

Garth Tarr
Mathematical Sciences Institute
Australian National University (ANU)

Visualising variable selection: the mplot package

This talk introduces the mplot R package which provides a collection of functions to aid exploratory variable selection. We have developed routines for modified versions of the simplified adaptive fence procedure (Jiang et al., 2009) as well as other graphical tools such as variable inclusion plots and model selection plots (Müller and Welsh, 2010; Murray et al., 2013). A browser based graphical user interface is provided to facilitate interaction with the results. These variable selection methods rely heavily on bootstrap resampling techniques. Fast performance for standard linear models is achieved using the branch and bound algorithm provided by the leaps package. Reasonable performance for generalised linear models and robust models can be achieved using sensible default tuning parameters and parallel processing. These methods enable users to view the stability of the variable selection process. We will give an overview of what has been achieved and discuss areas for future research.

Time permitting, a brief overview of the current state of parallel processing in R will be given with an application to bootstrap model selection.

This is joint work with Alan Welsh (ANU) and Samuel Müller (USYD).