Copula Regression Analysis of Networked Data

Peter X.-K. Song

Networked data are becoming increasingly popular in various fields. Such data refer to multivariate correlated measurements from one or multiple networks. One example of such data is unemployment rates measured over some social or community networks. In this talk, we will introduce a copula dependence model that accommodates relationship among nodes in a network on the basis of an adjacency matrix, with entry 1 indicating the presence of undirected edge between two nodes and 0 otherwise. We develop an innovative composite estimating functions approach to perform estimation and inference on the model parameters. We also incorporate a procedure of sparse network learning in the modeling and estimation of network topologies. The proposed model and analytic methods will be illustrated by simulation experiments and real-world data examples.

References

[1] J. Friedman, T. Hastie, and R. Tibshirani, *Sparse inverse covariance estimation with the graphical lasso*, Biostatistics 9, 432-441 (2008).

[2] E.D. Kolaczyk, *Statistical Analysis of Network Data*, Springer (2010).

[3] M.E.J. Newman, *Networks: An Introduction*, Oxford University Press (2010).

[4] J. Peng, P. Wang, N. Zhou, and J. Zhu, *Partial correlation estimation by joint sparse regression*. Journal of the American Statistical Association 104, 735-746 (2009).

[5] P.X.K. Song, *Correlated Data Analysis: Modeling, Analytics and Applications*, Springer (2007).

[6] C. Varin, N. Reid and D. Firth, *An Overview of Composite Likelihood Methods*, Statistica Sinica 21, 5-42 (2011).

[7] M. Yuan and Y. Lin, *Model selection and estimation in the Gaussian graphical model*, Biometrika 94, 19-35 (2007).

[1]Peter X.K. Song
Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029, USA
e-mail:pxsong@umich.edu