions	36
2.2	Tangent Spaces	37
2.3	Inner Product	40
2.4	Metric Tensor	41
2.4.1	Example 1, Normal distribution with known variance	42
2.4.2	Example 2, Normal distribution with known mean	42
2.4.3	Example 3, Normal distribution	43
2.4.4	Example 4, Multinomial distribution	45
2.4.5	Example 5, Generalized Linear Model	47
2.5	Affine Connection	48
2.6	α -connections	52
2.7	Statistical Interpretation of α -connections	52
2.7.1	Riemann Christoffel Curvature	54
2.8	Equivalence of α , δ and c	54
2.9	Bartlett's Equations	55
2.10	Interpretation of α in the one parameter case	57
2.10.1	Mixture Connection	58
2.10.2	Skewness Connection	59
2.10.3	Information Connection	60
2.10.4	'Normal' Connection	60
2.10.5	Exponential Connection	61
2.10.6	Note	61
2.10.7	Summary	62
2.11	Interpretation of α in the multi-parameter case	64
2.11.1	Mixture Connection	67
2.11.2	Skewness Connection	67
2.11.3	Information Connection	68

<i>CONTENTS</i>	iii
2.11.4 ‘Normal’ Connection	69
2.11.5 Exponential Connection	70
2.12 Dual Space	71
2.13 Generalized Linear Models	71
2.13.1 One-dimensional GLMS	74
2.14 Regression coefficients in GLMs	74
2.14.1 Imbedding	75
2.14.2 Imbedding Theorem	75
2.14.3 Normal Distribution	78
2.14.4 Normal Linear Models	79
2.14.5 Nonlinear Regression	79
2.14.6 Generalized Linear Models	81
2.14.7 Canonical Links	82
2.14.8 Summary	83
2.15 Exponential Connection and GLMs	84
2.15.1 Theorem	84
2.15.2 Preliminaries	84
2.15.3 Proposition	85
2.15.4 Proof	85
2.15.5 Interpretation	87
2.15.6 Canonical Link	88
2.15.7 Non-Canonical Link	88
2.15.8 Discussion	89
2.15.9 Link Adequacy	94
2.15.10 Summary	104
3 α-Curvatures	108
3.1 Introduction	108
3.1.1 Transformation Rule (Γ)	110
3.2 Curvatures	110

3.2.1	Derivation	111
3.2.2	Transformation Rule (H)	112
3.3	Projections	112
3.3.1	Normal Component	112
3.3.2	The Invariance of Intrinsic Curvature	114
3.3.3	Tangential Component	117
3.3.4	Scalar Parameter-effects Curvature	118
3.4	Decomposition	119
3.4.1	Decomposition of Scalar Curvature	120
3.5	Examples	121
3.5.1	Nonlinear Regression	121
3.5.2	Generalized Linear Models	123
3.6	Generalized Nonlinear Models	126
3.6.1	Definition	126
3.6.2	Curvatures	129
3.6.3	Note	130
3.7	Expected and Observed Geometries	132
4	Applications	134
4.1	Tensorial α -connections and GLMs	134
4.1.1	Example	135
4.2	Invariance of Parameter-Effects Curvature	137
4.2.1	Theorem	137
4.2.2	Short Form of Proof	142
4.3	Exponential Curvature	143
4.3.1	Preamble	143
4.3.2	Canonical Links in GLMs	144
4.4	The exponential form of α -curvature	146
4.5	Generalized Nonlinear Models	147
4.6	Bias and Covariance of Estimators	148

4.7	Variance Stabilizing Link Function	149
4.7.1	Other Link Functions	154
5	Extensions and Conclusion	157
5.1	Extensions	157
5.1.1	Leverage in Nonlinear Regression	157
5.1.2	Replication and Curvature	167
5.2	Overall Results	175
5.2.1	Summary	176
A	(Ch. 1)	184
A.1	The Hat Matrix for GLMs	184
A.1.1	Standardized Form	184
A.1.2	Raw Form	185
B	(Ch. 2)	187
B.1	Jeffreys' distance measure	187
B.1.1	Preamble	187
B.1.2	Derivation	187
B.2	Metric Tensor : alternative form	188
B.2.1	Derivation	188
B.3	Metric Tensor : results	189
B.3.1	Metric tensor	189
B.3.2	Affine connection	189
B.3.3	General tensors	190
B.3.4	Imbedding	190
B.4	Riemann Christoffel Curvature Tensor	190
B.5	Exponential Families and 1-connections	192
B.6	Wedderburn's Exponential Form	193
B.7	GLM Notation	198
B.8	Derivation of the Imbedding Theorem	198

