

Bootstrap model selection for possibly dependent and heterogeneous data

Alessio Sancetta

Received: 28 August 2006 / Revised: 20 February 2008 / Published online: 16 July 2008
© The Institute of Statistical Mathematics, Tokyo 2008

Abstract This paper proposes the use of the bootstrap in penalized model selection for possibly dependent heterogeneous data. The results show that we can establish (at least asymptotically) a direct relationship between estimation error and a data based complexity penalization. This requires redefinition of the target function as the sum of the individual expected predicted risks. In this framework, the wild bootstrap and related approaches can be used to estimate the penalty with no need to account for heterogeneous dependent data. The methodology is highlighted by a simulation study whose results are particularly encouraging.

Keywords Complexity regularization · Random penalty · Wild bootstrap