Journal of Multivariate Analysis 140 (2015) 139–161

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Robust spiked random matrices and a robust G-MUSIC estimator☆

Romain Couillet

Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506), CentraleSupelec - CNRS - Université Paris-Sud, 91192, Gif sur Yvette, France

ARTICLE INFO

Article history:
Received 30 April 2014
Available online 21 May 2015

AMS subject classifications:
15A52
62G35
62J10

Keywords:
Random matrix theory
Robust estimation
Spiked models
MUSIC

ABSTRACT

A class of robust estimators of scatter applied to information-plus-impulsive noise samples is studied, where the sample information matrix is assumed of low rank; this generalizes the study (Couillet et al., 2013) (restricted to a noise only setting) to spiked random matrix models. It is precisely shown that, as opposed to sample covariance matrices which may have asymptotically unbounded (eigen-)spectrum due to the sample impulsiveness, the robust estimator of scatter has bounded spectrum and may contain isolated eigenvalues which we fully characterize. We show that, if found beyond a certain detectability threshold, these eigenvalues allow one to perform statistical inference on the eigenvalues and eigenvectors of the information matrix. We use this result to derive new eigenvalue and eigenvector estimation procedures, which we apply in practice to the popular array processing problem of angle of arrival estimation. This gives birth to an improved algorithm based on the MUSIC method, which we refer to as robust G-MUSIC.

1. Introduction

The mathematical advances in the field of random matrix theory have recently allowed for the improvement of sometimes old statistical estimation methods when the data have population size N is commensurable with the sample size n, therefore disrupting from the traditional assumption $n \gg N$. One of the recent contributions of random matrix theory lies in the introduction of methods to retrieve information contained in low rank perturbations of large matrices with independent entries, which are referred to as spiked models. The initial study of such models [4] for matrices of the type $\hat{S}_N = \frac{1}{n}(I_N + A)XX^*(I_N + A^*)$, where $X \in \mathbb{C}^{N\times n}$ has independent and identically distributed (i.i.d.) zero mean, unit variance, and finite fourth moment entries and A has fixed rank L, has shown that, as $N, n \to \infty$ with $N/n \to c \in (0, \infty)$, $\hat{S}_N$ may exhibit up to L isolated eigenvalues strictly away from the *bounded* support of the limiting empirical distribution μ of $\hat{S}_N^\circ = \frac{1}{n}XX^*$, while the other eigenvalues of $\hat{S}_N$ get densely compacted in the support of μ. This result has triggered multiple works on various low rank perturbation models for Gram, Wigner, or general square random matrices [7,29,5] with similar conclusions. Of particular interest to us here is the information-plus-noise model $\hat{S}_N = \frac{1}{n}(X + A)(X + A)^*$ introduced in [7] which is closer to our present model. Other generalizations explored the direction of turning X into the more general $XT^{\frac{1}{2}}$ model for $T = \mathrm{diag}(\tau_1, \ldots, \tau_n) \succeq 0$, such that $\frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_{\tau_i} \to \nu$ weakly (with $\boldsymbol{\delta}_x$ the Dirac mass at x), where ν has bounded support $\mathrm{Supp}(\nu)$ and $\max_i\{\mathrm{dist}(\tau_i, \mathrm{Supp}(\nu))\} \to 0$ [8]. In this scenario again, thanks to the fundamental assumption that no τ_i can escape $\mathrm{Supp}(\nu)$ asymptotically, only finitely many eigenvalues of $\hat{S}_N$ can be found away from the support of the limiting spectral distribution of $\frac{1}{n}XTX^*$, and these eigenvalues are intimately linked to A.

☆ This work is jointly supported by the French ANR DIONISOS project (ANR-12-MONU-OOO3) and the GDR ISIS–GRETSI "Jeunes Chercheurs" Project.
E-mail address: romain.couillet@centralesupelec.fr.

http://dx.doi.org/10.1016/j.jmva.2015.05.009
0047-259X/

The major interest of the spiked models in practice is twofold. First, if the (non observable) perturbation matrix A constitutes the relevant information to the system observer, then the observable isolated eigenvalues and associated eigenvectors of $\hat{S}_N$ contain most of the information about A. These isolated eigenvalues and eigenvectors are therefore important objects to characterize. Moreover, since $\hat{S}_N$ has the same limiting spectrum as that of simple random matrix models, this characterization is usually quite easy and leads to tractable expressions and computationally efficient algorithms. This led to notable contributions to statistical inference and in particular to detection and estimation techniques for signal processing [24,27,17,11].

