Asymptotic Efficiency in High-Dimensional Covariance Estimation

Vladimir Koltchinskii

Abstract

We will discuss recent results on asymptotically efficient estimation of smooth functionals of covariance operator Σ of a mean zero Gaussian random vector X in a separable Hilbert space $\mathbb{H}$ based on n i.i.d. observations of X. This includes such functionals as linear forms of eigenvectors of Σ (principal components) and more general functionals of the form $\langle f(\Sigma), B\rangle$, where f is a smooth function in real line and B is a given operator with nuclear norm bounded by a constant. It is of interest to study the problem of efficient estimation of such functionals in a high-dimensional framework, where the dimension d of space $\mathbb{H}$ is allowed to grow with n and the goal is to determine the degree of smoothness of function f needed for efficient estimation of $\langle f(\Sigma), B\rangle$ depending on the rate of growth of d. The problem could be also studied in the case when the space $\mathbb{H}$ is infinite-dimensional and the complexity of covariance estimation is characterized by so called effective rank $\mathbf{r}(\Sigma)$ of unknown covariance. The development of asymptotically efficient estimators in such problems is based on new bias reduction methods and it rely on a variety of probabilistic and analytic tools including concentration inequalities for sample covariance operators, bounds on the remainder of Taylor expansions for operator functions as well as methods of approximate solution of integral equations on the cone of positively semi-definite self-adjoint operators.

The lectures will be mostly based on paper [8] with some results from papers [2-4] as well as preliminaries on operator theory [1] also discussed.

References

[1] A. Aleksandrov and V. Peller. Operator Lipschitz Functions. 2016, arXiv: 1611.01593.

[2] R. van Handel. Structured random matrices. In: *Convexity and Concentration (Carlen, Madiman, and Werner, eds.)*, IMA Vol. 161, Springer, 2017.

[3] V. Koltchinskii and K. Lounici. Asymptotics and concentration bounds for bilinear forms of spectral projectors of sample covariance. *Ann. Inst. H. Poincaré Probab. Statist.*, 2016, 52, 4, 1976–2013.

[4] V. Koltchinskii and K. Lounici. Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli*, 2017, 23, 1, 110–133.

[5] V. Koltchinskii and K. Lounici. Normal approximation and concentration of spectral projectors of sample covariance. *Annals of Statistics*, 2017, 45, 1, 121–157.

[6] V. Koltchinskii and K. Lounici. New Asymptotic Results in Principal Component Analysis. *Sankhya*, 2017, to appear.

[7] V. Koltchinskii, M. Löffler and R. Nickl. Efficient Estimation of Linear Functionals of Principal Components. 2017, *arXiv:1708.07642*.

[8] V. Koltchinskii. Asymptotically Efficient Estimation of Smooth Functionals of Covariance Operators. 2017, *arXiv:1710.09072*

School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332-0160
vlad@math.gatech.edu