Geodesically-convex optimization for averaging partially observed covariance matrices - Supplementary material -

Florian Yger* FLORIAN.YGER@DAUPHINE.FR
LAMSADE, CNRS, Université Paris-Dauphine, Université PSL
Paris, France

Sylvain Chevallier* SYLVAIN.CHEVALLIER@UVSQ.FR
Université Paris-Saclay, UVSQ, LISV
Vélizy-Villacoublay, France

Quentin Barthélemy Q.BARTHELEMY@FOXSTREAM.FR
Foxstream
Vaulx-en-Velin, France

Suvrit Sra SUVRIT@MIT.EDU
Laboratory for Information and Decision Systems, Massachusetts Institute of Technology
Cambridge, MA, USA

Abstract

This file contains the supplementary material for the article "Geodesically-convex optimization for averaging partially observed covariance matrices" (Yger et al., 2020).

1. Detailed proof on geodesic convexity of Section 3.4

For SPD matrices $\mathbb{I}_c \preccurlyeq \Sigma_1, \Sigma_2$, we have $\mathbb{I}_c \preccurlyeq (\Sigma_1 \sharp \Sigma_2)$. We consider a full column-rank matrix $\tilde{M} \in \mathbb{R}^{c\times(c-p)}$, such that $\tilde{M}^\top \tilde{M} \succcurlyeq \mathbb{I}_{(c-p)}$. On the one hand, we have:

$$\mathbb{I}_{(c-p)} \preccurlyeq \tilde{M}^\top \tilde{M} = \tilde{M}^\top \mathbb{I}_c \tilde{M} \preccurlyeq \tilde{M}^\top (\Sigma_1 \sharp \Sigma_2) \tilde{M} \ , \tag{1}$$

On the other hand, applying Theorem 2.8 of (Sra and Hosseini, 2015) with $\tilde{M}^\top(\cdot)\tilde{M}$ as a strictly positive map:

$$\tilde{M}^\top (\Sigma_1 \sharp \Sigma_2) \tilde{M} \preccurlyeq (\tilde{M}^\top \Sigma_1 \tilde{M}) \sharp (\tilde{M}^\top \Sigma_2 \tilde{M}) \ . \tag{2}$$

Combining Eq. (1) and (2):

$$\mathbb{I}_{(c-p)} \preccurlyeq \tilde{M}^\top (\Sigma_1 \sharp \Sigma_2) \tilde{M} \preccurlyeq (\tilde{M}^\top \Sigma_1 \tilde{M}) \sharp (\tilde{M}^\top \Sigma_2 \tilde{M}) \ , \tag{3}$$

and because $\|\mathrm{Log}(\cdot)\|_F^2 = \delta_R^2(\mathbb{I}, \cdot)$ is monotonically increasing above $\mathbb{I}$, we have:

$$\delta_R^2 \left(\mathbb{I}_{(c-p)}, \tilde{M}^\top (\Sigma_1 \sharp \Sigma_2) \tilde{M} \right) \leq \delta_R^2 \left(\mathbb{I}_{(c-p)}, (\tilde{M}^\top \Sigma_1 \tilde{M}) \sharp (\tilde{M}^\top \Sigma_2 \tilde{M}) \right) \ . \tag{4}$$

Using continuity, this midpoint convexity is sufficient to prove the geodesic convexity.

2. Dataset visualisation of Section 4.1

Figure 1 illustrates the dataset generated in the experiment described in Section 4.1.

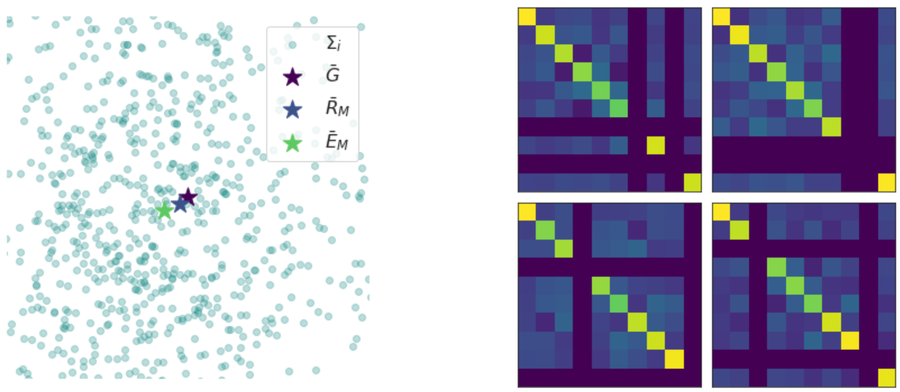

Figure 1: Left: visualization of the generated covariance matrices Σ_i (for $\sigma = 0.25$), projected in the tangent space at $\bar{G}$ and keeping the two highest principal components, with the groundtruth $\bar{G}$, the masked Riemannian mean $\bar{R}_M$ and the masked Euclidean mean $\bar{E}_M$. Right: visualization of some incomplete covariance matrices, for $p = 2$ missing variables.

3. Evaluation of convergence rate and robustness on synthetic dataset

The convergence rate is evaluated on a synthetic dataset: matrices are generated as in Section 4.1 of the article. The training samples are generated from a distribution that is defined by a reference $\bar{G}$ and a dispersion σ. In this experiment, data are generated with $m = 100$ matrices, $c = 30$ variables, and $p = 5$ missing variables. The robustness of the algorithm is evaluated, varying the noise $\sigma = \{0.25, 0.5, 1.0\}$. The solver used to estimate the masked Riemannian mean is a Riemannian conjugate gradient.

On the Fig. 2, the value of the cost f_{R_M} of the masked Riemannian mean (defined in Eq. (10)) is represented as a function of the iteration. The estimation of the masked mean $\bar{\Sigma}$ is repeated 50 times, with different initial points. The cost values f_{R_M} are plotted as green (light gray) lines. The averaged cost per iteration is shown as a strong blue (dark) line.

The Fig. 2 shows that the algorithm converges after 3 to 4 iterations, depending of the dispersion σ of the generated matrices. This experimental validation confirms that the introduced algorithm converges in practice for different levels of noise, and could be applied on various types of data.

References

S Sra and R Hosseini. Conic geometric optimization on the manifold of positive definite matrices. *SIAM J Optim*, 25(1):713–739, 2015.

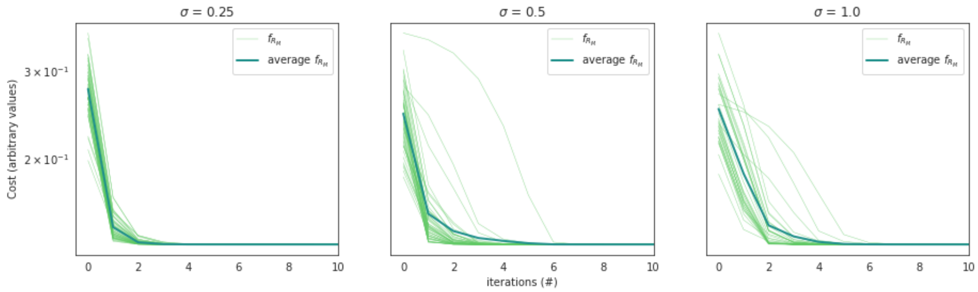

Figure 2: Evaluation of the convergence rate, for different σ values: cost f_{R_M} is represented as a function of the iteration.

F Yger, S Chevallier, Q Barthélemy, and S Sra. Geodesically-convex optimization for averaging partially observed covariance matrices. In *ACML*, pages 417–432, 2020.