Manifold Matching via Deep Metric Learning for Generative Modeling – Supplementary Material

Mengyu Dai*
Microsoft
mendai@microsoft.com

Haibin Hang*
University of Delaware
haibin@udel.edu

Given any space X associated with a probability measure μ s.t. $\text{supp}[\mu] = X$ and $p \geq 1$, then for any function $f : X \to \mathbb{R}$ we denote its L^p norm as

$$\|f\|_p := \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p}.$$

Then we have the following properties:

Lemma 0.1. *Denote* $\|f\|_\infty := \sup_{x\in X} f(x)$*, then*

1. *If* $p \leq q$*, then* $\|f\|_p \leq \|f\|_q$*;*
2. $\lim_{p\to\infty} \|f\|_p = \|f\|_\infty$.

Proof. Let $r = q/p$ and s be such that $1/r + 1/s = 1$. Let $a(x) = |f(x)|^p$ and $b(x) \equiv 1$, then by Holder's inequality, we have

$$\|ab\|_1 \leq \|a\|_r \|b\|_s$$

or

$$\int_{x\in X} |f(x)|^p d\mu(x) \leq \left(\int_{x\in X} |f(x)|^{p\cdot\frac{q}{p}} d\mu(x)\right)^{p/q}.$$

Now taking the p-th root on both sides we obtain $\|f\|_p \leq \|f\|_q$.

It is easy to see $\|f\|_p \leq \|f\|_\infty$ for any $p > 1$. Now we choose $\forall\epsilon > 0$ with $\epsilon < \|f\|_\infty$ and let $X_\epsilon := \{x \in X | f(x) \geq \|f\|_\infty - \epsilon\}$, then

$$\begin{aligned}&\left(\int_{x\in X} |f(x)|^p d\mu(x)\right)^{1/p}\\ \geq&\left(\int_{X_\epsilon} (\|f\|_\infty - \epsilon)^p d\mu(x)\right)^{1/p}\\ =&(\|f\|_\infty - \epsilon)\mu(X_\epsilon)^{1/p}\end{aligned}$$

When $p \to \infty$ we have $\lim_{p\to\infty} \|f\|_p \geq \|f\|_\infty - \epsilon$. Since ϵ is arbitrary we have $\lim_{p\to\infty} \|f\|_p \geq \|f\|_\infty$. □

*Equal contributions.

1. Properties of geometric descriptors

Given metric measure space $\mathcal{X} = (X, d, \mu)$, intuitively, the Fréchet mean set $\sigma(\mathcal{X})$ represents the center of $\mathcal{X}$ with respect to metric d. Particularly,

Lemma 1.1. *The Fréchet mean of measure* μ *with respect to* d_E *coincides with the mean* $\overline{\mu} := \int_{\mathbb{R}^D} y d\mu(y)$.

Proof. Suppose that elements in $\mathbb{R}^D$ are represented as column vectors. Then $\sigma(\mathbb{R}^D, d, \mu)$ equals the set of minimizers of

$$F(x) = \int_{\mathbb{R}^D} (y - x)^T (y - x) d\mu(y).$$

Since $\frac{\partial F}{\partial x} = -2\int_{\mathbb{R}^D} y d\mu(y) + 2x$, let $\frac{\partial F}{\partial x} = 0$ we have single minimizer of $F(x)$ to be $\int_{\mathbb{R}^D} y d\mu(y) = \overline{\mu}$. □

When the underlying metric d is not the Euclidean metric, the Fréchet mean provides a better option of centroid.

Let $x_1, x_2, \cdots, x_k, \cdots$ be a sequence of independent identically distributed points sampled from μ. Let $\mu_k = \frac{1}{k}\Sigma_{i=1}^k \delta_{x_i}$ denote the empirical measure. We can estimate the Fréchet mean of (X, d, μ) by a set of random samples:

Proposition 1.2. $\sigma(\mathcal{X}) = \lim_{k\to\infty} \sigma(X, d, \mu_k)$.

