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.

EFFICIENT BIASED ESTIMATION

AND

APPLICATIONS

TO

LINEAR MODELS

A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Statistics
at Massey University

Terry Moore

1981

81_13020

Abstract

Efficient Biased Estimation and Applications to Linear Models

In recent years biased estimators have received a great deal of attention because they can often produce more accurate estimates in multiparameter problems. One sense in which biased estimators are often more accurate is that the mean square error is smaller.

In this work several parametric families of estimators are examined and good values of the parameters are sought by approximate analytical arguments. These parametric values are then tested by computing and plotting graphs of the mean square error. In this way the risks of various estimators may be seen and it is possible to discard some estimators which have large risk.

The risk functions are computed by numerical integration - a method faster and more accurate than the usual simulation studies. The advantage of this is that it is possible to evaluate a greater number of estimators; however, the method only copes with spherically symmetric estimators.

The relationship of biased estimation to the use of prior information is made clear. This leads to discussion of partially spherically symmetric estimators and the fact that, although not uniformly better than spherically symmetric ones, they are usually better in a practical sense.

It is shown how the theoretical results may be applied to the linear model. The linear model is discussed in the very general case in which it is not of full rank and there are linear restrictions on the parameter. A kind of weak prior knowledge which is often assumed for such a model makes the partially symmetric estimators attractive.

Distributions of spherically symmetric estimators are briefly discussed.

Acknowledgements

I should like to thank Dr. Richard J. Brook and Prof. Brian Hayman of the department of Mathematics and Statistics, Dr. Brook for supervision of this project and Prof. Hayman for acting as second supervisor as well as giving more substantial help during Dr. Brook's sabbatical leave. I should also like to thank several members of the department with whom I had profitable discussions which helped clarify my thinking. These include Dr. Howard Edwards, Dr. mike Carter, Dr. Rob Pringle, Mr. John Reynolds, Mr. Adrian Swift and Dr. Peter Thomson. My thanks also go to Mrs. Gail Tyson for typing a small but significant part of the manuscript and to the long suffering Dr. Ted Drawneek of the computer unit who had to considerably modify his plotter package to cope with my requirements.

Preface

In recent years it has become apparent that biased estimators often give estimates which are more accurate than unbiased estimators. One way of measuring the accuracy of an estimator is by means of its mean square error. Stein was the first to show that the usual unbiased estimator for the mean of a multivariate normal distribution is inadmissible in the sense that there are estimators with smaller mean square error. In fact the mean square error of the James-Stein estimator which shrinks the usual estimates towards the origin, is often very much smaller. Of course, an estimator which is not unbiased is biased. This seems to be a bad property of an estimator - but so, it would seem, is the property of being inadmissible. In fact both words are technical terms and should not be thought of as having their everyday meanings. There are a variety of ways of measuring the bias of an estimator and a variety of ways of measuring its mean deviation from the true value. The properties of the estimator depend critically on how these things are measured.

In chapter 1 we review some of the properties of estimators and suggest Bayes and empirical Bayes estimators as tools for finding estimators with good properties with respect to repeated sampling. Some of the ways of doing so are surveyed and the results suggest the form which good estimators might take. These estimators shrink the usual estimates towards the origin as does the James-Stein estimator. There is little in this chapter which is new.

Chapter 2 leans heavily on the work of Stein and in particular we prove a result which Stein only proves asymptotically.

It is well known that a linear model can be transformed into the canonical form for which Stein proved his results. We show how to apply the James-Stein estimator directly to the general linear model whether or not of full rank and with or without linear restrictions imposed upon it. We then prove the result alluded to above which shows that separate shrinkages in several linear subspaces of the parameter space are generally better than one over-all shrinkage. The result also gives a bound on the loss of mean square error which may be incurred by such separate shrinkages. Graphs of the difference in risk for shrinkages in two subspaces and the risk for a single subspace shrinkage are plotted in three dimensions together with a contour map showing the region of improvement.

Chapter 3 is closely related to work of Lindley and Smith and to work of Tiao and Zellner. The results are again given for the non-full rank model with linear restrictions. This generalisation poses difficulties when stages of prior information are incorporated in a natural order. It is in this part of chapter 3 that the novelty lies.

Another approach to estimation, Theil's so called "minimum mean square error estimation", is the topic of chapter 4. This criterion does not lead to an estimator as the statistic calculated depends upon unknown parameters. How this statistic itself can be estimated, and the properties of the estimators thus obtained, are discussed. Some distributional properties of quadratic forms and their ratios are derived in a discussion of consistent estimation. The resulting estimators belong to a parametric family of estimators. The various approaches to estimation of the shrinkage factor suggest possible parameter values which are then tested by numerical computation of the risk function. Graphs of these are plotted and displayed at the end of chapter 5. This material is mostly the creation of the author.

Chapter 5 discusses iterative improvement of the estimators of chapter 4. Although this was originally discussed by Hemmerle, we consider several different and novel approaches and compute and plot the risk functions of the resulting estimators. Graphs of the risk functions of these estimators are plotted together with the graphs of the estimators of chapter 4.

The theoretical computation of the risk functions for shrunken estimators was postponed until chapter 6 so that it could first be seen for what class of estimators this should be done. A wide selection of different formulae for the risk are given with the proofs arranged in a systematic manner. If only a few of the formulae are required then the proofs can be simplified by ignoring certain previous results used for computing other forms of the risk. If this is done then more elegant proofs than those used given are obtained. Some generalisations to non-spherically symmetric estimators are given and these are new. These expressions lead to an easier proof of a minimaxity condition than that given by Strawderman in a generalisation of a theorem of Baranchik, and a non-minimaxity theorem of Efron and Morris is generalised to the non-spherically symmetric case.

In chapter 7 some risk estimate domination results of Efron and Morris are generalised by using an unbiased estimator for the risk in the manner of Efron and Morris. This generalisation is not completely successful but some results are obtained.

The distributions of James-Stein and other shrunken estimators have never been given. Possibly of more use is the distribution of the Studentised version. In chapter 8 this is shown to be a transformation of a multivariate t-distribution. Some of the results in chapter 4 on ratios of quadratic forms will lead, with tedious computations, to moments of the James-Stein estimator but this was not done as the Studentised version is of more value.

We have not given a complete bibliography of work in the general area covered by this work, nor have we referred to every paper in the more precise areas in this thesis. The works cited are directly related to the development of this work.

In order to make this work as self contained as possible we have given some standard results along with their proofs and have appended some general mathematical formulae which have been used heavily.

Equations and theorems have been numbered consecutively within each section and are referred to in that section by their numbers. When referenced outside their own section their numbers are prefixed by the chapter and section number. Diagrams are numbered consecutively throughout the whole thesis.

Contents