FY2025/2025年度

Functional Analytic Learning Team

Ha Quang Minh

関数解析的学習チーム ハクァンミン

RIKEN Center for Advanced Intelligence Project

Focus: Optimal Transport and Information Geometry in Statistics and Machine Learning

- Our focus is on Optimal Transport and Information Geometry in infinite dimension, in particular in the settings of infinite-dimensional probability measures on Hilbert space and stochastic processes with sample paths lying in infinite-dimensional function spaces. In the setting of Gaussian measures on Hilbert space and Gaussian processes, many quantities admit explicit formulas.
- Our theoretical results provide a rigorous mathematical framework for Gaussian process methods in machine learning and statistics. Examples of recent applications utilizing these results include Functional Bayesian Neural Networks and Linear Gaussian inverse problems on Hilbert space.

Examples of recent results: Geometric Jensen-Shannon divergence

- Rényi divergence of order r, $0 < r < 1$ $D_{\mathrm{R},r}(P||Q) = -\frac{1}{r(1-r)}\log\int_{\mathcal{X}} p^r(x)q^{1-r}(x)d\mu(x)$ where $p = \frac{dP}{d\mu}$, $q = \frac{dQ}{d\mu}$, $P \ll \mu$, $Q \ll \mu$
- $\lim_{r\to 1} D_{\mathrm{R},r}(P||Q) = \mathrm{KL}(P||Q)$ **(Kullback-Leibler divergence)**, $\lim_{r\to 0} D_{\mathrm{R},r}(P||Q) = \mathrm{KL}(Q||P)$, $\mathrm{KL}(P||Q) < \infty$ if and only if $P \ll Q$.
- $\mathrm{KL}(P||Q)$ is asymmetric and one form of symmetrization is the Jensen-Shannon divergence

$$\mathrm{JS}(P||Q) = \frac{1}{2}\mathrm{KL}\left(P\left\|\frac{P+Q}{2}\right.\right) + \frac{1}{2}\mathrm{KL}\left(Q\left\|\frac{P+Q}{2}\right.\right) = \frac{1}{2}\int_{\mathcal{X}} p(x)\log\left(\frac{2p(x)}{p(x)+q(x)}\right)d\mu(x) + \frac{1}{2}\int_{\mathcal{X}} q(x)\log\left(\frac{2q(x)}{p(x)+q(x)}\right)d\mu(x)$$

  where $P \ll \mu$, $Q \ll \mu$, $p = \frac{dP}{d\mu}$, $q = \frac{dQ}{d\mu}$, e.g.$\mu = \frac{P+Q}{2}$. $\mathrm{JS}(P||Q)$ is **always finite**, with $0 \leq \mathrm{JS}(P||Q) \leq \log 2$

- $\frac{P+Q}{2}$ is no longer Gaussian when P and Q are both Gaussian measures and there is **no known closed form formula** for $\mathrm{JS}(P||Q)$.
- For **any pair** of **mutually singular** probability measures P, Q (in particular, two mutually singular Gaussian measures on an infinite-dimensional Hilbert space) $\mathrm{JS}(P||Q) = \log 2$, thus the divergence is finite but degenerate in this case.
- **Weighted abstract means**, e.g. *Arithmetic mean* $A_\alpha(x,y) = (1-\alpha)x + \alpha y$; *Geometric mean* $G_\alpha(x,y) = x^{1-\alpha}y^\alpha$; *Harmonic mean* $H_\alpha(x,y) = \frac{xy}{(1-\alpha)y+\alpha x}$.
- **Interpolation of probability measures via abstract weighted means**. Let $\alpha \in [0,1]$, M_α be a given weighted mean. Let $P, Q \in \mathcal{P}(\mathcal{X})$, μ be a Borel measure on $\mathcal{X}$ such that $P \ll \mu, Q \ll \mu$. Let $p = \frac{dP}{d\mu}$, $q = \frac{dQ}{d\mu}$. The M_α-interpolation $(pq)^M_\alpha$, or α-weighted M-mixture, of p and q is defined to be

$$(pq)^M_\alpha(x) := \frac{M_\alpha(p(x),q(x))}{Z^M_\alpha(p:q)}, \quad \text{where } Z^M_\alpha(p:q) = \int_{\mathcal{X}} M_\alpha(p(x),q(x))\,d\mu(x).$$

  We call $(PQ)^M_\alpha \in \mathcal{P}(\mathcal{X})$with $\frac{d(PQ)^M_\alpha}{d\mu} = (pq)^M_\alpha(x)$ the α-weighted M-mixture of P and Q.

