

Sparse Hierarchical Interactions

Jacob Bien
Stanford University

Abstract

Building predictive interaction models is a challenging problem, especially when the number of variables is large. Statisticians commonly demand that an interaction only be included in a model if both variables are marginally important. We study the problem of identifying hierarchical two-way interaction models from the viewpoint of the Lasso (i.e., ℓ_1 -penalized regression). We show that by adding a set of convex constraints to the Lasso problem, we can produce sparse interaction models that honor the hierarchy restriction. In contrast to stepwise procedures that are most commonly used for building interaction models, our formulation is convex, and its solution is completely characterized by a set of optimality conditions. This makes it easier to study as a statistical estimator. We argue that restricting to hierarchical interactions can be advantageous both statistically and computationally. Our proposal extends more generally to any problem in which "hierarchical sparsity" is desired (i.e., one parameter is forced to be zero if another is zero). For example, in (univariate) polynomial regression, hierarchical sparsity yields low-order polynomial fits. Underlying our work is the observation that there is more to interpretability and simplicity in a model than sparsity alone.

Monday, February 6, 2012
Sherrerd Hall, Room 101
4:30 PM