

Title:

Regularization Path Algorithms for Detecting Gene Interactions

Author(s):

Mee Young Park and Trevor Hastie

Technical Report number (Dept. of Statistics, Stanford Univ.):

2006-13

Date:

September 2006

Abstract:

In this study, we consider several regularization path algorithms with grouped variable selection for modeling gene-interactions. When fitting with categorical factors, including the genotype measurements, we often define a set of dummy variables that represent a single factor/interaction of factors. Yuan & Lin (2006) proposed the group-Lars and the group-Lasso methods through which these groups of indicators can be selected simultaneously. Here we introduce another version of group-Lars. In addition, we propose a path-following algorithm for the group-Lasso method applied to generalized linear models. We then use all these path algorithms, which select the grouped variables in a smooth way, to identify gene-interactions affecting disease status in an example. We further compare their performance to that of L_2 penalized logistic regression with forward stepwise variable selection discussed in Park & Hastie (2006).