R-NL: Covariance Matrix Estimation for Elliptical Distributions Based on Nonlinear Shrinkage

Simon Hediger, Jeffrey Näf, and Michael Wolf

***Abstract*—We combine Tyler's robust estimator of the dispersion matrix with nonlinear shrinkage. This approach delivers a simple and fast estimator of the dispersion matrix in elliptical models that is robust against both heavy tails and high dimensions. We prove convergence of the iterative part of our algorithm and demonstrate the favorable performance of the estimator in a wide range of simulation scenarios. Finally, an empirical application demonstrates its state-of-the-art performance on real data.**

***Index Terms*—Heavy tails, nonlinear shrinkage, portfolio optimization.**

I. Introduction

MANY statistical applications rely on covariance matrix estimation. Two common challenges are (1) the presence of heavy tails and (2) the high-dimensional nature of the data. Both problems lead to suboptimal performance or even inconsistency of the usual sample covariance estimator $\hat{\mathbf{S}}$. Consequently, there is a vast literature on addressing these problems.

Two prominent ways to address (1) are (Maronna's) M-estimators of scatter [1], as well as truncation of the sample covariance matrix; for example, see [2]. There also appear to be two main approaches to solving problem (2). The first is to assume a specific structure on the covariance matrix to reduce the number of parameters. One example of this is the "spiked covariance model", as explored e.g., in [3], [4], [5], a second is to assume (approximate) sparsity and to use thresholding estimators [6], [7], [8], [9]. We also refer to [2] who present a range of general estimators under heavy tails and extend to the case $n > p$, by assuming specific structures on the covariance matrix. If one is not willing to assume such structure, a second approach is to leave the eigenvectors of the sample covariance

One promising line of research to address both problems at once is to extend (Maronna's) M-estimators of scatter [1] with a form of shrinkage for high dimensions. This approach is in particular popular with a specific example of M-estimators called "Tyler's estimator" [17], which is derived in the context of elliptical distributions. Several papers have studied this approach, using a convex combination of the base estimator and a target matrix, usually the (scaled) identity matrix. We generally refer to such approaches as robust linear shrinkage estimators. For instance, [18], [19], [20], [21] combine the linear shrinkage with Maronna's M-estimators, whereas [22], [23], [24], [25] do so with Tyler's estimator. Since this approach of combining linear shrinkage with a robust estimator entails choosing a hyperparameter determining the amount of shrinkage, the second step often consists of deriving some (asymptotically) optimal parameter that then can be estimated from data. The approach results in estimation methods that are generally computationally inexpensive and it also enables strong theoretical results on the convergence of the underlying iterative algorithms.

Despite these advantages, several problems remain. First, the performance of these robust methods sometimes does not exceed the performance of the basic linear shrinkage estimator of [12] in heavy-tailed models, except for small sample sizes n (say $n < 100$). In fact, the theoretical analysis of [19], [26] shows that robust M-estimators using linear shrinkage are asymptotically equivalent to scaled versions of the linear shrinkage estimator of [12]. Depending on how the data-adaptive hyperparameter is chosen, the performance can even deteriorate quickly as the tails get lighter, as we demonstrate in our simulation study in Section IV. Second, some robust methods cannot handle the case when the dimension p is larger than the sample size n, such