IN PURSUIT OF A SUITABLE ALTERNATIVE TO LEAST SQUARES ESTIMATION IN NORMAL LINEAR MODELS

A THESIS PRESENTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN STATISTICS
AT MASSEY UNIVERSITY

ROBERT HUGH FLETCHER

1979

ABSTRACT

A search for an estimator of β in the Normal Linear Model which has better mean squared error properties than the usual least squares estimator is undertaken. The properties of some classical techniques such as restricted least squares, which includes the selection of a subset of the independent variables, are examined, along with more recent techniques such as ridge regression and Bayesian estimators. Most of these can be shown analytically to improve over least squares only when the true parameter vector β is in some subspace of the parameter space. Empirical Bayes estimators are in general difficult to handle analytically, and so several of these are studied by Monte Carlo methods. A particular modification of one of these empirical Bayes estimators is found to improve over least squares over a large region of the parameter space, and it's use is demonstrated on a small data set. Some suggestions for further improvement of this estimator are given and some techniques for further study of estimators by Monte Carlo methods are recommended.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Richard Brook for his guidance and encouragement during the preparation of this thesis, and also my wife Carol, for bearing with me over the last few months.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

1. INTRODUCTION

This thesis is concerned with estimators of β in the Normal Linear Model,

$$y = X\beta + \xi$$

where y is an nx1 vector of observed variables, X is an nxp matrix of known constants, β is an unobservable px1 vector of coefficients and ξ is an nx1 vector of unobservable random errors assumed to be independently and identically Normally distributed with zero mean and constant, but generally unknown variance σ^2. This is written

$$\xi \sim N(0, \sigma^2 I).$$

The more general case where $\xi \sim N(0, \sigma^2 V)$ will not be explicitly discussed here since, if V is of full rank, the model can be reparameterised to conform to the simpler Normal Linear Model.

It is well known that among unbiased estimators of β, the least squares estimate

$$b_0 = (X^T X)^{-1} X^T y$$

has minimum variance. However this does not guarantee that the variance of the least squares estimate will be small and it is for this reason that some biased estimation techniques are considered here. The criterion adopted as a measure of the goodness of an estimator will be mean squared error , mse or MSE, where for a particular estimator b,

$$\mathrm{mse}(b) = E(b - \beta)^T (b - \beta)$$

and $$\mathrm{MSE}(b) = E(b - \beta)(b - \beta)^T.$$

E has the usual meaning of "the expected value of". It is hoped to find an estimator which has good mean squared error properties when compared with least squares.

Various biased estimators, including some commonly used variants of least squares, are discussed in Chapter 2. The emphasis is on their mean-squared error performance over different regions of the parameter space for β and σ^2. Then in Chapter 3, several simulation experiments are reported. These experiments investigate the mean-squared error performance of two Stein-type estimators and several stochastic ridge estimators over a wide range of the parameters. As the experiments progress a heuristically modified stochastic ridge estimator is developed which appears to have good mean-squared error properties. A numerical example demonstrating the use of this estimator is given in Chapter 4. Finally a summary, including suggestions for further study of stochastic ridge estimators and a discussion of several aspects of simulation studies and mean-squared error performance of estimators is presented in Chapter 5.