Hypothesis Testing for Network Data Objects, With Applications in Functional Neuroimaging

Eric D. Kolaczyk*
Boston University, Boston, Massachusetts, USA – kolaczyk@bu.edu

In recent years, it has become common practice in many areas of science to use networks to summarize relational information in a set of measurements. Yet almost uniformly, the attention in this field has been focused upon the analysis of (usually large) individual networks. However, in the context of big data another paradigm is emerging, one in which it will also be critical to be able to analyse large collections of (sub)networks, i.e., datasets of network objects. For example, in the field of neuroscience, networks are being used to summarize functional or structural relationships between regions of interest in the brain in databases of 1000s of patients. One of the most basic tasks of interest in the analysis of such data is the testing of hypotheses, in answer to questions such as "Is there a difference between the networks of these two groups of subjects?" In the classical setting, where the unit of interest is a scalar or a vector, such questions are answered through the use of familiar two-sample testing strategies. Networks, however, are not Euclidean objects, and hence classical methods do not directly apply. We address this challenge by developing a framework for asymptotic inference with network data objects, drawing on concepts and techniques from geometry, shape analysis, and high-dimensional statistical inference. Our work relies on a precise geometric characterization of the space of graph Laplacian matrices and a nonparametric notion of averaging due to Frechet. We motivate and illustrate our resulting methodologies for testing in the context of networks derived from functional neuroimaging data on human subjects from the 1000 Functional Connectomes Project.
In particular, we show that this global test is more statistical powerful, than a mass-univariate approach. This is joint work with Cedric Ginestet, Lizhen Lin, and Steve Rosenberg.

Keywords: networks; shape analysis; geometry; asymptotics.