- **Generalized Jensen-Shannon divergence via abstract means**. Let M be a given abstract mean. Let $P, Q \in \mathcal{P}(\mathcal{X},\mu)$. The M-Jensen-Shannon divergence between P and Q is defined to be

$$\mathrm{JS}_{M_\alpha}(P||Q) = (1-\alpha)\mathrm{KL}(P||(PQ)^M_\alpha) + \alpha\mathrm{KL}(Q||(PQ)^M_\alpha), \quad 0 \leq \alpha \leq 1.$$

- **Geometric interpolation between probability measures**. For $M_\alpha = G_\alpha$, $Z^M_\alpha(p:q) = \int_{\mathcal{X}} p^{1-\alpha}(x)q^\alpha(x)\,d\mu(x) = \exp(-\alpha(1-\alpha)D_{R,\alpha}(Q||P)) > 0$ if and only if $D_{R,\alpha}(Q||P) < \infty$. Thus $(PQ)^G_\alpha$ is well-defined, independent of the dominating measure μ, if and only if $D_{R,\alpha}(Q||P) < \infty$

Example of recent results: Geometric Jensen-Shannon divergence between Gaussian measures on Hilbert space

Theorem 1 (Geometric interpolation of two equivalent Gaussian measures on Hilbert space). *Let $\mu_* \in \mathcal{H}$ and $C_* \in \mathrm{Tr}(\mathcal{H})$ be fixed. Let $\mathrm{SymHS}(\mathcal{H}) = \{S \in \mathrm{Sym}(\mathcal{H}) \cap \mathrm{HS}(\mathcal{H}) \ :\ I - S > 0\}$. Let $\mu_i = \mathcal{N}(m_i, C_i) \in \mathrm{Gauss}(\mathcal{H},\mu_*)$, $i = 0,1$, where $C_i = C_*^{1/2}(I-S_i)C_*^{1/2}$, $S_i \in \mathrm{SymHS}(\mathcal{H})_{<I}$. Let $p_i = \frac{d\mu_i}{d\mu_*}$. Then*

$$(p_0p_1)^G_\alpha(x) = \frac{d\mu_\alpha}{d\mu_*}, \quad \alpha \in [0,1], \ \text{where } \mu_\alpha = \mathcal{N}(m_\alpha, C_\alpha) \in \mathrm{Gauss}(\mathcal{H},\mu_*), C_\alpha = C_*^{1/2}(I-S_\alpha)C_*^{1/2}$$
$$S_\alpha = I - [(1-\alpha)(I-S_0)^{-1} + \alpha(I-S_1)^{-1}]^{-1} \in \mathrm{SymHS}(\mathcal{H})_{<I},$$
$$m_\alpha = m_* + C_*^{1/2}(I-S_\alpha)[(1-\alpha)(I-S_0)^{-1}C_*^{-1/2}(m_0-m_*) + \alpha(I-S_1)^{-1}C_*^{-1/2}(m_1-m_*)].$$

Theorem 2 (Geometric Jensen-Shannon divergence between equivalent Gaussian measures). *Let $\mu_i = \mathcal{N}(m_i, C_i) \in \mathrm{Gauss}(\mathcal{H},\mu_*)$, $i = 0,1$, with $C_i = C_*^{1/2}(I-S_i)C_*^{1/2}$, $S_i \in \mathrm{SymHS}(\mathcal{H})_{<I}$. Then we have the closed form expression*

$$\mathrm{JS}_{G_\alpha}(\mu_0||\mu_1) = (1-\alpha)\mathrm{KL}(\mu_0||\mu_\alpha) + \alpha\mathrm{KL}(\mu_1||\mu_\alpha)$$
$$= \frac{(1-\alpha)}{2}||C_\alpha^{-1/2}(m_0-m_\alpha)||^2 - \frac{1-\alpha}{2}\log\det{}_2[(I-S_\alpha)^{-1}(I-S_0)] + \frac{\alpha}{2}||C_\alpha^{-1/2}(m_1-m_\alpha)||^2 - \frac{\alpha}{2}\log\det{}_2[(I-S_\alpha)^{-1}(I-S_1)].$$

Recent publications

[1] H.Q. Minh, F. Nielsen. *Geometric Jensen-Shannon divergence between Gaussian measures on Hilbert space*, **International Conference on Geometric Science of Information (GSI 2025).**

[2] G. Braun, B. Loureiro, H.Q. Minh, M. Imaizumi. *Fast Escape, Slow Convergence: Learning Dynamics of Phase Retrieval under Power-Law Data*, **International Conference on Learning Representations (ICLR 2025)**, accepted for publication.