Estimation and Selection in Survival Models for Individuals with Spatial Frailty

Kyeongeun Kim[1], Joonho Shin[2], **Chae Young Lim[1]**

[1]*Department of Statistics, Seoul National University, Korea*

[2]*School of Mathematics, Statistics and Data Science, Sungshin Women's University, Korea*

ABSTRACT

In this talk, we introduce an estimating equation approach for a right-censored survival model with spatial random effects called spatial frailty. Our method accommodates multiple individuals at each site and incorporates spatial dependence of the individuals across sites. Additionally, we consider penalization for group variable selection, enabling the identification of key group variables, including categorical covariates, that influence survival times. We establish consistency and asymptotic normality as well as the oracle property of our estimator. A simulation study with various scenarios supports our theoretical findings. We also apply the proposed approach to real data, analyzing the survival of businesses in developing commercial districts of Seoul, South Korea, with a focus on spatial dependency using location-based map data. Our results suggest that the proposed method is a useful tool for analyzing spatially correlated survival data.

Keywords: Estimating equation; Spatial frailty; Spatial survival analysis; Variable selection

A Spatio-Temporal Modeling Approach for Wind Speed Data from a Regional Climate Model

Eva Murphy[1], Whitney Huang[2], Ting Fung Ma[3]

[1]Department of Statistical Sciences, Wake Forest University

[2]School of Mathematical and Statistical Sciences, Clemson University

[3]Department of Statistics, University of South Carolina

ABSTRACT

Statistical models that describe how physical processes vary across space and time are essential in environmental studies. We propose a spatio-temporal modeling framework for wind speed that is continuous in space and discrete at regular time intervals. The model adopts an additive decomposition: a smooth space–time function captures the mean structure, incorporating temporal periodicity associated with the annual cycle, while a combination of empirical orthogonal functions (EOFs) and a first-order dynamical Gaussian process enables a nonstationary spatial covariance structure and a parsimonious temporal dependence. The proposed approach is applied to regional climate model data from Canada, demonstrating improved and more consistent predictive performance compared to a baseline method. This is joint work with Eva Murphy at Wake Forest University and Ting Fung Ma from University of South Carolina.

Keywords: Spatio-temporal modeling; Empirical orthogonal functions; Gaussian processes; Wind Speeds

A Comparative Study of Neural Network Adaptations for Spatial Data

Bo-yu Chen[1] and Hao Zhang[2]

[1]*Department of Statistics, Purdue University*

[2]*Department of Statistics and Probability, Michigan State University*

ABSTRACT

In this study, we evaluated the predictive performance of various deep neural network approaches for spatial data. Recent advances in neural networks have led to a surge in deep learning methods and applications. Many of these methods were developed implicitly for independent data, and it is not trivial to incorporate spatial correlation into them. However, several strategies have been proposed to adapt neural networks to account for spatial correlation. Using both simulated and real-world datasets, we compared spatial methods developed from fully connected deep neural networks, analyzing the impact of different spatial adaptations. Our findings aim to inform future research directions in this evolving field.

Keywords: Deep learning, Kriging, Neural network, Prediction, Spatial statistics

Spectral Radii of Kernel Matrices and Applications to Kernel Score Tests

Hao Zhang

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

ABSTRACT

In spatial models and linear mixed-effects models, a central question is whether the data exhibit independence or, alternatively, whether spatial or random effects are present. One approach, derived from the score function (the derivative of the log-likelihood), was first developed in the context of linear mixed-effects models using the dot-product kernel. This framework was later generalized to incorporate the Gaussian kernel and other kernels, broadening its applicability. A notable recent application is in spatial transcriptomics, where the test is used to identify spatially varying genes. A practical challenge, however, is the selection of the kernel bandwidth (also called the range parameter in spatial statistics), which is often chosen empirically. Our recent theoretical results provide guidance on this choice. In this talk, I will give an overview of these results and their implications.

Key words: dependence, score tests, spatial models, spatial transcriptomics