Abstract

Spatial regression is an important topic in current days research and is widely applied in diverse disciplines including public health, epidemiology, environmental sciences, forestry and engineering. The existing spatial regression techniques assume the underlying spatial structure is known and common regression surface throughout the study region. In many practical applications neither of these assumptions holds. This is particularly true for large study regions that are associated with LARGE data. We develop an innovative method based on clustering where the spatial structure need not be known and different spatial region can have different regression surface. The method is computationally fast, in the order of $O(n)$ instead of $O(n^3)$. Since the first step of regression is selecting a set of suitable covariates, we also develop spatial variable selection. The theory of variable selection differs from the existing theory of variable selection. We develop increasing domain asymptotic to proof the consistency.