INTERNATIONAL CENTRE OF EXCELLENCE FOR EDUCATION IN MATHEMATICS

Comparisons between Sliced Inverse Regression and Sliced Average Variance Estimation

Callum McLean, Department of Statistical Science, La Trobe University

Visualization in regression analysis can become extremely inhibited when faced with high dimensional sets of data. Because of this, methods of dimension reduction are becoming increasingly important. One type of dimension reduction model that describes the relationship between the response y and explanatory variable $\boldsymbol{x}$ is

$$y = f(\beta_1^T \boldsymbol{x}, \ldots, \beta_K^T \boldsymbol{x}, \varepsilon)$$

where the β_i's are unknown column vectors, $\boldsymbol{x}$ is a p-dimensional random vector, the error term ε is independent of $\boldsymbol{x}$ and f is an unknown *link function* with $K < p$. The span of $\beta_1,..,\beta_K$ is known as the effective dimension reduction (e.d.r.) space and any vector element of this space is an e.d.r. direction.

Under some mild conditions Sliced Average Regression (SIR, Li, 1991) and Sliced Average Variance Regression (SAVE, Cook and Weisberg, 1991) are two dimension reduction methods that seek to estimate the e.d.r. space. The proofs for SIR and SAVE were studied as part of the vacation scholarship.

The Boston Housing data collected by Harrison and Rubinfeld (1978) for the estimation of house prices (y) was analysed by Belsley and Welsch (1980) with respect to multiple linear regression. However, plots of y versus the reduced explanatory space (estimated first by SIR and then SAVE) reveal that the model is perhaps of a more complicated form. This clearly highlights the usefulness of dimension reduction techniques as a pre-step to model fitting.

I have greatly enjoyed the AMSI scholarship program and I am continuing study in this area for my honours thesis.

References

BELSLEY, D.A., KUH E., & WELSCH, R.E. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons, Inc.

COOK, R.D. & WEISBERG, S. 1991. Discussion of sliced inverse regression for dimension reduction. *Journal of the American statistical association*, **86**(414), 328-332.

HARRISON, D., & RUBINFELD, D.L. 1978. Hedonic housing prices and the demand for clean air. *J. environ. Econ manage,* **5**, 81-102.

LI, K-C. 1991. *Journal of the American statistical association*, **86**(414), 316-342.

Callum received an ICE-EM Vacation Scholarship in January 2005.
See www.ice-em.org.au/students.html#scholarship.