Statistica Sinica **35** (2025), 1-21
doi:https://doi.org/10.5705/ss.202022.0221

A HOMOGENEOUS LIKELIHOOD RATIO MEASURE FOR HIDDEN JUMP-SETS IN GENERALIZED SPATIAL REGRESSION

Pei-Sheng Lin*[1], Jun Zhu[2] and Katherine J. Curtis[2]

[1]*National Health Research Institutes and* [2]*University of Wisconsin-Madison*

Abstract: Hidden structures indicative of additional patterns relevant to the scientific inquiry are generally ignored and thus, the classical spatial regression analysis could miss important information carried by the latent variables. We develop novel methodology for uncovering some of such possible structures and patterns in spatial regression analysis. Our approach is to simultaneously model regression terms and hidden jump sets that occur abruptly across space in the presence of spatial dependence. An inequality for the homogeneity measure is derived by which we establish the consistency of jump-set selection. We devise a three-step computational algorithm based on a quasi-likelihood function and homogeneity measure to uncover patterns related to jump coefficients. Under suitable regularity conditions, we prove that the identification procedure is consistent when the hidden jump sets, covariates, and spatial correlation are incorporated into the model from the outset. The simulation study also shows sound finite-sample properties. In a case study, we examine closely county-based poverty rates in relation to industrial and racial compositions prior to the decline of manufacturing in the Upper Midwest of the U.S. Our case study reveals important socio-economic factors on poverty and additionally interesting structures and patterns not detected in classical spatial regression.

Key words and phrases: Homogeneity measure, quasi-likelihood ratio, spatial statistics.

1. Introduction

Spatial regression analysis is widely used in many scientific disciplines such as the social sciences and public health fields, for relating a response variable to explanatory variables across space while assuming spatially correlated errors (Cressie, 1993). In practice, the relation between the response and the explanatory variables is viewed as of primary interest, while accounting for spatial correlation in the error is understood to be important for a proper inference about the relation. Although sensible and popular, we believe this way of conducting regression analysis for spatial data can miss structures in the data indicative of additional patterns relevant to a scientific study. The objective of this paper is to

*Corresponding author. E-mail: pslin@nhri.edu.tw