

1 Introduction

Estimation the covariance or correlation matrix of the variables is a common practice for researchers interested in a broad spectrum of statistical applications, ranging from understanding the relationship among variables, perform classification or regression and even form groups or clusters or features. The most common choice of an estimator is the sample covariance or correlation matrix which is also the Maximum Likelihood estimate. While this estimator works fine for $n > p$ cases, it has extremely high approximation error with respect to the population covariance/correlation matrix due to its low rank structure when $n \ll p$.

In 2003, Ledoit and Wolf proposed an estimator that is well conditioned and has much lesser approximation error than the sample correlation matrix [?] [?] in particular under $n \ll p$ scenarios. This approach was further developed and generalized by Schäfer and Strimmer, who besides proposing new shrinkage estimators, also provided analytic calculation of the optimal shrinkage intensity [?]. The idea was to fit a convex combination of the empirical sample covariance matrix (S) along with a chosen target matrix T, which can be chosen to be an identity matrix or constant correlation matrix. The mixing proportion δ in the convex combination $\delta T + (1 - \delta)S$ is usually selected to minimize the expected error of approximation of the shrunk estimate. The above papers used a single target for shrinking, but a multi-target covariance shrinkage approach was recently proposed - see Lancewicki and Aladjem 2014 [?].

In this paper, we propose three versions of an alternative method called *CorShrink* which assumes multiple targets T_1, T_2, \dots, T_K , all of which are noisy versions of the identity matrix and the noise variation increases with each k . We adaptively determine the amount of shrinkage by optimally determining the shrinkage weights for each target and assuming that the set of targets cover the range of variation of the data well. We will discuss about the noise structure and the model fit in more details in the Methods section. We also perform comparisons of our model performance with respect to the Schäfer and Strimmer approach and the Graphical LASSO algorithm developed by Friedman et al [?] for sparse representation of the correlation and primarily inverse correlation matrices used for building causal networks. We show that *CorShrink* performs marginally better than the Schäfer and Strimmer shrinkage in terms of eigenspace approximation to the population covariance when $n \ll p$, and both *CorShrink* and Shafer-Strimmer method perform much better as correlation shrinkage methods compared to GLASSO. We also show an application our method on a single cell mouse pre-implantation RNA-seq data due to Deng et al. 2014 [?].