B.8.1	Proof	199
B.9	Equivalence	199
C (Ch. 3)		201
C.1	The derivation of α -curvature	201
C.2	The transformation rule for α -curvature	202
C.2.1	Proof	202
C.3	Tensorial normal α -curvature	203
C.3.1	Proof	203
C.3.2	Alternative Derivation	204
C.4	Lemma	205
C.4.1	Proof	206
C.5	Non-tensorial tangential α -curvature	206
C.5.1	Proof	207
D (Ch. 5)		208
D.1	GLIM Output : Test Problem 1	208
D.2	GENSTAT Output : Test Problem 2	212
D.3	Replication results	221
D.3.1	Introduction	221
D.3.2	Metric tensor	221
D.3.3	α -connection	222
D.3.4	Exponential family	223
D.3.5	Curved exponential family	224

List of Figures

1.1	Ratkowsky Problem.	8
1.2	Solution Locus : Linear Model.	9
1.3	Solution Locus : Non-linear Model.	11
1.4	Solution Locus : View 1, Example 2.	14
1.5	Solution Locus : View 2, Example 2.	16
1.6	Solution Locus : Example 3	18
2.1	The tangent space T in parameter space S	37
2.2	Basis vectors span the tangent space.	38
2.3	Neighbouring tangent spaces.	39
2.4	Vector addition for neighbouring parameter spaces.	40
2.5	The basis vectors for neighbouring tangent spaces.	48
2.6	The correspondence between neighbouring basis vector spaces.	49
2.7	Example 1 : Reciprocal Link	105
2.8	Example 3 : Log Link	106
2.9	Example 4 : Square Root Link	107
5.1	Solution Locus (solid curve), Tangent (line) and GLM approxi- mant (crosses $[+]$). The data are shown by the box(\square).	163
5.2	Sum of squares plotted against the parameter $\theta : N = 1$	179
5.3	Sum of squares plotted against the parameter $\theta : N = 2$	180
5.4	Sum of squares plotted against the parameter $\theta : N = 5$	181
5.5	Sum of squares plotted against the parameter $\theta : N = 100$	182
5.6	Solution locus : replication experiment	183

List of Tables

- 1.1 Illustrative Data Set 7
- 1.2 Problem A : Draper and Smith (1981) 13
- 1.3 Two Parameter Example. 15

- 2.1 The interpretation of α , δ and c 55
- 2.2 Conditions for the interpretation of α : single parameter. 63
- 2.3 Conditions for the interpretation of α : multi-parameter. 70
- 2.4 Link functions used in the Examples. 97
- 2.5 Example 1 : Poisson data with reciprocal link 99
- 2.6 Example 1 : Skewness and the standard error of the coefficients . . 99
- 2.7 Example 2 : Poisson data with identity link 100
- 2.8 Example 2 : Skewness and the standard error of the coefficients . . 100
- 2.9 Example 3 : Poisson data with log link 101
- 2.10 Example 3 : Skewness and the standard error of the coefficients . . 101
- 2.11 Example 4 : Poisson data with square root link 102
- 2.12 Example 4 : Skewness and the standard error of the coefficients . . 102

- 3.1 The functions p and q for generalized nonlinear models. 128

- 4.1 The corner—point and group means parameterizations. 136
- 4.2 Alternative symbols for the key values of α 143
- 4.3 Constant information link functions. 150
- 4.4 The canonical parameter function $b(\theta)$ and its derivatives. 152
- 4.5 Link functions for key values of $\alpha(\delta)$ 155

5.1	Summary output : Test problem 1	162
5.2	Data Set with replication	164
5.3	Results summary : Data Set with replication	165
5.4	Leverages summary : Data set with replication	166
5.5	Square root model – replication experiment	169
5.6	‘Typical’ data generated for the replication experiment	169
5.7	Results(averages) for the simulation replication experiments	170
5.8	Results for the ‘typical’ data	170
B.1	Key values of δ	193

Abstract

First addressed by Beale (1960), the use of curvature measures of nonlinearity in nonlinear regression has been elucidated most comprehensively by Bates and Watts (1980). They used differential geometric results that exploit features of the Euclidean space imposed by the Normality assumption. The partitioning of these measures into intrinsic effects (due to the model) and parameter effects (due to the form or parameterization of the model) allows a proper assessment of model departures from linearity. Indeed, the term ‘linear’ has become synonymous with a lack of *both* of these effects, since the commonly designated ‘linear model’ with Normal disturbance does not contain either effect. These curvature measures are used to unravel the effects of model reformulation on convergence of fitting procedures, and on the appropriateness of confidence regions based on the linearization assumption. For model criticism using residual analysis, the presence of intrinsic curvature in a nonlinear regression model can distort the visual assessment procedures borrowed from linear modelling, since the fundamental basis of these procedures can be undermined when the model is nonlinear.

When the disturbances are non-Normal, the consequent geometry is no longer Euclidean, necessitating a different approach, as outlined by Amari (1982a). The required approach generalizes the Euclidean inner product to a metric, and the ordinary derivative to an α -connection. The concept of these α -connections is fundamental to a proper understanding of the role of differential geometry to the investigation of estimator behaviour in the case of non-Normal errors. These connections provide the general method for comparing nearby points in the parameter

space, for general classes of error distributions. In these cases, such a comparison is complicated by the difficulty of the existence of different bases for the neighbouring tangent spaces derived from the likelihood. The exception or special case is the linear model with Normal errors, where no such difficulty arises.

