Ann Inst Stat Math (2013) 65:807–832
DOI 10.1007/s10463-012-0396-3

On constrained and regularized high-dimensional regression

Xiaotong Shen · Wei Pan · Yunzhang Zhu · Hui Zhou

Received: 2 July 2012 / Revised: 12 October 2012 / Published online: 12 January 2013

Abstract High-dimensional feature selection has become increasingly crucial for seeking parsimonious models in estimation. For selection consistency, we derive one necessary and sufficient condition formulated on the notion of degree of separation. The minimal degree of separation is necessary for any method to be selection consistent. At a level slightly higher than the minimal degree of separation, selection consistency is achieved by a constrained L_0-method and its computational surrogate—the constrained truncated L_1-method. This permits up to exponentially many features in the sample size. In other words, these methods are optimal in feature selection against any selection method. In contrast, their regularization counterparts—the L_0-regularization and truncated L_1-regularization methods enable so under slightly stronger assumptions. More importantly, sharper parameter estimation/prediction is realized through such selection, leading to minimax parameter estimation. This, otherwise, is impossible in the absence of a good selection method for high-dimensional analysis.

Keywords Constrained regression · Parameter and nonparametric models · Nonconvex regularization · Difference convex programming · (p, n) versus fixed p-asymptotics