

Score matching for non-Euclidean data

Eduardo García-Portugués • Department of Statistics, UC3M

Score matching is an important method for fitting parametric distribution and regression models without needing to compute their normalizing constant — a common challenge in high-dimensional or complex models and a requirement for applying maximum likelihood estimation. Originally introduced by Aapo Hyvärinen in 2005 (Hyvärinen, 2005), score matching has since evolved, with many generalizations aimed at handling data coming from non-Euclidean supports, such as data comprised by directions on the sphere or compositions on the simplex (Mardia et al., 2016; Scealy et al., 2024; Scealy & Wood, 2023; Xu et al., 2025), where normalizing constants are especially complicated, even in low-dimensional situations, because of the more complex domains of integration.

This thesis will focus on reviewing the main developments in score matching framework for directional and/or compositional data, digesting the mathematical foundations of these contributions, performing numerical studies to compare with the maximum likelihood inference framework and method of moments, and proposing new real data applications.

References

- Hyvärinen, A. (2005). Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research*, 6(Apr), 695–709. <http://jmlr.org/papers/v6/hyvarinen05a.html>
- Mardia, K. V., Kent, J. T., & Laha, A. K. (2016). Score matching estimators for directional distributions. *arXiv:1604.08470*. <https://doi.org/10.48550/arXiv.1604.08470>
- Scealy, J. L., Hingee, K. L., Kent, J. T., & Wood, A. T. A. (2024). Robust score matching for compositional data. *Statistics and Computing*, 34(2), 93. <https://doi.org/10.1007/s11222-024-10412-w>
- Scealy, J. L., & Wood, A. T. A. (2023). Score matching for compositional distributions. *Journal of the American Statistical Association*, 118(543), 1811–1823. <https://doi.org/10.1080/01621459.2021.2016422>
- Xu, J., Scealy, J. L., Wood, A. T. A., & Zou, T. (2025). Generalized score matching. *Journal of Multivariate Analysis*, 210, 105473. <https://doi.org/10.1016/j.jmva.2025.105473>