Model selection in regression type models\*

Model selection is so fundamental to the practical application of statistics that there are a number of different approaches and a substantial literature on the subject. We focus on resampling methods such as the bootstrap. The purpose of model selection is to choose one or more models from a set of given models with specified desirable properties.

Model selection typically involves specifying an estimator, using the estimator to fit models in the class of possible models and then comparing the fitted models. We broaden the usual approach by considering different types of estimators of each of the models (such as the least squares estimator, various M-estimators, the MM-estimator etc.). Our viewpoint is that the minimal requirements for a useful linear regression model are that the model should (i) parsimoniously describe the relationship between the sample data and the exploratory data and (ii) be able to predict independent new observations. We propose new approaches to the selection of linear, partially linear, and generalized linear regression models and we show that these model selection procedures are consistent under some conditions and work well in our simulations.