Kernel Averaging Estimators and Parsimonious Model Averaging with a Diverging Number of Parameters

Xinyu Zhang
Academy ofMathematics and Systems Science, Chinese Academy of Sciences

ABSTRACT of "**Kernel Averaging Estimators**"
Bandwidth or smoothness parameter selection has always been a key issue for kernel regression analysis despite the fact that kernel regression has rapidly become a regular tool for explorative empirical research. The issue of bandwidth selection is a fundamental model selection problem stemming from the uncertainty about the smoothness of the regression. In this paper, we advocate a model averaging approach to circumvent the problem caused by this uncertainty. Our new approach involves averaging across a series of Nadaraya-Watson kernel estimators each under different bandwidth, with weights for these different estimators chosen such that a least-squares cross validation criterion is minimised. We prove that the resultant combined-kernel estimator achieves the smallest possible asymptotic aggregate squared error. The superiority of the new estimator over estimators based on widely accepted conventional bandwidth choices in finite samples is demonstrated in a simulation study and a real data example.

ABSTRACT of "**Parsimonious Model Averaging with a Diverging Number of Parameters**"
Model averaging generally provides better predictions than model selection, but the existing model averaging methods cannot lead to parsimonious models. Parsimony is an especially important property when the number of parameters is large. To achieve a parsimonious model averaging coefficient estimator, we suggest a novel criterion for choosing weights. Asymptotic properties are derived in two practical scenarios: (i) one or more correct models in candidate model set; and (ii) all candidate model misspecified. Under the former scenario, it is proved that our method can put the weight one to the smallest correct model and the resulting model averaging estimators of coefficients have many zeros and thus lead to a parsimonious model. The asymptotic distribution of the estimators is also provided. Under the latter scenario, prediction is mainly focused on and we prove that the proposed procedure is asymptotically optimal in the sense that its squared prediction loss and risk are asymptotically identical to those of the best – but infeasible – model averaging estimator. Numerical analysis shows the promise of the proposed procedure over existing model averaging and selection methods.