



Exact statistical inference for the Wasserstein distance by selective inference

Selective Inference for the Wasserstein Distance

Vo Nguyen Le Duy^{1,2} · Ichiro Takeuchi^{2,3}

Received: 5 January 2022 / Revised: 7 May 2022 / Accepted: 10 May 2022 /

Published online: 28 June 2022

© The Institute of Statistical Mathematics, Tokyo 2022

Abstract

In this paper, we study statistical inference for the Wasserstein distance, which has attracted much attention and has been applied to various machine learning tasks. Several studies have been proposed in the literature, but almost all of them are based on *asymptotic* approximation and do *not* have finite-sample validity. In this study, we propose an *exact (non-asymptotic)* inference method for the Wasserstein distance inspired by the concept of conditional selective inference (SI). To our knowledge, this is the first method that can provide a valid confidence interval (CI) for the Wasserstein distance with finite-sample coverage guarantee, which can be applied not only to one-dimensional problems but also to multi-dimensional problems. We evaluate the performance of the proposed method on both synthetic and real-world datasets.

Keywords Selective inference · Wasserstein distance · Confidence interval · Optimal transport · Uncertainty quantification

✉ Ichiro Takeuchi
ichiro.takeuchi@mae.nagoya-u.ac.jp

Vo Nguyen Le Duy
duy.mllab.nit@gmail.com

¹ Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan

² RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

³ Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan