Score matching vs. maximum likelihood

Eduardo García-Portugués • Department of Statistics, UC3M

Parametric distribution and regression models rely on accurately estimating probability distributions from data. Under certain smoothness assumptions, maximum likelihood estimation is known to be asymptotically efficient, i.e., it achieves the lowest possible variance among all consistent estimators as the sample size increases. However, maximum likelihood estimation often requires computing the normalizing constant of the distribution, which can be computationally challenging in high-dimensional models or even in one-dimensional models. Besides, maximum likelihood estimation often leads to non-explicit estimators that require from numerical optimization, even at simple models like a Gamma distribution, Beta distribution or distributions in the exponential family, and the regression models featuring them.

Score matching is an important method for fitting distribution models, offering a way to estimate parameters without needing to compute the normalizing constant and offering closed-form estimates where maximum likelihood does not. Originally introduced by Aapo Hyvärinen in 2005 (Hyvärinen, 2005) to deal with continuous random vectors on $\mathbb{R}^p$, score matching has been generalized to accommodate more general supports (Hyvärinen, 2007; Liu et al., 2022; Yu et al., 2019).

This thesis will focus on reviewing the foundations of score matching for data on $\mathbb{R}^p$, comparing side-to-side with maximum likelihood estimation in a variety of distribution models. Particularly, the thesis will focus on evaluating the computational and mathematical effort required to implement both methods, and on quantifying the asymptotic efficiency loss of score matching with respect to maximum likelihood estimation, with the aim of building a comprehensive summary table. The study will be done by numerical experiments.

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

Hyvärinen, A. (2007). Some extensions of score matching. *Computational Statistics & Data Analysis*, *51*(5), 2499–2512. https://doi.org/10.1016/j.csda.2006.09.003

Liu, S., Kanamori, T., & Williams, D. J. (2022). Estimating density models with truncation boundaries using score matching. *Journal of Machine Learning Research*, *23*(186), 1–38. http://jmlr.org/papers/v23/21-0218.html

Yu, S., Drton, M., & Shojaie, A. (2019). Generalized score matching for non-negative data. *Journal of Machine Learning Research*, *20*(76), 1–70. http://jmlr.org/papers/v20/18-278.html