Covariance Estimation on Manifold

December 10, 2020

The disadvantage of maximizing the likelihood

Estimation in both settings often involves solving an optimization problem, usually maximizing the likelihood of a data set. Defining this objective function requires prescribing a specific probability distribution for the data, which may not be readily available. Moreover, maximum likelihood is not a proper notion of distance on the manifold of symmetric positive-(semi) definite matrices.

Basic tools

Covariance function and family: A one-parameter covariance function is a map $\phi : \mathbb{R} \to S_+(n)$; its corresponding covariance family is the image of ϕ.

Spectral functions: Let A_1 be a matrix in $S_+(n)$. A function $F(A_1)$ is a spectral function if it is a differentiable and symmetric map from the eigenvalues of A_1 to the reals. The function F can be understood as a composition of λ the eigenvalue function and a differentiable and symmetric map f; that is, $F(A_1) = \lambda(A_1)$.

Some spectral functions are introduced below.

- Natural distance in $S_+(n)$:

$$d(A_1, A_2) = \sqrt{\sum_{k=1}^{n} \log^2(\lambda_K)}$$

Problem statement

Let $y_1, y_2, \dots y_q$ be independent and identically distributed observations from some distribution with density function $p_Y(\cdot, A\#_t B)$, the maximum likelihood estimate of t is

$$t^{ML} = \arg \min_{t \in (-\infty, \infty)} \sum_{i=1}^{q} \log p_Y(y_i, A\#_t B)$$

Suppose the sample covariance matrix $\hat{C}$ of $\{y_i\}_{i=1}^{q}$ is full-rank, in this case, we can minimize the geodesic distance $d(\cdot, \cdot)$ defined below

$$t^* \in \arg \min_{t \in (-\infty, \infty)} d(A\#_t B, \hat{C})$$