MGGD Parameter Estimation on the Space of SPD Matrices

Zois Boukouvalas [1] **Jialun Zhou** [2] **Mark Fuge** [1] **Salem Said** [2]

1. Problem Statement

Due to its simple parametric form, the family of multivariate generalized Gaussian distributions (MGGD) has been widely used for modeling vector-valued signals. Therefore, efficient estimation of its parameters is of significant interest for a number of machine learning tasks. The MGGD probability density functions are given by (Kotz, 1975)

$$p(\mathbf{x};\boldsymbol{\Sigma},\beta,m) = \eta\frac{\beta}{m^{\frac{p}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}}\times\exp\left[-\frac{1}{2m^{\beta}}\left(\mathbf{x}^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{x}\right)^{\beta}\right],$$

where $\mathbf{x}\in\mathbb{R}^p$, $\eta = \frac{\Gamma\left(\frac{p}{2}\right)}{\pi^{\frac{p}{2}}\Gamma\left(\frac{p}{2\beta}\right)2^{\frac{p}{2\beta}}}$, $m>0$ is a scale parameter, $\beta>0$ is a shape parameter that controls the distribution's peakedness and spread, and $\boldsymbol{\Sigma}\in\mathbb{R}^{p\times p}$ is a symmetric positive definite (spd) matrix, called the scatter matrix.

For a random sample $\{\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N\}$ of $p-$dimensional observation vectors, the computation of the maximum likelihood (ML) estimates $\hat{\beta}$, $\hat{\boldsymbol{\Sigma}}$, and $\hat{m}$ lies in solving the nonlinear equation given by

$$\boldsymbol{\Sigma} = \sum_{i=1}^{N}\frac{p}{u_i + u_i^{1-\beta}\sum\limits_{i\neq j}u_j^{\beta}}\mathbf{x}_i\mathbf{x}_i^{\top}, \qquad (1)$$

where $u_i = \mathbf{x}_i^{\top}\boldsymbol{\Sigma}^{-1}\mathbf{x}_i$. The method of moments (MoM) and ML estimation techniques have been proposed (Verdoolaege & Scheunders, 2011; 2012; Bombrun et al., 2012; Sra & Hosseini, 2013; Pascal et al., 2013) for estimating the scatter matrix. However, their accuracy suffers when β becomes large, making them unsuitable for many applications.

Here, we present an effective algorithm on the space of spd matrices $\mathcal{S}_+^p$—Riemannian-averaged fixed point algorithm (RA-FP)—that accurately estimates $\boldsymbol{\Sigma}$ for any β.

2. RA-FP Algorithm

Boukouvalas et al. (2015) formulated (1) as a fixed point equation by defining the right hand side as a function on $\mathcal{S}_+^p$ and used a FP algorithm $\boldsymbol{\Sigma}_{k+1} = f(\boldsymbol{\Sigma}_k)$ for $k=0,1,2,\ldots$ to estimate $\hat{\boldsymbol{\Sigma}}$. The algorithm's convergence requires f be contractive, which numerical experiments showed is not the case when $\beta\geq 2$. By taking advantage of the Riemannian geometry of $\mathcal{S}_+^p$, RA-FP overcomes this difficulty.

Precisely speaking, given $\boldsymbol{\Sigma}_k$, the new estimate $\boldsymbol{\Sigma}_{k+1}$ is:

$$\boldsymbol{\Sigma}_{k+1} = \boldsymbol{\Sigma}_k \#_{t_k} f(\boldsymbol{\Sigma}_k), \qquad (2)$$

where the right hand side of (2) denotes the Riemannian average with ratio t_k between $\boldsymbol{\Sigma}_k$ and $\boldsymbol{\Sigma}_{k+1}$. Thus, RA-FP implements Riemannian averages of successive fixed point iterates, preventing them from diverging when β increases. For a full discussion of RA-FP, as well as it proof of convergence, we refer the reader to Boukouvalas et al. (2015).

3. Numerical Experiments

To numerically verify RA-FP's effectiveness, Fig. 1 shows the Frobenius norm of the difference between the estimated and original scatter matrices, when $\boldsymbol{\Sigma}$ and β are jointly estimated.

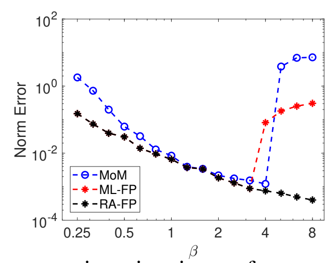

Figure 1. Scatter matrix estimation performance for different values of the shape parameter when $\boldsymbol{\Sigma}$ and β have been jointly estimated. $N = 10000$.

4. Future Directions

Future work will focus on high-dimensional cases, specifically how to estimate MGGD parameters when dimension p increases. Providing globally convergent algorithms that also scale is non-trivial. However, one promising approach is averaged constant-step-size Riemannian stochastic gradient descent—an MCMC method with a geometric mixing property that converges exponentially to the stationary distribution. The online averaging (as in RA-FP) stabilises the Markov chain to a unique deterministic limit, which experimentally approximates the true parameter values well.

[1]Department of Mechanical Engineering, University of Maryland, College Park, Maryland, USA [2]Laboratoire IMS, Universit de Bordeaux, Bordeaux, France. Correspondence to: Zois Boukouvalas <zoisb@umd.edu>.

Copyright 2018 by the author(s).

Acknowledgements

ZB and MF were partially supported through ONR N00014-17-1-2108.

References

Bombrun, L., Pascal, F., Tourneret, J.-Y., and Berthoumieu, Y. Performance of the maximum likelihood estimators for the parameters of multivariate generalized Gaussian distributions. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 3525–3528, March 2012. doi: 10.1109/ICASSP.2012.6288677.

Boukouvalas, Z., Said, S., Bombrun, L., Berthoumieu, Y., and Adal, T. A new riemannian averaged fixed-point algorithm for mggd parameter estimation. *IEEE Signal Processing Letters*, 22(12):2314–2318, Dec 2015. ISSN 1070-9908. doi: 10.1109/LSP.2015.2478803.

Kotz, Samuel. *Multivariate distributions at a cross road*. Springer, 1975.

Ollila, E., Tyler, D.E., Koivunen, V., and Poor, H.V. Complex elliptically symmetric distributions: Survey, new results and applications. *Signal Processing, IEEE Transactions on*, 60(11):5597–5625, Nov 2012. ISSN 1053-587X. doi: 10.1109/TSP.2012.2212433.

Pascal, F., Bombrun, L., Tourneret, J.-Y., and Berthoumieu, Y. Parameter estimation for multivariate generalized Gaussian distributions. *Signal Processing, IEEE Transactions on*, 61(23):5960–5971, Dec 2013. ISSN 1053-587X. doi: 10.1109/TSP.2013.2282909.

Sra, Suvrit and Hosseini, Reshad. Geometric optimisation on positive definite matrices for elliptically contoured distributions. In *Advances in Neural Information Processing Systems*, pp. 2562–2570, 2013.

Verdoolaege, Geert and Scheunders, Paul. Geodesics on the manifold of multivariate generalized Gaussian distributions with an application to multicomponent texture discrimination. *International Journal of Computer Vision*, 95(3):265–286, 2011.

Verdoolaege, Geert and Scheunders, Paul. On the geometry of multivariate generalized Gaussian models. *Journal of Mathematical Imaging and Vision*, 43(3):180–193, 2012.

Zhang, Teng, Wiesel, Ami, and Greco, Maria Sabrina. Multivariate generalized Gaussian distribution: Convexity and graphical models. *Signal Processing, IEEE Transactions on*, 61(16):4141–4148, 2013.