Proof. Since the empirical measure μ_k weakly converges to μ, the result following by the definition of weak convergence. □

The following lemma explains its geometric meaning of p-diameters:

Lemma 1.3. *For any metric measure space* $\mathcal{X} := (X, d, \mu)$ *with* $\text{supp}[\mu] = X$*,*

1. $\text{diam}_p(\mathcal{X}) \leq \text{diam}_q(\mathcal{X})$ *for any* $p \leq q$*;*
2. $\lim_{p\to\infty} \text{diam}_p(\mathcal{X}) = \sup_{x,x'\in X} d(x, x')$.

Proof. It follows directly from Lemma 0.1 □

Similarly, we can estimate the p-diameter of a metric measure space $\mathcal{X} := (X, d, \mu)$ by a set of random samples:

Proposition 1.4.

$$\mathrm{diam}_p(\mathcal{X}) = \lim_{k\to\infty} \mathrm{diam}_p(X, d, \mu_k).$$

Proof. Since the empirical measure μ_k weakly converges to μ, the result following by the definition of weak convergence. $\square$

Proposition 1.5. *Given a metric measure space $\mathcal{X} := (X, d, \mu)$ and map $g : X \to \mathbb{R}^n$. If $d = g^* d_E$, then*

$$\sigma(\mathcal{X}) = g^{-1}(\overline{g_* \mu}).$$

Proof. By definition, $\sigma(\mathcal{X})$ is the set of minimizers of function

$$\begin{aligned} F(x) :&= \int_X (g^* d_E)^2 (x, y) d\mu(y) \\ &= \int_X d_E^2(g(x), g(y)) d\mu(y) \\ &= \int_{\mathbb{R}^n} d_E^2(g(x), z) d(g_*\mu)(z). \end{aligned}$$

It is easy to see that x_0 is a minimizer of $F(x)$ iff. $g(x_0)$ is a minimizer of $G(w) := \int_{\mathbb{R}^n} d_E^2(w, z) d(g_*\mu)(z)$. By Lemma 3.4, $\overline{g_*\mu} = \operatorname{argmin} G(w)$, hence we have $\sigma(\mathcal{X}) = \operatorname{argmin} F(x) = g^{-1}(\overline{g_*\mu})$. $\square$

Proof of Proposition 3.10. It follows directly from Proposition 1.5 $\square$

2. Network Architectures

For unconditional image generation task, we used a ResNet data generator and a deep convolutional net metric generator:

f_θ: convt(128) → upres(128) → upres(128) → upres(128) → upres(128) → bn → conv(128) → sig;

g_w: conv(32) → leaky-relu(0.2) → conv(64) → leaky-relu(0.2) → conv(128) → leaky-relu(0.2) → conv(256) → leaky-relu(0.2) → conv(512) → leaky-relu(0.2) → maxpool → dense(10).

For super-resolution task, the architecture of Triplet embedding network is presented as below:

g_w: (conv(x→2x, x_0=32) → prelu → maxpool) * 7 → dense(256) → prelu → dense(256) → prelu → dense(32),

where conv, convt, upres, bn, relu, leaky-relu, prelu, maxpool, dense and sig refer to nn.Conv2d, nn.ConvTranspose2d, up-ResidualBlock, nn.BatchNorm2d, nn.ReLU, nn.LeakyReLU, nn.PReLu, nn.MaxPool2d, nn.Linear and nn.Sigmoid layers in Pytorch framework respectively.

3. More Evaluation Details

Perception-Based Evaluations in SISR: we adopted the function $niqe$ in Matlab for computing NIQE scores, and python package $lpips$ for computing LPIPS.

FID Evaluation: we used 50000 randomly generated samples comparing against 50000 random samples from real data sets for testing. Features were extracted from the pool3 layer of a pre-trained Inception network. FID was computed over 10 bootstrap resamplings.