Neural Optimal Transport Auto-encoders 👾

Evgeny Lagutin

Adviser: Alexey Kolosov

×

https://github.com/sverdoot/NOT_AE

Optimal Transport formulations 🚕

Let $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ be a **cost** function.

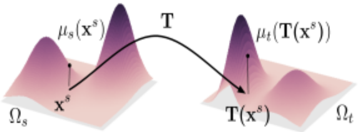
Monge's formulation:

$$\mathrm{Cost}(\mathbb{P}, \mathbb{Q}) = \inf_{T_{\#}\mathbb{P} = \mathbb{Q}} \int_{\mathcal{X}} c(x, T(x)) d\mathbb{P}(x)$$

Kantorovich formulation:

$$\mathrm{Cost}(\mathbb{P}, \mathbb{Q}) = \inf_{\pi \in \Pi(\mathbb{P}, \mathbb{Q})} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y) d\pi(x, y)$$

if minimizer π^* is deterministic (i.e. $\pi^*(x) = [x, T^*(x)]_{\#}\mathbb{P}$) then T^* is a solution of Monge's OT formulation



Weak Optimal Transport 🛴

Let $C : \mathcal{X} \times \mathcal{P}(\mathcal{Y}) \to \mathbb{R}$ be a **weak cost** function.

$$\text{Cost}(\mathbb{P}, \mathbb{Q}) = \inf_{\pi \in \mathcal{P}(\mathcal{Y})} \int_{\mathcal{X}} C(x, \pi(\cdot \mid x)) d\mathbb{P}(x)$$

Weak formulation generalizes the strong with

$$C(x, \mu) = \int_{\mathcal{Y}} c(x, y) d\mu(y).$$

Dual formulation:

$$\text{Cost}(\mathbb{P}, \mathbb{Q}) = \sup_{f} \int_{\mathcal{X}} f^C(x) d\mathbb{P}(x) + \int_{\mathcal{Y}} f(y) d\mathbb{Q}(y),$$

$$\text{where } f^C(x) = \inf_{\mu \in \mathcal{P}(\mathcal{Y})} \{C(x, \mu) - \int_{\mathcal{Y}} f(y) d\mu(y)\}$$

"one-to-one" transport plan learning 🦊↔️🦁

Assume the optimal plan is deterministic, i.e. $\pi^*(x) = [x, T^*(x)]_{\#}\mathbb{P}$. For $C(x, \mu) = \int_{\mathcal{Y}} c(x, y) d\mu(y)$

$$\text{Cost}(\mathbb{P}, \mathbb{Q}) = \sup_{f} \left[\int_{\mathcal{X}} \inf_{\mu \in \mathcal{P}(\mathcal{Y})} \left\{ \int_{\mathcal{Y}} c(x, y) d\mu(y) - \int_{\mathcal{Y}} f(y) d\mu(y) \right\} d\mathbb{P}(x) + \int_{\mathcal{Y}} f(y) d\mathbb{Q}(y) \right] =$$

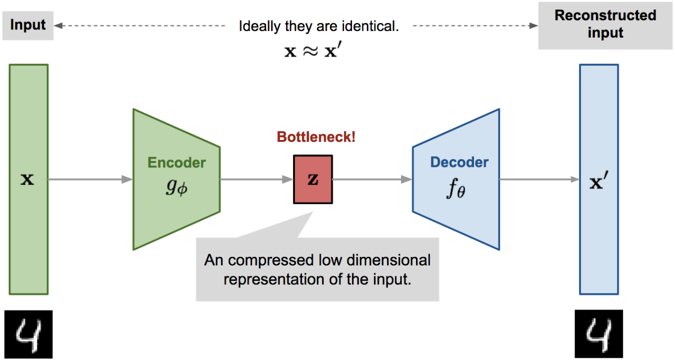
$$\sup_{f} \inf_{T} \left[\int_{\mathcal{X}} \{ c(x, T(x)) - f(T(x)) d\mathbb{P}(x) \} + \int_{\mathcal{Y}} f(y) d\mathbb{Q}(y) \right]$$

Auto-encoder setting 🦊↔️🦊

We aim to formulate an auto-encoder task as an optimal transport plan from $\mathbb{P}$ to $\mathbb{P}$

(the optimal transport plan is $[x, x]_{\#}\mathbb{P}$)

with explicit restriction on transport map: the intermediate dimension is smaller than $\dim(\mathcal{X})$.



Results 🎀

We measure LPIPS (perceptual similarity of original image and reconstruction) and FID ($W_2^2(\mathbb{P}_{\text{data}}, \mathbb{P}_{\text{rec}})$ in assumption that both $\mathbb{P}_{\text{data}}$ and $\mathbb{P}_{\text{rec}}$ are normal).

Method	Cost	test LPIPS (↓)	test FID (↓)
AE	L2	0.23	71.8
NOT-AE	L2	**0.14**	**58.4**
AE	L1	0.21	71.1
NOT-AE	L1	0.22	71.4

Table 1: Results on CelebA 32×32 dataset after 5k training iterations (50k backward steps of T. Intermediate dimension: 64 (48 times smaller than original)

Code for experiments with ArtBench10 dataset and Perceptual cost prepared, but results are not ready yet..

NOT - Neural Optimal Transport

Results 🎀

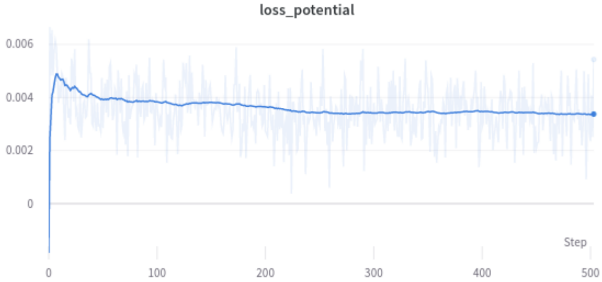


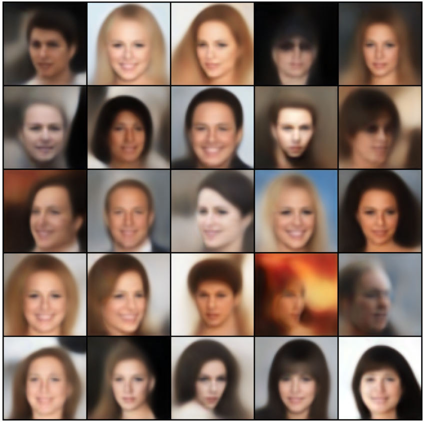
Figure 1: Potential's loss dynamic during training with L1 cost.

Observation: it is tough to balance stable training of map T and training of potential f (which is done by tuning of number of gradient steps for inner optimization problem)

Same problem for perceptual cost (pretrained VGG-11 as a feature encoder).

Results 🎀

(a) CelebA



(b) AE, L2 cost

Results 🎀

(a) CelebA

(b) NOT-AE, L2 cost

Roles

- Evgeny - preparing code for NOT, datasets and metrics, running experiments

- Alexey - consultations on model training, scaling experiments to run on a cluster, verification of experimental results

Thanks!