Geometry of the Jensen-Shannon metric

1 Introduction

There are several metrics available on the set $\mathcal{P}(\mathbb{R}^d)$ of Borel probability measures on $\mathbb{R}^d$. The performance of several algorithms, useful in data analysis, depend on the geometry of the chosen metric. One specific metric, known as the Jensen-Shannon (JS) metric, plays an important role in the theoretical analysis of Generative Adversarial Networks (GANs) introduced in Goodfellow et al. (2014). The JS metric between probability measures μ and ν is defined by

$$\mathcal{J}(\mu,\nu) = \left[\frac{1}{2}D\left(\mu\|\frac{\mu+\nu}{2}\right) + \frac{1}{2}D\left(\nu\|\frac{\mu+\nu}{2}\right)\right]^{\frac{1}{2}},$$

where $D(.\|.)$ stands for the Kullback-Leibler divergence defined by

$$D(\mu\|\nu) = \begin{cases} \int_{\mathbb{R}^d}\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)\mathrm{d}\nu & \text{if } \mu \ll \nu, \\ +\infty & \text{otherwise.}\end{cases}$$

The proof that $\mathcal{J}$ defines a metric was first given in Endres and Schindelin (2003). Some theoretical properties of GANs, in connection to metric $\mathcal{J}$, were also recently studied in Biau et al. (2018). Next are listed a few research perspectives.

2 Projects

2.1 Geometry of $\mathcal{J}$

A first goal of this project is to further understand the geometric properties of this metric. For instance, one objective would be to study whether the set $\mathcal{P}(\mathbb{R}^d)$, endowed with $\mathcal{J}$, is a geodesic metric space, characterize its shortest paths and eventually its curvature properties.

2.2 Barycenters

A related field of investigation is to study barycenters relative to metric $\mathcal{J}$. Given a probability measure P on $\mathcal{P}(\mathbb{R}^d)$ (i.e. P is a probability measure on the set of probability measures), a barycenter of P for the metric $\mathcal{J}$ is any probability measure $\nu^* \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\int \mathcal{J}(\mu,\nu^*)^2\,\mathrm{d}P(\mu) = \min_{\nu}\int \mathcal{J}(\mu,\nu)^2\,\mathrm{d}P(\mu).$$

One natural question is whether such barycenters exist, are unique and if one can estimate them in a consistent way. These questions could be studied in the light of the recent paper by Le Gouic and Loubes (2017).

2.3 Variance inequalities

Given the recent work by Ahidar-Coutrix et al. (2018), another interesting question is whether one can establish so called variance inequalities for the metric $\mathcal{J}$. More specifically, we would like to known if, a probability measure P on $\mathcal{P}(\mathbb{R}^d)$ with barycenter $\nu^\star$, there exist constants $K > 0$ and $\beta \in (0,1]$ such that for all $\nu \in \mathcal{P}(\mathbb{R}^d)$,

$$\mathcal{J}(\nu,\nu^*)^2 \le K\left(\int (\mathcal{J}(\mu,\nu)^2 - \mathcal{J}(\mu,\nu^*)^2)\,\mathrm{d}P(\mu)\right)^{\beta}.$$

Such an inequality was shown to guarantee fast convergence rates for certain estimators of barycenters in Ahidar-Coutrix et al. (2018) and would be of great interest from a statistical point of view.

References

A. Ahidar-Coutrix, T. Le Gouic, and Q. Paris. On the rate of convergence of empirical barycentres in metric spaces: curvature, convexity and extendible geodesics. *arXiv preprint arXiv:1806.02740*, 2018.

G. Biau, B. Cadre, M. Sangnier, and U. Tanielian. Some theoretical properties of gans. *arXiv:1803.07819*, 2018.

D.M. Endres and J.E. Schindelin. A new metric for probability distributions. *IEEE Transactions on Information Theory*, 49:1858–1860, 2003.

I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'14, pages 2672–2680. MIT Press, 2014.

T. Le Gouic and J.-M. Loubes. Existence and consistency of Wasserstein barycenters. *Probability Theory and Related Fields*, 168:901–917, 2017.