Simulating the effect of Riemannian curvature on statistical inference using Brownian motion

Nicholas Wisniewski
Department of Medicine
University of California, Los Angeles
10833 Le Conte Ave, A2-237 CHS, Los Angeles, CA 90095-1679
e-mail: nw@ucla.edu

A number of information geometric distance measures, when used as test statistics, are known to asymptotically follow a χ^2 distribution. Examples of such statistics are the likelihood ratio, Wald statistic, Kullback-Leibler divergence, and geodesic distance. The asymptotic distribution is derived using Euclidean geometry, since as sample size $n \to \infty$, deviations are confined to the tangent plane of the hypothesis being tested. However, their non-asymptotic properties are much less understood; as deviations become large, the effect of the Riemannian curvature becomes apparent, and exact calculation becomes prohibitively difficult.

We investigated the effect of curvature on statistical inference in the family of normal distributions by numerically simulating Brownian motion as a generalization of the central limit theorem. Brownian motion on Riemannian manifolds differs from Euclidean diffusion due to Christoffel forces that arise from the curvature. Because the family of normal distributions forms a manifold of constant negative curvature, Brownian motion accelerates at farther distances, leading to a thicker tail than the χ^2 distribution. Additionally, because the Riemannian curvature of Gaussian models becomes more negative with higher dimensionality, this effect increases as more parameters are being estimated simultaneously.

The result shows how curvature effects can lead to significant deviations from asymptotic theory, which comparatively underestimates the potential for large fluctuations, and therefore overestimates statistical significance in hypothesis testing. To illustrate the effect on different test statistics, we computed the empirical distribution of several infomation geometric distance measures commonly used in diffusion tensor imaging, using simulated Brownian motion on Gaussian manifolds of increasing dimension. We show the least distortion is experienced by the geodesic and log-Euclidean distances, followed by J-divergence, and lastly the Frobenius norm.

Keywords: Brownian motion, Riemannian curvature, diffusion tensor