Electronic Journal of Statistics
Vol. 9 (2015) 1723–1726
ISSN: 1935-7524
DOI: 10.1214/15-EJS998

Comment on "Hypothesis testing by convex optimization"*

Philippe Rigollet†

Department of Mathematics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139-4307, USA
e-mail: rigollet@math.mit.edu

Received February 2015.

With the growing size of problems at hand, convexity has become preponderant in modern statistics. Indeed, convex relaxations of NP-hard problems have been successfully employed in a variety of statistical problems such as classification [2, 16], linear regression [7, 5], matrix estimation [8, 12], graphical models [15, 9] or sparse principal component analysis (PCA) [10, 4]. The paper "Hypothesis testing by convex optimization" by Alexander Goldenshluger, Anatoli Juditsky and Arkadi Nemirovski, hereafter denoted by GJN, brings a new perspective on the role of convexity in a fundamental statistical problem: composite hypothesis testing. The role of this problem is illustrated in the light of several interesting applications in Section 4 of GJN.

One of the key insights in GJN is that there exists a pair of distributions, one in each of the composite hypotheses and on which the statistician should focus her efforts. Indeed, Theorem 2.1(ii) guarantees that any test that is optimal for this simple hypothesis problem is also near optimal for the composite hypothesis problem. Moreover, this pair can be found by solving a convex optimization problem. While convexity does not necessarily imply tractability, the convex problem considered here may become simple to the point that closed solutions exist even though no succinct description of the hypothesis sets may be known. This point is illustrated below.

Unlike the papers cited above, where the original problem to be solved is non-convex, GJN assumes given convex hypotheses (or finite unions of convex hypotheses). Hereafter, we investigate the performance of the proposed test when convexity is artificial and arises as a relaxation of a non-convex problem.

Let us consider two examples that fall under the umbrella of *combinatorial testing problems* [1]. Such problems are defined as follows. Assume that one observes a Gaussian random vector $X \sim \mathcal{N}(\mu, I_n)$ for some $\mu \in \mathbb{R}^n$. Let $\mathsf{p} \in \{0,1\}^n$ be a *sparsity pattern* [17]. Given a class $\mathcal{P} \subset \{0,1\}^n$ of sparsity patterns,

*Main article 10.1214/15-EJS1054.

†Supported in part by NSF grants DMS-1317308, CAREER-DMS-1053987.