The Annals of Applied Statistics
2009, Vol. 3, No. 4, 1285–1294
DOI: 10.1214/09-AOAS312E
Main article DOI: 10.1214/09-AOAS312

DISCUSSION OF: BROWNIAN DISTANCE COVARIANCE

BY ARTHUR GRETTON[1], KENJI FUKUMIZU AND BHARATH K. SRIPERUMBUDUR[2]

Carnegie Mellon University, MPI for Biological Cybernetics, The Institute of Statistical Mathematics, Department of Electrical and MPI for Biological Cybernetics and Computer Engineering, UCSD

1. Introduction. A dependence statistic, the Brownian Distance Covariance, has been proposed for use in dependence measurement and independence testing: we refer to this contribution henceforth as SR [we also note the earlier work on this topic of Székely, Rizzo and Bakirov (2007)]. Some advantages of the authors' approach are that the random variables X and Y being tested may have arbitrary dimension $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively; and the test is consistent against all alternatives subject to the conditions $\mathbf{E}\|X\|_p < \infty$ and $\mathbf{E}\|X\|_q < \infty$.

In our discussion we review and compare against a number of related dependence measures that have appeared in the statistics and machine learning literature. We begin with distances of the form of SR, equation (2.2), most notably the work of Feuerverger (1993); Kankainen (1995); Kankainen and Ushakov (1998); Ushakov (1999), which we describe in Section 2: these measures have been formulated only for the case $p = q = 1$, however. In Section 3 we turn to more recent dependence measures which are computed between mappings of the probability distributions $\mathbf{P}_x$, $\mathbf{P}_y$, and $\mathbf{P}_{xy}$ of X, Y, and (X, Y), respectively, to high dimensional feature spaces: specifically, reproducing kernel Hilbert spaces (RKHSs). The RKHS dependence statistics may be based on the distance [Smola et al. (2007), Section 2.3], covariance [Gretton et al. (2005a, 2005b, 2008)], or correlation [Dauxois and Nkiet (1998); Bach and Jordan (2002); Fukumizu, Bach and Gretton (2007); Fukumizu et al. (2008)] between the feature mappings, and make smoothness assumptions which can improve the power of the tests over approaches relying on distances between the unmapped variables. When the RKHSs are characteristic [Fukumizu et al. (2008); Sriperumbudur et al. (2008)], meaning that the feature mapping from the space of probability measures to the RKHS is injective, the kernel-based tests are consistent for all probability measures generating (X, Y).

[1]Supported by Grants DARPA IPTO FA8750-09-1-0141, ONR MURI N000140710747, and ARO MURI W911NF0810242.

[2]Supported by Max Planck Institute (MPI) for Biological Cybernetics, NSF Grant DMS-MSPA 0625409, the Fair Isaac Corporation, and the University of California MICRO program.

Key words and phrases. Independence testing, Brownian distance covariance, covariance operator, kernel methods, reproducing kernel Hilbert space.