However, from the discussion of the first paragraph, these works have a few severe practical limitations in that: (i) the support of the limiting spectral distribution of $\hat{S}_N$ must be bounded for isolated eigenvalues to be detectable and exploitable and (ii) no eigenvalue of $\hat{S}_N^\circ$ (the unperturbed model) can be isolated, to avoid risking a confusion between isolated eigenvalues of $\hat{S}_N$ arising from A and isolated eigenvalues of $\hat{S}_N$ intrinsically linked to $\hat{S}_N^\circ$. This therefore rules out the possibility to straightforwardly extend these techniques in practice to impulsive noise models $XT^{\frac{1}{2}}$ where $T = \operatorname{diag}(\tau_1, \ldots, \tau_n)$ with either τ_i i.i.d. arising from a distribution with unbounded support or $\tau_i = 1$ for all but a few indices i. In the former case, the support of the limiting spectrum of $\hat{S}_N^\circ$ is unbounded [12, Proposition 3.4], therefore precluding information detection, while in the latter spurious eigenvalues in the spectrum of $\hat{S}_N$ may arise that are also found in $\hat{S}_N^\circ$ and therefore constitute false information (note that this case can be seen as one where low rank perturbations are present *both in the population and in the sample directions* which cannot be discriminated). Such impulsive models are nonetheless fundamental in many applications such as statistical finance or radar array processing, where impulsive samples are classically met.

Traditional statistical techniques to accommodate for impulsive samples fall in the realm of robust estimation [22], the study of which has long remained limited to the assumption $n \gg N$. Recently though, in a series of articles [15,14,13], the author of the present work and his coauthors provided a random matrix analysis of robust estimation, i.e., assuming $N, n \to \infty$ and $N/n \to c \in (0, 1)$, which revealed that a large family of so-called robust M-estimates $\hat{C}_N^\circ$ of scatter (or covariance) matrices can be fairly easily analyzed through simpler equivalent random matrix models. In [14], a noise-only setting of the present article is considered, i.e., with $A = 0$, for which it is precisely shown that robust estimators of scatter can be assimilated as special models of the type of $\hat{S}_N^\circ$.[1] Besides, it importantly appears that the limiting spectrum distribution of $\hat{C}_N^\circ$ always has bounded support, irrespective of the impulsiveness of the samples. Also, it is proved (although not mentioned explicitly) that, asymptotically, isolated eigenvalues of $\hat{C}_N^\circ$ (arising from isolated τ_i) can be found but that none of the eigenvalues can exceed a fixed finite value.

In the present work, we extend the model studied in [14] by introducing a finite rank perturbation A to the robust estimator of scale $\hat{C}_N^\circ$, the resulting matrix being denoted $\hat{C}_N$. As opposed to non-robust models, it shall appear (quite surprisingly on the onset) that $\hat{C}_N$ now allows for finitely many isolated eigenvalues to appear beyond the aforementioned fixed finite value (referred from now on to as the detection threshold), these eigenvalues being related to A. This holds even if $\frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_{\tau_i}$ has unbounded support in the large n regime. As such, any isolated eigenvalue of $\hat{C}_N$ found below the detection threshold may carry information about A or may merely be an outlier due to an isolated τ_i (as in the non-robust context) but any eigenvalue found beyond the detection threshold necessarily carries information about A. This has important consequences in practice as low rank perturbations *in the sample direction* are now appropriately harnessed by the robust estimator while the (more relevant) low rank perturbations *in the population direction* can be properly estimated. We shall introduce an application of these results to array processing by providing two novel estimators for the power and steering direction of signals sources captured by a large sensor array under impulsive noise.

Our contribution thus lies on both theoretical and practical grounds. We first introduce in Theorem 1 the generalization of Couillet et al. [14] to the perturbed model $\hat{C}_N$ which we precisely define in Section 2. This generalization demands to ensure that the (implicit) equation defining the noise only matrix $\hat{C}_N^\circ$ is only marginally affected by the introduction of the finite rank perturbation A, which we prove. The main results are then contained in Section 3. In this section, Theorem 2 provides the localization of the eigenvalues of $\hat{C}_N$ in the large system regime along with associated population eigenvalue and eigenvector estimators when the limiting distribution for $\frac{1}{n}\sum_{i=1}^n \boldsymbol{\delta}_{\tau_i}$ is known. This result is based on standard spiked methods, however for a somewhat new random matrix model which calls for a thorough inspection of the classical proof steps. The result is then extended in Theorem 3 thanks to a two-step estimator where the τ_i are directly estimated. The resulting form of the estimator is quite novel and leaves some degrees of freedom to control the degree of robustness of the method. A practical application to the context of steering angle estimation for array processing is then provided, leading to an improved algorithm referred to as robust G-MUSIC. This application is quite appropriate to radar array processing settings where elliptical assumptions are generally considered to model the impulsive noise clutter [10,1]. Simulation results in this context are then displayed that confirm the improved performance of using robust schemes versus traditional sample covariance matrix-based techniques. To the best of the author's knowledge, the robust G-MUSIC algorithm is the first array processing technique binding together robust estimation theory and the random matrix regime. Besides, this approach is theoretically efficient in impulsive noise scenarios where standard all G-MUSIC approaches known so far provably fail in principle. We finally close the article with concluding remarks in Section 4.

[1] These models are special in that XTX^* becomes now $XTVX^*$ for a diagonal matrix V which makes VT bounded in norm. However, V contains non-observable information about T, which makes $\hat{S}_N^\circ$ only observable through its approximation by $\hat{C}_N^\circ$.