

DISCUSSION: “A SIGNIFICANCE TEST FOR THE LASSO”

BY PETER BÜHLMANN, LUKAS MEIER AND SARA VAN DE GEER

ETH Zürich

We congratulate Richard Lockhart, Jonathan Taylor, Ryan Tibshirani and Robert Tibshirani for a thought provoking and interesting paper on the important topic of hypothesis testing in potentially high-dimensional settings.

1. A short description of the test procedure. We start by presenting the proposed test procedure in a slightly different form than in the paper. Let

$$\hat{\beta}(\lambda) := \arg \min \frac{1}{2} \|y - X\beta\|_2^2 + \lambda \|\beta\|_1$$

be the Lasso estimator with tuning parameter equal to λ . The paper uses the Lasso path $\{\hat{\beta}(\lambda) : \lambda > 0\}$ to construct a test statistic for the significance of certain predictor variables.

For a subset $S \subseteq \{1, \dots, p\}$, let $\hat{\beta}_S(\lambda)$ be the Lasso solution using only the variables in S :

$$\hat{\beta}_S(\lambda) := \arg \min_{\beta_S \in \mathbb{R}^{|S|}} \frac{1}{2} \|y - X_S \beta_S\|_2^2 + \lambda \|\beta_S\|_1.$$

The covariance test is based on the difference

$$T(S, \lambda) := [\|y - X_S \hat{\beta}_S(\lambda)\|_2^2 + \lambda \|\hat{\beta}_S(\lambda)\|_1] / \sigma^2 \\ - [\|y - X \hat{\beta}(\lambda)\|_2^2 + \lambda \|\hat{\beta}(\lambda)\|_1] / \sigma^2.$$

If $T(S, \lambda)$ is large, then the solution using only the values in S does not have a very good fit, and this may support evidence against the hypothesis $H_S : A^* \subseteq S$, where $A^* = \text{support}(\beta^*)$ is the true active set.

Let $\infty =: \hat{\lambda}_0 > \hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots$ be the knots of $\hat{\beta}(\lambda)$. For $k \geq 1$, let $\hat{A}_k := \text{support}(\hat{\beta}(\hat{\lambda}_k))$. We put “hats” on these quantities to stress that they are random variables depending (only) on the data. Thus, $T(\hat{A}_k, \hat{\lambda}_k) = 0$ and by continuity arguments also $T(\hat{A}_{k-1}, \hat{\lambda}_k) = 0$. The authors suggest to use the test statistic

$$T_k := T(\hat{A}_{k-1}, \hat{\lambda}_{k+1})$$

for the hypotheses $H_{\hat{A}_{k-1}}$. They derive the interesting result that under certain conditions, the test statistic has an asymptotic exponential distribution.

Received December 2013.

MSC2010 subject classifications. Primary 62J07; secondary 62J12, 62F25.

Key words and phrases. High-dimensional linear model, multiple hypotheses testing, semiparametric efficiency, sparsity.