Large Spatial Data Analysis: Multi-Resolution Spatial Methods on the Sphere and Bootstrap Confidence Intervals for Estimation on $\mathbb{R}^2$

黃灝勻
國立清華大學統計學研究所

Abstract

In the first part of the work, we investigate spatial prediction on the sphere in the presence of measurement errors, where data may be irregularly located. We first develop a special class of basis functions in the thin-plate-spline (TPS) function space on the sphere. These basis functions are ordered according to their level of smoothness from large-scale features to small-scale details, providing a multi-resolution representation and an orthogonal transformation of the data. Theoretically, we show that the number of basis functions selected by conditional Akaike information criterion is small, and the resulting reduced-rank estimate achieves a good convergence rate to the target function. In addition, we develop a multi-resolution mixed-effects spatial model on the sphere, by including a Gaussian spatial process to capture fine-scale information. Since large-scale features are captured by leading basis functions, the small-scale spatial process tends to have a short spatial dependence range, leading to a universal kriging estimate that allows rapid computations. A simulation experiment is performed, and an application to global sea-surface-temperature data observed from a satellite is given to demonstrate the effectiveness of the proposed method. In the second part of the work, we presents a comparative investigation of two bootstrap methods, percentile bootstrap (PB) and reverse percentile interval (RPI) for spatially corrected data, focusing on estimating confidence intervals for the parameters of the Mat´ern model. The paper addresses inference challenges under the infill asymptotic framework. It highlights the difficulties associated with the maximum likelihood estimators of the model parameters, which may not be consistent, and their asymptotic distributions are not known under infill asymptotics. The primary objective is to evaluate the performance of PB and RPI by assessing the coverage rate of confidence intervals and the interval score, considering both the accuracy and width of the intervals. From a theoretical perspective, we demonstrate that PB outperforms RPI in confidence estimation under some regularity. Simulation experiments confirm PB's superior performance over RPI, exhibiting higher accurate coverage rates and lower interval scores. By drawing connections to the micro-ergodic parameter, the study provides valuable insights into the factors influencing these outcomes, particularly when dealing with large spatial data.