RISK CONSISTENCY OF CROSS-VALIDATION
WITH LASSO-TYPE PROCEDURES

Darren Homrighausen and Daniel J. McDonald

Department of Statistics, Colorado State University
Department of Statistics, Indiana University

Abstract: The lasso and related sparsity inducing algorithms have been the target of substantial theoretical and applied research. Correspondingly, many results are known about their behavior for a fixed or optimally chosen tuning parameter specified up to unknown constants. In practice, however, this oracle tuning parameter is inaccessible so one must use the data to select one. Common statistical practice is to use a variant of cross-validation for this task. However, little is known about the theoretical properties of the resulting predictions with such data-dependent methods. We consider the high-dimensional setting with random design wherein the number of predictors $p$ grows with the number of observations $n$. Under typical assumptions