Combining Smoothing Spline with Conditional Gaussian Graphical Model for Density and Graph Estimation

Yuedong Wang

University of California-Santa Barbara
E-mail: yuedong@pstat.ucsb.edu

Abstract: Multivariate density estimation and graphical models play important roles in statistical learning. The estimated density can be used to construct a graphical model that reveals conditional relationships whereas a graphical structure can be used to build models for density estimation. Our goal is to construct a consolidated framework that can perform both density and graph estimation. Denote Z as the random vector of interest with density function f(z). Splitting Z into two parts, Z=(X,Y) and writing f(z)=f(x)f(y|x) where f(x) is the density function of X and f(y|x) is the conditional density of Y|X=x. We propose a semiparametric framework that models f(x) nonparametrically using a smoothing spline ANOVA (SS ANOVA) model and f(y|x) parametrically using a conditional Gaussian graphical model (cGGM). Combining flexibility of the SS ANOVA model with succinctness of the cGGM, this framework allows us to deal with high-dimensional data without assuming a joint Gaussian distribution. We propose a backfitting estimation procedure for the cGGM with a computationally efficient approach for selection of tuning parameters. We also develop a geometric inference approach for edge selection. We establish asymptotic convergence properties for both the parameter and density estimation. The performance of the proposed method is evaluated through extensive simulation studies and real data applications.