

Bayesian Quantile Semiparametric Mixed-Effects Double Regression Models for Analyzing Longitudinal Data with Non-Ignorable Missing Responses

Mai Dao, *Wichita State University*

Semiparametric mixed-effects double regression models have demonstrated satisfactory efficacy across diverse applications of longitudinal studies. However, the estimation of these models often relies on the assumptions of normally distributed errors and complete data with no missing values. Therefore, such restrictions may limit the practical usage of these models when analyzing longitudinal data that exhibit heavy-tailed behaviors and/or contain missing values. In this talk, we introduce a Bayesian quantile regression-based semiparametric mixed-effects double regression model for examining longitudinal data with non-ignorable missing responses. Here, the quantile regression is used to address non-normality issues and the missing mechanism is defined through a logistic regression model. Our proposed algorithm can concurrently model both the mean and variance of the mixed effects as functions of predictors while investigating the covariate effects at different quantiles of interest. Additionally, we utilize the Bayesian adaptive LASSO hierarchical model to devise an effective Metropolis-Hastings-within-Gibbs computation algorithm for both parameter estimation and variable selection purposes. Finally, we conduct different simulation studies and a real-data example to demonstrate the successful implementation of our proposed Bayesian methodology.

Explained Kernel Embedding Variation for Sufficient Dimension Reduction

Chenlu Ke, *Virginia Commonwealth University*

We propose a new framework for sufficient dimension reduction based on explained kernel embedding variation. The main idea is to view dimension reduction through a model-free ANOVA decomposition in reproducing kernel Hilbert spaces and to estimate low-dimensional linear combinations of predictors that maximize regression information. This formulation unifies several SDR targets: with characteristic kernels it targets the central subspace, while with linear kernels it reduces to central mean subspace estimation. On the computational side, we combine a sequential procedure that builds an ordered set of directions with a joint refinement step over the Grassmann manifold. We also develop a cross-validation-based dimension selection procedure using held-out kernel embedding prediction loss. The proposed method is flexible, model-free, and naturally accommodates different response kernels depending on the inferential target. Numerical studies illustrate its ability to recover both mean and distributional sufficient directions, and comparisons with existing sufficient dimension reduction methods demonstrate its competitive performance.

Comparison of Small Area Procedures Based on Gamma Distributions with Extension to Informative Sampling

Yanghyeon Cho, *University of Idaho*

The gamma distribution is a useful model for small area

prediction of a skewed response variable. We study the use of the gamma distribution for small area prediction. We emphasize a model, called the gamma-gamma model, in which the area random effects have gamma distributions. We compare this model to a generalized linear mixed model. Each of these two models has been proposed independently in the literature, but the two models have not yet been formally compared. We evaluate the properties of two mean square error estimators for the gamma-gamma model, both of which incorporate corrections for the bias of the estimator of the leading term. Finally, we extend the gamma-gamma model to informative sampling. We conduct thorough simulation studies to assess the properties of the alternative predictors. We apply the proposed methods to data from an agricultural survey.

A Scalable Bayesian Spatial Topic Model for Large-Scale Multiplexed Imaging Analysis

Xiyu Peng, *Texas A&M University*

Recent progress in multiplexed tissue imaging is deepening our understanding of tumor microenvironments related to treatment response and disease progression. However, analyzing whole-slide images with millions of cells remains computationally challenging, and few methods provide a principled approach for integrative analysis across images. Here, we introduce SpatialTopic, a spatial topic model designed to decode high-level spatial tissue architecture from multiplexed images. This algorithm integrates both cell type and spatial information within a topic modelling framework, originally developed for natural language processing and adapted for computer vision. The model employs an efficient collapsed Gibbs sampling algorithm for inference. We benchmarked SpatialTopic across diverse single-cell spatial transcriptomic and proteomic imaging platforms spanning multiple tissue types. We show that SpatialTopic is highly scalable to large-scale images, along with high precision and interpretability. It consistently identifies biologically and clinically significant spatial topics, such as tertiary lymphoid structures, and tracks spatial changes over disease progression. Its computational efficiency and broad applicability will enhance the analysis of large-scale imaging datasets.

