Do the magic properties of the Benjamini-Hochberg procedure hold in the context of multiple regression ?

Malgorzata Bogdan, Wroclaw University of Technology

Benjamini and Hochberg procedure [2] is a widely applied tool for multiple testing. It is now known that apart from controlling FDR is has some asymptotic optimality properties in the context of parameter estimation and minimizing the Bayes risk in sparse mixtures ([1],[3],[4]). These known results will be briefly summarized and then we will concentrate on the related problem of identifying „causal" regressors in a high dimensional sparse regression. Inherent limitations with respect to FDR control in the context of multiple regression will be presented, which however still allow to control this parametr under very low sparsity and when the columns of the design matrix are only weakly correlated. A new convex optimization procedure, SLOPE ([5],[6]), will be presented, which uses BH idea and has been empirically shown to control FDR in the context of multiple regression with weakly correlated regressors. We will also present simulation results illustrating very promising predictive properties of SLOPE, which suggest that the asymptotic optimality results of [1] carry over into the context of sparse high dimensional regression.

References

[1] Abramovich F., Benjamini Y., Donoho D. L. and Johnstone I. M. (2006). Adapting to unknown sparsity by controlling the false discovery rate. Ann. Statist. 34: 584--653.

[2] Benjamini Y. and Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Statist. Soc. Ser. B. 57: 289-300.

[3] Bogdan M., Chakrabarti A., Frommlet F. and Ghosh J.K, (2011). Asymptotic Bayes Optimality under sparsity of some multiple testing procedures, Ann. Statist. 39: 1551—1579.

[4] Frommlet F. and Bogdan M. (2013). Some optimality properties of FDR controlling rules under sparsity, Electronic Journal of Statistics, 7: 1328—1368.

[5] Bogdan M., van den Berg E., Su W. and Candes E.J. (2013). Statistical estimation and testing via the ordered L_1 norm, arXiv:1310.1969.

[6] Bogdan M., van den Berg E., Sabatti C., Su W. and Candes E. J. (2014). SLOPE -- Adaptive Variable Selection via Convex Optimization, arXiv:1407.3824.