

Variable Selection with The Modified Buckley–James Method and The Dantzig Selector for High–dimensional Survival Data

Md Hasinur Rahaman Khan*
ISRT, University of Dhaka, Bangladesh hasinur@isrt.ac.bd

J. Ewart H. Shaw
Department of Statistics, University of Warwick, UK Ewart.Shaw@warwick.ac.uk

Abstract

We develop a group of algorithms for variable selection using the accelerated failure time (AFT) models that are based on the synthesis of the Buckley–James estimating method and the Dantzig selector. In particular, first two algorithms are based on two modified Buckley–James estimating methods that are developed for high–dimensional data. The last three algorithms are based on a weighted Dantzig selector that uses weights obtained from the two synthesis based algorithms and another obtained from a proposed form. The adaptive Dantzig selector is chosen because it satisfies the oracle properties. The methods are easy to understand and scalable and they do estimation and variable selection simultaneously. They also can deal with collinearity among the covariates and the groups of covariates. We conducted several simulation studies and one empirical analysis with a microarray dataset.

Keywords: Accelerated failure time, Buckley–James estimating equation, Dantzig selector, Variable selection