Joint penalization of multiple scatter matrices

D.E. Tyler[1*], E. Ollila [2], I. Soloveychik[3], and A. Wiesel[3]

[1] *Department of Statistics & Biostatistics, Rutgers, The State University of New Jersey, USA*
[2] *Department of Signal Processing and Acoustics, Aalto University, Finland*
[3] *School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel*
[*] *Presenting author*

Keywords. *Geodesic convexity; Regularized M-estimators of covariance matrices.*

Consider sampling p-dimensional observations from k distinct groups, with sample sizes $n_j, j = 1, \ldots k$ respectively. A common assumption in the k-group problem is the equality of the covariance matrix over the different groups. Such an assumption is helpful in the under-sampled scenario, that is when the sample sizes of the different groups are relatively small, and in particular when $n_j < p$ or even $n_j = 1$ for some groups. Rather than assume equal covariance matrices, we consider in this paper estimating the covariance or scatter matrices under the assumption that they may be simply close to each other in some metric, and hence deviate from some common positive definite "center".

We consider two penalized M-estimation approaches. The first approach begins with a pooled M-estimator of scatter based on all the data, followed by a penalized M-estimator of scatter for each group, with the penalty term chosen so that they groups scatter matrices are shrunk towards the pooled scatter matrix. In the second approach, we minimize the sum of the M-estimation cost functions over the groups along with an additive joint penalty enforcing some similarity, i.e. with shrinkage towards a mutual center.

In both approaches, we utilize the concept of geodesic convexity to prove the existence and uniqueness of the penalized solution under general conditions. We then consider three specific penalty functions based on the Euclidean, the Riemannian, and the information theoretic (Kullback-Leibler) distances In the second approach, the distance based penalties are shown to lead estimators of the mutual center that are related to the arithmetic, the Riemannian and the harmonic means of positive definite matrices, respectively. We also consider a penalty based on an ellipticity measure for positive definite matrices, which shrinks the individual estimators toward a common shape matrix rather than a common scatter matrix.