Statistics > Methodology

On false discovery rate thresholding for classification under sparsity

Pierre Neuvial (SG), Etienne Roquain (LPMA)

(Submitted on 30 Jun 2011)

We study the properties of false discovery rate (FDR) thresholding, viewed as a classification procedure. The "0"-class (null) is assumed to have a known, symmetric log-concave density while the "1"-class (alternative) is obtained from the "0"-class either by translation (location model) or by scaling (scale model). Furthermore, the "1"-class is assumed to have a small number of elements w.r.t. the "0"-class (sparsity). Non-asymptotic oracle inequalities are derived for the excess risk of FDR thresholding. In a regime where Bayes power is away from 0 and 1, these inequalities lead to explicit rates of convergence of the excess risk to zero. Moreover, these theoretical investigations suggest an explicit choice for the nominal level α_m of FDR thresholding, in function of m. Our oracle inequalities show theoretically that the resulting FDR thresholding adapts to the unknown sparsity regime contained in the data. This property is illustrated with numerical experiments, which show that the proposed choice of α_m is relevant for a practical use.

Subjects: **Methodology (stat.ME)**
Cite as: **arXiv:1106.6147v1 [stat.ME]**

Submission history

From: Etienne Roquain [view email]
[v1] Thu, 30 Jun 2011 08:55:18 GMT (2051kb,D)

Which authors of this paper are endorsers?