

On L0-regularization in high-dimensional regression

Xiaotong Shen, Wei Pan, YunZhang Zhu and Hui Zhou

University of Minnesota

School of Statistics

Division of Biostatistics

`xshen@stat.umn.edu`

Abstract

ABSTRACT HERE - Feature (variable) selection is crucial for high-dimensional regression, which enables to enhance a model's interpretability and accuracy of its parameter estimation/prediction. In this talk, I will present some sharp asymptotic results of the L0-regularization with regard to consistent feature selection, the optimal rate of convergence of parameter estimation/prediction, and unbiased inference, particularly when the number of candidate features greatly exceeds the sample size. To overcome computational difficulty with the L0-penalty for high-dimensional data, we construct a piecewise linear continuous penalty as its surrogate for feature selection. Theoretically, we show that this surrogate simultaneously possesses the “oracle” properties of the L0-penalty. Computationally, we develop a method through difference convex programming and regular subdifferentials to treat non-convex minimization. As suggested by our theoretical and numerical analyses, the proposed penalty advances three disparate objectives—accurate feature selection, accurate parameter estimation/prediction, and unbiased inference.