Electronic Journal of Statistics
Vol. 10 (2016) 67–70
ISSN: 1935-7524
DOI: 10.1214/15-EJS1081A

Discussion of "Estimating structured high-dimensional covariance and precision matrices: Optimal rates and adaptive estimation"*

Alexandre B. Tsybakov

Centre de Recherche en Economie et Statistique (UMR CNRS 9194), ENSAE
e-mail: alexandre.tsybakov@ensae.fr

Received January 2016.

Statistical inference for large covariance and precision matrices is a novel and interesting topic emerged in the last decade. The paper by Tony Cai, Zhao Ren and Harry Zhou (further referred to as [CRZ]) summarizes the key recent achievements in this rapidly developing area where the authors are among the leading contributors. The focus is on fundamental decision theoretic aspects, namely, on the following questions: (a) what are the best attainable rates of convergence of estimators in a minimax sense on various classes of matrices, and (b) how to construct data-driven adaptive procedures attaining these rates without the knowledge of the parameters of the classes. When the dimension of the covariance matrix is greater than the sample size, accurate estimation is problematic unless some assumptions are imposed on the structure of the matrix. A wealth of such structure assumptions is presented in the paper, most of them having the form of sparsity or approximate sparsity constraints. Sparsity here is understood either as a small number of non-zero entries or of non-zero columns/rows of the matrix, or as a small ℓ_q-norm of columns/rows, or as a low rank of the matrix, or as a combination of these properties.

The questions addressed in the paper have analogs in the classical Gaussian mean (Gaussian sequence) model, which is now extensively studied, cf., e.g., [2]. A key problem there is to construct minimax optimal and adaptive estimators of vectors on the ℓ_q-balls based on observation of the unknown vector in Gaussian noise. A straightforward matrix extension of this classical problem is estimation of a sparse matrix $\Sigma \in \mathbb{R}^{p\times p}$ from the observation

$$Y = \Sigma + \varepsilon W \tag{1}$$

where W is a random noise matrix with i.i.d. standard Gaussian entries and $\varepsilon > 0$ is the noise level that we can set as $\varepsilon = 1/\sqrt{n}$ in order to explore similarities

*Main article 10.1214/15-EJS1081.