

Best subset selection via distance covariance

潘文亮 中国科学院数学与系统科学研究院

邀请人：王典朋

摘要： Best subset selection is an important problem in regression analysis, which has many applications in computer science and medicine. However, the existing best subset selection methods have some limitations, such as strict conditions on modeling the relationship or their performances rely on the forms of models. Motivated by these problems, we propose a novel selection procedure to directly identify the best subset of predictors via distance covariance. Based on it, we develop a computational efficient algorithm that can be available to high-dimensional data with guaranteed convergence. We show that the estimator from the proposed algorithm is consistency in the sparsity selection under wild regularity conditions. Simulation and real data analysis show our algorithm drastically improves the computation efficiency across various settings compared with the existing methods.

个人简介： 潘文亮，中国科学院数学与系统科学研究院副研究员，博士生导师。2016年毕业于中山大学统计科学系，获理学博士。曾于2014-2016年在美国北卡罗莱纳大学教堂山分校留学。主要从事统计学习算法、医学图像数据分析、度量空间的非参数方法等领域的研究。目前已在统计学顶级杂志Annals of Statistics, Journal of the American Statistical Association等发表了学术论文二十余篇。主持国家自然科学基金面上与青年基金、广东省纵向协同和博士启动项目，参与国家自然科学基金面上项目两项。

时间：2023年3月25日 11:00-12:00

地点：文萃楼E311