

Random projection pursuit regression for high-dimensional complex data

Sijian Wang

Rutgers University

E-mail: sijian.wang@stat.rutgers.edu

Abstract: Projection pursuit regression (PPR) adapts the additive models in that it first projects the data matrix of explanatory variables in the optimal direction before applying smoothing functions to these explanatory variables. As a consequence of this, PPR can be quite flexible to approximate a complicated regression function. It also has connections to neuronnetwork (deep learning) and boosting. One possible limitation of PPR is the complicated and heavy computation, especially when the number of variables is large. In this talk, borrowing the spirit of random forest, we introduce a random projection pursuit regression, which is not only as flexible as PPR, but also has an advantage in computing complexity. The method is demonstrated with both simulation studies and real data analysis. The connections to neuronnetwork and boosting will be also discussed.