Spatial Statistics 20 (2017) 221–243

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Is a matrix exponential specification suitable for the modeling of spatial correlation structures?

Magdalena E. Strauß [a], Maura Mezzetti [b,*], Samantha Leorato [b]

[a] MRC Biostatistics Unit, University of Cambridge, Robinson Way, Cambridge CB2 0SR, UK

[b] Department of Economics and Finance, Università Tor Vergata, Via Columbia 2, 00133 Rome, Italy

ARTICLE INFO

Article history:
Received 14 June 2016
Accepted 11 April 2017
Available online 27 April 2017

Keywords:
Matrix exponential
Covariance matrix
Spatial correlation

ABSTRACT

This paper investigates the adequacy of the matrix exponential spatial specifications (MESS) as an alternative to the widely used spatial autoregressive models (SAR). To provide as complete a picture as possible, we extend the analysis to all the main spatial models governed by matrix exponentials comparing them with their spatial autoregressive counterparts.

We propose a new implementation of Bayesian parameter estimation for the MESS model with vague prior distributions, which is shown to be precise and computationally efficient. Our implementations also account for spatially lagged regressors. We further allow for location-specific heterogeneity, which we model by including spatial splines. We conclude by comparing the performances of the different model specifications in applications to a real data set and by running simulations. Both the applications and the simulations suggest that the spatial splines are a flexible and efficient way to account for spatial heterogeneities governed by unknown mechanisms.

* Corresponding author.
E-mail addresses: magdalena.strauss@mrc-bsu.cam.ac.uk (M.E. Strauß), maura.mezzetti@uniroma2.it (M. Mezzetti), samantha.leorato@uniroma2.it (S. Leorato).

http://dx.doi.org/10.1016/j.spasta.2017.04.003
2211-6753/

1. Introduction

Data collected from geographic areas such as countries, regions, states, or individual points in space often exhibit spatial dependence, and require specific estimation methods to account for the lack of independence among the data. In recent years the economics literature has seen an increasing number of theoretical and applied econometric studies involving spatial dependence. While the interest in spatial models in economics is relatively recent, spatial models have a long history in the regional science, epidemiology and geography literature (see Anselin and Florax, 1995 for detailed references).

The widely used spatial autoregressive (SAR) approach was first introduced by Whittle (1954) and refers to the autoregressions that occur simultaneously at each data location. One drawback of the SAR model is that it requires specialized techniques for large samples. Ciu et al. (1996) first proposed exponential operators to specify a covariance matrix and also pointed out some advantages of the matrix exponential, but focused on general (non-spatial) covariance matrices. Later, LeSage and Pace (2007) proposed to apply the matrix exponential specification in a spatial context, as an alternative to the widely used SAR model. The resulting matrix exponential spatial specification model (MESS) replaces the conventional geometric decay of influence over space with an exponential pattern of decay. The MESS model has advantages, relative to the SAR, deriving from the characteristics of the matrix exponential reviewed in Section 2. However, it also has some disadvantages, the first of which is the difficult interpretation of the *correlation* parameter, which was also noted by LeSage and Pace (2007). One further concern related to the use of the MESS model, was raised by Rodrigues et al. (2014), who recently showed that it often induces opposite signs for the marginal and conditional correlations between two areas. We briefly discuss these two aspects in Appendix, which we devote to a comparison between MESS and SAR marginal effects and covariance structures.

Matrix exponentials can be introduced to define either the interaction between dependent variables or the spatial covariance of the errors. To these different approaches correspond two subclasses of the MESS models, usually referred to as MESS models and MESS error models, respectively. These models are alternatives to the SAR and spatial error models (SEM). The final goal of this paper is to contribute to the literature on the matrix exponential model, by assessing its validity, both on its own and relative to its main competitor, the SAR model. To take up this challenge and to allow a wider comparison with the SAR models, in Section 3 the different specifications of the spatial models with matrix exponential covariance are illustrated. As a possible way to allow MESS error models to account for location-specific heterogeneity, in Section 3.2, we moreover explore the effects of introducing spatial splines to cope with uncertainty of the spatial structure, which is acknowledged to be one common weak point of spatial linear regression models. We focus in particular on Bayesian estimation of the models: in Section 4 we propose a new implementation of Bayesian parameter estimation with vague prior distributions for both MESS and MESS error models.

To our knowledge, Bayesian approaches have never been used for the estimation of the latter. In fact, ever since the work of LeSage and Pace (2007), the literature has mainly focused on the first class, and MESS error models have been neglected, except for a brief description in LeSage and Pace (2009).

In contrast to previous model implementations, our method does not use Taylor series expansion with a fixed number of terms to approximate the matrix exponential; using an appropriate R package, the method used in our approach ensures that our approximation to the matrix exponential is always within a given fixed small interval around the true value. For the MESS model, we use an algorithm based on Krylov subspaces techniques developed by Sidje and Stewart (1999). Like the Taylor expansion method discussed in LeSage and Pace (2009), Sidje and Stewart (1999)'s algorithm directly computes the action of a matrix exponential on a vector without computing the matrix exponential itself. It was shown to be very efficient in Sidje and Stewart (1999), and it avoids lacking control on roundoff errors that may occur in Taylor series approximations due to alternating signs of the terms in the series. Because of these differences, a reasonable comparison between our implementation with those based on Taylor approximation of the exponential series, as the one proposed by LeSage and Pace (2009), is in terms of computation time. This is presented in Section 5.6.

Moreover in Section 5, by applying the different MESS models in an econometric application and in simulated data, we could assess its predictive ability with different weight matrices, and we find comparable performances relative to the SAR model with the same weight matrix choices.