Analyzing the Sensitivity of a Spatial Filter Model Specification when Modeling Land Use Change

Parmanand Sinha

The University of Texas at Dallas, 800 W. Campbell rd., Richardson, Texas, USA

Email: parmanand.sinha@utdallas.edu

Abstract: *This study compares the effect of changing the spatial resolution of the eigenvectors from finer to coarser on the model parameters in eigenvector spatial filtering specification model for multinomial discrete choice model. It has been applied to land use change dataset with four unordered response variables.*

Keywords: Eigenvector spatial filtering, land-use modelling, spatial discrete choice model.

Spatial discrete choice models best suit unordered categorical response variables like land use change, but there is very little literature available on their application on large datasets. Nearly all the literature available is about binary or ordered response variables. The principal reason is computational issues regarding the estimators. A few papers are available use pseudo-maximum likelihood, generalized method of moments, Bayesian MCMC-based estimator and maximum approximate composite marginal likelihood (MACML) estimation (Sidharthan and Bhat 2012) to estimate the parameters of spatial discrete choice models. But all these approaches require sophisticated computing, and have been applied to relatively small datasets. Eigenvector spatial filtering (ESF) (Griffith 2004) has been applied to a binary dataset, and requires shorter and more straightforward computation compared to other estimation techniques (Wang, Kockelman, and Wang 2013). In the case of a multinomial dataset, the computation time significantly increases. Multinomial auto logistic regression using ESF can be written as

$$U_{ij} = X_i\beta_j + E_K\beta_E + \varepsilon_{ij}$$

$$P(Y_{ij} = 1) = \frac{exp(U_{ij})}{\sum_{j=1}^{n} exp(U_{ij})}$$

where U_{ij} is a latent dependent variable representing the underlying utility from choosing a given alternative j, $P(Y_{ij} = 1)$ is the probability of land parcel i have land use j, X_i is a matrix of explanatory variables, E_K is an n-by-K matrix containing K eigenvectors, β_E is the corresponding vector of regression parameters, and ε_{ij} is a vector of identically distributed errors. The K eigenvectors are selected from the candidate set using stepwise multinomial logistic regression maximizing model fit at each step, which requires lots of computation time as n increases. One possible way to reduce this computation time is to reduce the number of eigenvectors in the candidate set, which can be done by making the spatial resolution coarser for the eigenvectors while keeping the same spatial resolution for the response variable as shown in figure 1 and 2 where nine cells have been aggregated to one for eigenvectors. This study compares the effect of changing the resolution of the eigenvectors from finer to coarser on the model parameters and computation time.

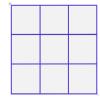*Figure 1. showing finer spatial resolution for response variable*

Figure 2. showing coarser spatial resolution for eigen-vectors

Study Area and Data

This procedure is applied to a 2000 to 2010 land use change dataset using four mutually exclusive land use categories: (1) Residential (R); (2) Commercial (R); (3) Open Space (OS); and (4) Vacant. The study area selected for this research is part of the City of Colony and Frisco, TX. With a total area of 32.28 square

miles, this is a rapidly developing suburban area located in the northern part of the Dallas-Fort Worth metroplex.

For the spatial econometric analysis, the area is divided into 10,201 square cells, each of size 300 ft × 300 ft. Five covariates are used in this regression: (1) Proximity to Dallas North Toll way & SR121; (2) Proximity to Eldorado Pkwy; (3) Proximity to arterial road; (4) Proximity to water bodies and lakes; (5) Proximity to shopping center; (6) Proximity to nearest school; (7) Proximity to city center; (8) Proximity to Exide battery plant. Out of 6844 vacant cells in 2000, 33.88 % cells were converted to residential, 7.66% cells were converted to commercial and 12.03% were converted to parks and open spaces in 10 years.
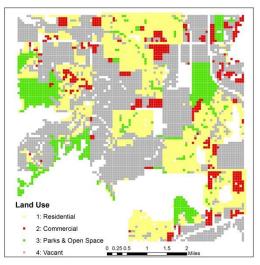

Figure 3.Study Area showing land use change between 2000 and 2010

Results

Initial results show that the computation time decreases as the spatial resolution of the eigenvectors is made coarser. The predicted value gets better when more eigenvectors are added to the model. Figure 4 and 5 compares the estimates and standard errors of estimates obtained by higher and lower resolution eigenvectors.
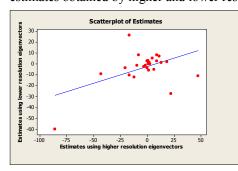

Figure 4. Scatterplot of estimates using higher spatial resolution eigenvectors and lower resolution eigenvectors.
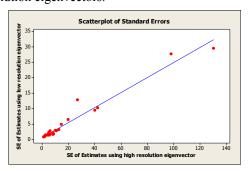
Figure 5. Scatterplot of standard error of estimates using higher spatial resolution eigenvectors and lower resolution eigenvectors.

The research currently under progress, aims to examine how much this increased coarseness affects the coefficient estimates for the covariates, and what spatial resolution can perform optimally during the modeling of land use change for large areas.

Reference

Griffith, Daniel A. 2004. "A Spatial Filtering Specification for the Autologistic Model." *Environment and Planning A* 36 (10).

Sidharthan, Raghuprasad, and Chandra R. Bhat. 2012. "Incorporating Spatial Dynamics and Temporal Dependency in Land Use Change Models." *Geographical Analysis* 44 (4): 321–49.

Wang, Yiyi, Kara M. Kockelman, and Xiaokun (Cara) Wang. 2013. "Understanding Spatial Filtering for Analysis of Land Use-Transport Data." *Journal of Transport Geography* 31 (July): 123–31.