Mutual distance covariance

Gábor J. Székely, **Tamás Fegyverneki**
Eötvös Loránd Tudományegyetem, Budapest, Hungary

Distance correlation is a multivariate dependence coefficient, analogous to the classical Pearson product-moment correlation. It was defined by Gábor Székely [2]. Distance covariance and correlation are applicable to random vectors of arbitrary and not necessarily equal dimension, and characterize independence, so they equal zero if and only if the random variables are independent. Distance correlation can be generalized by using α-th powers of Euclidean distances, having the same benefits if $0 < \alpha < 2$. (Distance correlation with $\alpha = 2$ leads to the absolute value of the classical product-moment correlation and thus does not characterize independence.)

Brownian covariance is also a measure of dependence between random vectors. It is a special case of a covariance with respect to a stochastic process, using the Brownian motion. It is shown that distance covariance coincides with Brownian covariance, so they can be called Brownian distance covariance [1]. Brownian covariance is also a natural extension for classical covariance, as we obtain Pearson's product-moment covariance by replacing Wiener process in the definition with identity. Replacing the Wiener process with Lévy fractional Wiener process with Hurst exponent $\frac{\alpha}{2}$ leads to α-distance covariance.

Unlike other dependence coefficients that characterize independence, distance covariance and distance correlation can easily be computed. The empirical computation formula applies to all sample sizes $n \geq 2$, is not constrained by the dimension and does not require parameter estimation or matrix inversion. The simplicity of the formula allows us to easily define the extension of empirical distance covariance for more than two random vectors. We extend the bivariate formula in [2] for three random vectors and discuss the meaning of the resulting mutual distance correlation. A general formula is also claimed for arbitrary many random vectors.

[1] G. J. Székely, M. L. Rizzo, Brownian distance covariance, *Annals of Applied Statistics* **3** (2009), 1236–1265.

[2] G. J. Székely, M. L. Rizzo, N. K. Bakirov, Measuring and testing dependence by correlation of distances, *Annals of Statistics* **35** (2007), 2769–2794.