

DETECT COMPLETE DEPENDENCE VIA TRACE CORRELATION IN THE PRESENCE OF MATRIX-VALUED RANDOM OBJECTS

Delin Zhao^{1,2} and Liping Zhu^{*2}

¹*Fuzhou University* and ²*Renmin University of China*

Abstract: Various metrics have been developed to test for statistical independence and measure the degree of nonlinear dependence between two random objects. Most of these metrics achieve their lower bound if and only if the two random objects are independent. However, it is often unclear how the two random objects are dependent if they attain their upper bound. Moreover, how to implement these metrics when one of the objects is matrix-valued is rarely touched in the literature. To address these issues, we introduce a new metric called trace correlation, which ranges from zero to one. It equals zero only if the two random objects are independent and attains one only if one random object is functionally dependent on the other. In addition, trace correlation allows one of the random objects to be matrix-valued. We estimate trace correlation using standard U-statistic theory and thoroughly study the asymptotic properties of resultant estimates. Furthermore, we adapt trace correlation in the reproducing kernel Hilbert space. Extensive simulations and an application to the MNIST dataset demonstrate the effectiveness and usefulness of trace correlation.

Key words and phrases: Complete dependence, independence test, matrix-valued object, trace correlation.

1. Introduction

Testing for statistical independence and measuring the degree of nonlinear dependence are fundamental issues in both statistics and machine learning communities. In classification problems, to what extent the features are predictive can be evaluated by quantifying the degree of nonlinear dependence between the features and the class labels. Let us use the MNIST dataset accessible at <http://yann.lecun.com/exdb/mnist/> as an example to illustrate this phenomenon. It comprises of 60,000 training images and 10,000 test images, each labeled with an integer between 0 and 9. The images have a resolution of 28×28 pixels and their values are scaled to the range of $[0, 1]$. We hide parts of the image each time and contaminate each image with standard Gaussian white noise. The contaminated pixels are subsequently replaced with 1 if they exceed 1, and set to 0 if they are negative, ensuring that pixel values fall within $[0, 1]$. Our objective

^{*}Corresponding author. E-mail: zhu.liping@ruc.edu.cn