A Wasserstein-type distance in the space of Wrapped Gaussian Mixtures

Michael Wilson

Abstract

We present a closed form expression for the Wasserstein Distance between two Wrapped Normal Distributions on the sphere. We then show how to use this distance to extend some results for mixtures of Gaussians in $\mathbb{R}^d$ to mixtures of Wrapped Normal distributions on the sphere.

1 Theoretical Details

Let $\mu = \mathcal{WN}(m, \Sigma)$ be a wrapped normal distribution defined on the $d-1$ dimensional sphere S^{d-1}. We can define the density of μ in terms of the tangent space $T_m(S_{d-1})$, with orthonormal basis B as;

$$\rho_\mu(x) = det(2\pi\Sigma)exp(-\frac{1}{2}\langle BExp_m^{-1}(x), \Sigma^{-1}BExp_m^{-1}(x)\rangle)$$

Here exp corresponds to the exponential function, while Exp corresponds to the Exponential Map for S^{d-1}.

1.1 Wrapped Normal Wasserstein Distance

Let $\mu_i = N(0, \Sigma_i)$, $i = \{0, 1\}$ be normal distributions defined on $\mathbb{R}^{d-1}$, with random variables $X_i \sim \mu_i$. For some $m_i \in S^{d-1}$, let $\tilde{\mu}_i \sim Exp(m_i)_\# \mu_i$, and let $\tilde{X}_i \sim \tilde{\mu}_i$. Then the Wasserstein distance between $\tilde{\mu}_0$ and $\tilde{\mu}_1$ can be calculated as;

$$W_2(\tilde{\mu}_0, \tilde{\mu}_1) = cos^{-1}(\langle m_0, m_1\rangle) + (\Sigma_0 + \Sigma_1 - 2(\Sigma_0^{\frac{1}{2}}\Sigma_1\Sigma_0^{\frac{1}{2}})^{\frac{1}{2}}) \tag{1}$$

Proof:

$$W_2(\tilde{\mu}_0, \tilde{\mu}_1) = \inf_{\pi\in\Pi(\mu_0,\mu_1)} E||\tilde{X}_1 - \tilde{X}_2|| \tag{2}$$

$$= \inf_{\pi \in \Pi(\mu_0,\mu_1)} \int_{S^{d-1}\times S^{d-1}} cos^{-1}(\langle x,y\rangle)\ d\pi(x,y) \tag{3}$$

Now, since $cos^{-1}(\langle x,y\rangle) = ||Exp_{m_0}^{-1}(y) - Exp_{m_0}^{-1}(x)||$ and $R_\alpha(Exp_{m_1}^{-1}(y)) = Exp_{m_0}^{-1}(y) - Exp_{m_0}^{-1}(m_1)$, we have that

$$cos^{-1}(\langle x,y\rangle) = ||R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x) + Exp_{m_0}^{-1}(m_1)|| \tag{4}$$

Thus,

$$\int_{S^{d-1}\times S^{d-1}} cos^{-1}(\langle x,y\rangle)d\pi(x,y) = \int_{S^{d-1}\times S^{d-1}} ||R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x) + Exp_{m_0}^{-1}(m_1)||d\pi(x,y) \tag{5}$$

$$= \int_{S^{d-1}\times S^{d-1}} ||R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x)\big)|| + \big\langle R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x)), (Exp_{m_0}^{-1}(m_1)\big\rangle + ||Exp_{m_0}^{-1}(m_1)||d\pi(x,y) \tag{6}$$

$$\begin{aligned} = ||Exp_{m_0}^{-1}(m_1)|| + \int_{S^{d-1}\times S^{d-1}} ||R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x)||d\pi(x,y) \\ + \int_{S^{d-1}\times S^{d-1}} \big\langle R_\alpha(Exp_{m_1}^{-1}(y)) - Exp_{m_0}^{-1}(x)), (Exp_{m_0}^{-1}(m_1)\big\rangle d\pi(x,y) \end{aligned} \tag{7}$$

$$= cos^{-1}(\langle m_0, m_1\rangle) + \int_{S^{d-1}\times S^{d-1}} cos^{-1}(\langle Exp_{m_0}(R_\alpha(Exp_{m_1}^{-1}(y)), x)\rangle)\ d\pi(x,y) \tag{8}$$

2 Implementation Details

References

[1] Computational optimal transport. *Foundations and Trends in Machine Learning*, 11(5-6):355–607, 2019.

[2] Julie Delon and Agnes Desolneux. A wasserstein-type distance in the space of gaussian mixture models, 2019.

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

[4] Asuka Takatsu. On wasserstein geometry of the space of gaussian measures, 2008.

[2] [4] [3] [1]