Continuous Fused Optimal Transport

Junhyoung Chung[*]
Department of Statistics
Seoul National University
Seoul 08826, Republic of Korea
junhyoung0534@gmail.com

Abstract

The abstract paragraph should be indented ½ inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

1 Introduction

2 Fused Optimal Transport

Notations. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and (S, d_S) a compact metric space. A measurable map $X : \Omega \to S$ is called a random element with distribution $\mathbb{P}_X := \mathbb{P} \circ X^{-1}$. We also introduce a feature space $M \subset \mathbb{R}^d$ which is compact, and call any one-to-one and continuous $f : S \to M$ a feature function. For two probability measures μ, ν on S, denote by

$$\Pi(\mu, \nu) \;=\; \{\pi \text{ on } S \times S : \text{ the marginals are } \mu \text{ and } \nu\}$$

the set of all couplings between μ and ν. Lastly, we say a measurable map $T : S \to S$ pushes forward μ to ν if $\mu(T^{-1}(A)) = \nu(A)$ for all $A \in \mathcal{B}(S)$, where $\mathcal{B}(S)$ is the Borel σ-algebra of S. We denote $T_{\#}\mu = \nu$ if T pushes forward μ to ν.

Fused optimal transport plan. For $0 \leq \alpha \leq 1$ and a feature function f, we consider

$$\begin{aligned} \inf_{\pi \in \Pi(\mathbb{P}_X, \mathbb{P}_Y)} & (1-\alpha)\mathbb{E}_{(X,Y)\sim\pi}\big[\|f(X) - f(Y)\|_2^2\big] \\ & + \alpha \mathbb{E}_{\substack{(X,Y)\sim\pi \\ (X',Y')\sim\pi}}\Big[K_h(X,X')\big|d_S(X,X') - d_S(Y,Y')\big|^2\Big], \end{aligned} \tag{1}$$

where $K_h : S \times S \to [0, \infty)$ is a bounded, continuous, and symmetric kernel with a bandwidth $h > 0$. We write $\mathcal{S}_\alpha$ for the set of minimizers of (1).

Proposition 1 (Existence of a minimizer)**.** *For each $0 \leq \alpha \leq 1$, the problem* (1) *admits at least one minimizer; that is, $\mathcal{S}_\alpha \neq \emptyset$.*

An equivalent formulation in feature space. Let $Z := f(X)$ and $W := f(Y)$, and consider couplings $\gamma \in \Pi(\mathbb{P}_Z, \mathbb{P}_W)$ on $M \times M$. Define $K_h^f(z, z') := K_h(g(z), g(z'))$ with $g := f^{-1}$ on $f(S)$. Then, (1) is equivalent to

$$\begin{aligned} \min_{\gamma \in \Pi(\mathbb{P}_Z, \mathbb{P}_W)} & (1-\alpha)\mathbb{E}_{(Z,W)\sim\gamma}\big[\|Z - W\|_2^2\big] \\ & + \alpha \mathbb{E}_{\substack{(Z,W)\sim\gamma \\ (Z',W')\sim\gamma}}\Big[K_h^f(Z,Z')\big|d_S\big(g(Z), g(Z')\big) - d_S\big(g(W), g(W')\big)\big|^2\Big]. \end{aligned} \tag{2}$$

[*]https://junhyoung-chung.github.io/

39th Conference on Neural Information Processing Systems (NeurIPS 2025).

Thus, π is feasible for (1) if and only if $\gamma = (f \times f)_{\#}\pi$ is feasible for (2) with the same objective value. In the sequel, we focus on analyzing (2) for mathematical simplicity.

References

Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

A Appendix