

Residual Importance Weighted Transfer Learning For High-dimensional Linear Regression

赵俊龙 北京师范大学

Transfer learning is an emerging paradigm for leveraging multiple source data to improve the statistical inference on a single target data. In this paper, we propose a novel approach named residual importance weighted transfer learning (RIW-TL) for high-dimensional linear models built on LASSO. Compared to existing methods such as Trans-Lasso that selects source data in an all-in-all-out manner, RIW-TL includes samples via importance weighting. To determine the weights, remarkably RIW-TL only requires one-dimensional density estimation by weighting residuals, thus overcoming the curse of dimensionality of having to estimate high-dimensional densities in naive importance weighting. We show that the oracle RIW-TL provides faster rate than its competitors and develop a cross-fitting procedure to estimate this oracle. We discuss variants of RIW-TL by adopting different choices for residual weighting. The theoretical properties of RIW-TL and its variants are established and compared with those of LASSO and Trans-Lasso. Extensive simulations and a real data analysis confirm the advantages of RIW-TL.

报告人简介：赵俊龙，北京师范大学统计学院教授，主要从事数理统计和机器学习相关研究，包括：高维数据分析、稳健统计、统计机器学习等。部分研究成果发表在统计学国际顶级期刊 *Journal of the Royal Statistical Society: Series B* (JRSSB)、*The Annals of Statistics* (AOS)、*Journal of American Statistical Association*(JASA), *Biometrika* 上。主持参与多项国家自然科学基金项目。