Geometry of Gaussian Measures

Joseph Anderson

The Ohio State University

andejose@cse.ohio-state.edu

24 March 2014

Curves

Let (X, d) be a metric space.

Definition

Let $\gamma : [0, 1] \to X$ be continuous. We say that γ is a *curve*, and the *length* of γ is

$$L_d(\gamma) = \sup_{0=t_1<t_2<\cdots<t_N=1} \sum_{i=1}^{N} d(\gamma(t_i), \gamma(t_{i+1})).$$

The curve γ is said to be *rectifiable* if its length is finite.

Length Spaces

Definition

Let $x, x' \in X$ and let $\Gamma(x, x')$ be the family of curves joining x and x'. The *intrinsic metric* on X is defined as

$$d^*(x, x') = \inf(L_d(\gamma) \mid \gamma \in \Gamma(x, x')).$$

If $\Gamma(x, x') = \emptyset$, we say that $d^*(x, x') = \infty$.

- If $d = d^*$, then we say that d is intrinsic and we call (X, d^*) a *path metric space* or *length space*.
- One calls (X, d^*) *geodesic* if for any pair of points x, x' there exist $\gamma \in \Gamma(x, x')$ so that $L_d(\gamma) = d(x, x')$.

Space of Measures

Let $\mathcal{P}_2^{ac}(\mathbb{R}^d)$ be the set of all absolutely continuous measures on $\mathbb{R}^d$ with finite second moment.

- Absolutely continuous: $\mu << \lambda$ if $\lambda(A) = 0 \implies \mu(A) = 0$
- Finite second moment: $\int_{\mathbb{R}^d} d(x, x_0) d\mu(x) < \infty$ for all x_0.

Gaussian Measures

Recall the Gaussian measure on $\mathbb{R}^d$, $\phi_{\mu,\Sigma}$, where

$$\phi_{\mu,\Sigma}(A) = \frac{1}{\sqrt{\det(2\pi\Sigma)}^d} \int_A \exp\left(\frac{-1}{2} \langle x - \mu, \Sigma^{-1}(x - \mu) \rangle \right) d\lambda_n(x).$$

- Let $(\mathcal{N}^d, d_{\mathcal{W},2}^{\mathbb{R}^d})$ be the space of all Gaussian measures on $\mathbb{R}^d$ with the (restriction of the) 2-Wasserstein distance.
- Note: Gaussian measures are square integrable and absolutely continuous with respect to the Lebesgue measure, so $\mathcal{N}^d \subset \mathcal{P}_2^{ac}$

Some Basic Questions about $\mathcal{N}$

Let $\phi_1, \phi_2 \in \mathcal{N}^d$.

- Is it true that we can find a measure $\mu \in \mathcal{M}(\phi_1, \phi_2)$?
- How about $\mathcal{N}^{d^2} \cap \mathcal{M}(\phi_1, \phi_2)$?
- Does $\mathcal{N}^d$ have recognizable structure?

One-Dimensional Warm Up

Let $\phi_1, \phi_2 \in \mathcal{N}_1$ with mean and variance μ_1, σ_1^2 and μ_2, σ_2^2, respectively.

- Couplings exist, and are Gaussian on $\mathbb{R}^2$
- We can assume that $\mu_1 = \mu_2 = 0$
- We can compute

$$\begin{aligned} d_{\mathcal{W},2}(\phi_1, \phi_2) &= \left(\inf_{\mu \in \mathcal{M}(\phi_1, \phi_2)} \iint |x - y|^2 d\mu(x, y) \right)^{1/2} \\ &= |\sigma_1 - \sigma_2|. \end{aligned}$$

- Furthermore, this even forms a length space

Known Results

Theorem ([GM96])

Let μ, ν be Borel probability measures on $\mathbb{R}^d$. Then

1. *there exists a convex function ψ on $\mathbb{R}^d$ whose gradient $\nabla\psi$ pushes μ forward to ν*
2. *the gradient of ψ is determined up to μ-measure 0*
3. *the measure $\pi = (\mathrm{id} \times \nabla\psi)_{\#}\mu$ is optimal*
4. *π is the only optimal measure in $\mathcal{M}(\mu, \nu)$ unless $d_{\mathcal{W},2}(\mu, \nu) = +\infty$*

Theorem ([GS84])

For $\phi_{m,V}, \phi_{n,U} \in \mathcal{N}^d$, we have

$$d^2_{\mathcal{W},2} = \|m - n\|^2 + \mathrm{tr}\,(V) + \mathrm{tr}\,(U) - 2\mathrm{tr}\left(U^{\frac{1}{2}} V U^{\frac{1}{2}}\right)^{\frac{1}{2}}$$

Geometry of $\mathcal{N}^d$ and Applications

- Proved by [GM96] that $\mathcal{N}_0^d$ (mean 0 Gaussian measures) is geometrically convex as a subspace of $\mathcal{P}_2^{ac}(\mathbb{R}^d)$.
- The sectional curvature of $\mathcal{N}_0^d$ can be computed, c.f. [Tak08]
- Known curvature can help learning algorithms which rely on regularization
- Distance between "soft" shapes with Gaussian-like geodesic distances becomes easily computable and interpolated

References

Wilfrid Gangbo and Robert J McCann.
The geometry of optimal transportation.
Acta Mathematica, 177(2):113–161, 1996.

Clark R Givens and Rae Michael Shortt.
A class of wasserstein metrics for probability distributions.
The Michigan Mathematical Journal, 31(2):231–240, 1984.

Asuka Takatsu.
On wasserstein geometry of the space of gaussian measures.
2008.