

2007 ABSTRACTS

STATISTICS AT LA TROBE

No. 2007-1

Large Sample Confidence Intervals in Regression Utilising Prior Information

Paul Kabaila and Khageswor Giri

La Trobe University

Abstract

We consider a linear regression model with regression parameter $\beta = (\beta_1, \beta_2)$ and independent and identically $N(0, \sigma^2)$ distributed errors. For simplicity, assume that σ^2 is known. We prove rigorously (by a rather delicate argument) that σ^2 known corresponds to σ^2 unknown when the sample size is large. Suppose that the parameter of interest is β_1 . Also suppose that we have prior information (which is not a certainty) that β_2/σ is close to 0. We present a method of computing a $1 - \alpha$ confidence interval for β_1 , with endpoints that are smooth functions of the data, that has the following properties. This interval has minimum weighted expected length where the largest weight is given to this expected length when $\beta_2/\sigma = 0$. Also, this interval coincides with the standard $1 - \alpha$ confidence interval when the data happens to strongly contradict this prior information. The properties of the resulting confidence interval are analysed computationally in detail.

No. 2007-2

Sensitivity of Principal Hessian Direction Analysis

Luke Prendergast and Jodie A. Smith

La Trobe University

Abstract

We provide sensitivity comparisons for two competing versions of the dimension reduction method principal Hessian directions (pHd). These comparisons consider the effects of small perturbations on the estimation of the dimension reduction subspace via the influence function. We show that the two versions of pHd can behave completely differently in the presence of certain observational types. Our results also provide evidence that outliers in the traditional sense may or may not be highly influential in practice. Since influential observations may lurk within otherwise typical data, we consider the influence function in the empirical setting for the efficient detection of influential observations in practice.

Trimming influential observations for improved single-index model sufficient summary plots

Luke A. Prendergast

Department of Mathematics and Statistics

La Trobe University

Abstract

The detection of influential observations has been well covered with respect to ordinary least squares (OLS) under an assumed multiple linear regression (MLR) model. It is common for researchers to use diagnostics such as Cook's Distance to detect, and perhaps subsequently treat, problematic observations that have large influence. However, the usefulness of OLS extends beyond the MLR model and the method may be utilized as a dimension reduction tool. A single-index model regression analysis seeks a single dimensional transformation of the multi-dimensional predictor that effectively summarizes the necessary regression structure. However, existing influence diagnostics developed with the MLR model should not be used in this setting where influential observations are termed as such because they have a high influence over the resulting predictor transformation. We introduce an influence diagnostic that may be utilized to detect such observations. Evidence supporting the usefulness of this diagnostic will be provided and the effect of trimming on resulting summary plots are considered with respect to some simulated models.