Casting the generalization as being from Normal to non-Normal errors, the extension can be considered to cause an ‘unbundling’ of the statistical properties of estimators, which in the case of Normal errors can be enjoyed simultaneously by the same estimator. In the general non-Normal case, such behaviour can no longer be guaranteed, implying that all properties may need to be considered separately, since, in the general case, specific properties of the estimator are associated with particular values of α .

This thesis outlines the fundamentals of the generalization of curvature measures to models of exponential type, in particular curved exponential families for which generalized linear models are an important subclass. This approach is used to generate insights into the properties of generalized linear models, with particular reference to the *canonical* link function as the non-Normal generalization of a linear model with Normal errors.

Indeed, the underlying ‘theme’ of this study is the investigation of the generalization of ‘linearity’ for the Normal error linear model to the non-Normal error nonlinear model. The potential simultaneity of estimator properties for the Normal distribution does not carry over to the generalization from the Normal to the non-Normal, since now each property has to be investigated separately, for each particular value of α .

As shown in Chapter 2, this individual treatment involves the statistical interpretation of each α -connection to demonstrate how key values of α are associated with estimator properties such as unbiasedness, stability of variance, lack of skewness, ‘normal’ likelihood and sufficiency. In terms of data analysis, all of these investigations need to be performed on the regression coefficients rather than on the fitted value (expectation parameter) scale. This requires the use of curved exponential families involving an imbedding of the regression coefficients in the

original expectation space.

One of the properties of Normal error linear models is estimator sufficiency, which for generalized linear models implies a canonical link function. The associated α -connection is the exponential or Efron connection. This connection could be considered as the springboard for the generalization of Normal error linear models to non-Normal error nonlinear models, since for generalized linear models it mimics the special case of Normal errors, by the conditions under which it vanishes. The investigation of this connection and its special relationship with generalized linear models has generated in Chapter 2 a test of adequacy for canonical link functions, based on the skewness of the regression coefficients.

The generalization of curvature follows a similar path to the α -connections, being a function of them in terms of the expectation parameters. In line with the decomposition demonstrated by Bates and Watts (1980) for Normal errors, generalized α -curvature decomposes into intrinsic and parameter-effects curvature; now, each particular α -curvature is associated with individual properties of the model, depending on the value of α . The other main change from the curvature measures of Bates and Watts is that, in the general case, a contribution to curvature is made from the error distribution as well as from the model and its parameterization. A major new result in Chapter 3 has been the proof of the invariance of intrinsic α -curvature in the general case, using a coordinate based system. A consequence of examining the generalization has been to define in Chapter 3 a class of models, *generalized nonlinear models*, having a non-Normal error distribution and a general nonlinear response function. The relationship of this class with classes of known models such as generalized linear models again raises the question of what is meant by ‘nonlinearity’ in general. Several related derivations such as the invariance of parameter-effects curvature in generalized linear models, and results involving exponential curvature, generalized linear models and generalized nonlinear models verify expected behaviour and highlight the generalizations that are possible.

The generalized curvature measures are shown in Chapter 4 to be related to

quantities of statistical interest such as the bias and covariance of estimators for curved exponential families, mirroring the known situation for nonlinear regression. For generalized linear models, alternative link functions to the canonical can be chosen on the basis of properties such as variance stabilization, ‘normal’ likelihood and lack of skewness. As expected, these links have been shown in Chapter 4 to be associated with specific α -connections. A table is presented of those link functions that produce the required properties on the expected value scale for each error distribution in a generalized linear model.

The special relationship between curvature measures, nonlinear regression and generalized linear models is further demonstrated in Chapter 5 by the use of a new method for nonlinear regression based on a second order approximant to the nonlinear function by means of a special generalized linear model. As expected, such an approximation follows the true function more closely than linearization; this is demonstrated empirically from calculations of leverage, parameter estimates and corresponding interval estimation. All these effects are predicted from considerations based on curvature measures, both intrinsic and parameter-effects.

The effect of replication on curvature is known empirically and theoretically in the case of nonlinear regression. In Chapter 5 it is shown that replication has two implications for the effects of curvature in a generalized nonlinear model. Firstly, the central limit theorem produces convergence to the Normal distribution, so that the error contribution to general α -curvature becomes zero asymptotically. The effect of replication on the model contribution is less clear, since the general limiting case is nonlinear regression if only the error component of α -curvature is considered. Locally, the generalized nonlinear model will be well approximated by a linear model. Secondly, under some conditions, a generalized nonlinear model will converge locally to a generalized linear model with canonical link. However, when the error component and the model component are considered, the overall effect of intense replication will be to produce locally a linear model with Normal errors.