Theory and Application of Dynamic Spatial Time Series Models

Andree, B.P.J.

2020

document version
Publisher's PDF, also known as Version of record

Chapter 8

Conclusion

The models that researchers estimate are necessarily an idealization of a complex reality. Advances in our capacity to compute, along with continued increases in the dimensions of datasets, have enormously increased both the complexity of what we attempt to achieve in analysis and the models that we use to pursue those goals. The aim of the basic theory with which we opened the introduction of this thesis was to provide clearly formulated and generalizable interpretation to standard empirical results. Given the advances in data and complexity, it is clear that analysis must acknowledge that the models ideally estimated aim at achieving a greater degree of idealization than was held possible when the theory of linear estimation of a parameter from a modest numbers of observations was first developed. With the general Consistency and Normality results for M-estimators that were introduced, there was much more freedom to think about more complex models that might provide a better description of reality. This thesis was devoted to exploring dynamic spatial time series models that can provide a better fit to the data using minimal complexity.

Chapter 3 first characterized spatial heterogeneity. This was done from the perspective of the data generating process itself. Specifically, we used a spatial model based on an economic rationale and parametrized it based on estimates from the literature. This was used to simulate likely economic outcomes at a grid-cell level. While inherently not a problem related to statistical inference in the way it was discussed in the introduction of this thesis, the analysis produced several useful insights. Specifically, we saw that by imposing simple linear relationships at a high resolution, aggregate system behavior tended to follow nonlinear patterns. This is important as, in reality, we tend to observe economic outcomes at a coarse scale while processes are arguably driven by the total sum of interactions between a large number of individual economic actors. Furthermore, we saw that the geophysical nature of our landscape plays an important role in economic processes. In particular, the natural organization in geological factors tends to contribute to spatial clustering, even when spatial interdependencies across various distances are not explicitly parameterized in the data generating process. This is also important, as it is easy to miss out on one or several unobserved common factors, that may follow this type of spatial organization, in empirical applications. This immediately implies that the residuals in simple cross-sectional regressions are likely to be spatially correlated and may follow structural patterns that vary by types of regimes. In the introduction of this thesis we had already emphasized the crucial role that neutralizing residuals plays in rendering the parameter distributions approximately normal.

In Chapter 4, we tackled the problem of spatial dependence in time series. Specifically, we specified the spatial autoregressive time series model discussed in the introduction of the thesis and studied it in more detail. Building on our notion that the linearity assumption may be too restrictive, especially as the spatial dimensions grow, we extended the model to allow the parameter that determines dependence between neighbors to vary across time and space in an idiosyncratic manner. This allows dependence to vary over different regimes that may be covered by the cross-sectional data. The model allowed each observation in the cross-section to have a different history of attraction to its neighbors and the magnitude of the induced feedback effects to vary continuously over time. This type of dynamic behavior could not be understood under standard dynamic time series theory provided in the introduction. We therefore extended the theory to allow for dynamic multivariate time series and provided a general theory that allowed the nonlinear dynamics to become spatial. We applied the model to a short spatial time series of urban densities and saw that the linear spatial model was not able to handle both the urban and rural dynamics in a single framework, causing the model to severely underestimate urban densities and overestimate rural densities. These regime-specific dependencies could, however, correctly be captured by the nonlinear model, allowing to analyze transitory effects across both the urban, rural, and urban gradients in one single framework. We also applied the nonlinear model to a long financial time series, and saw that it was able to fit both periods of financial stability during which spatial dependence was flat and periods of financial unrest in which there was substantially stronger idiosyncratic behavior.

In Chapter 5 we dropped the parametric assumption, and worked in a non-parametric framework in which the exact form of the nonlinearities did not have to be assumed. Instead of modeling the dependence between spatial observations to describe clustering in the data endogenously, we allowed for the flexibility to let dependence on exogenous variables vary nonlinearly across levels in the data. This resulted in rich dependence structures in which individual observations are part of different spatial and temporal regimes, each having possibly unique relationships with the outcome variable. We learned that there are methods that can approximate any type of nonlinearities arbitrarily well, while the estimation problem could still be solved linearly. In particular, the Kernel model mapped the input to a higher dimensional feature space, from where linear relationships could be established with the outcome variable. The growing number of local parameters used in those type of approximation strategies, however, violate the standard compactness assumption introduced in the introduction of the thesis that was used to obtain existence and measurability of the estimator. Hence, the uniform convergence that was obtained from point-wise convergence and stochastic equicontinuity on a compact parameter space was also lost. We saw that to estimate these models, it was necessary to regulate the size of the parameter space appropriately which ensured that there was sufficient data to support the degrees of freedom. The regularization method effectively ensured that the parameter space grew at an appropriate rate as the data grew. This delivered a type of consistency that had a different interpretation than what was discussed in the introduction of the thesis. In particular, the limit result depended on the user-defined tolerance for complexity, which was determined by a hyper-parameter that was not estimated by the criterion function itself. The appendix of this chapter discussed the implication of this external influence on the interpretation of the result and concluded that standard interpretation to the results is supported as long as the hyper-parameter was tuned by optimizing the criterion out-of-sample.

Chapter 6 moved away from the nonlinear world, and moved back into the linear one. In this chapter we focused on multivariate interactions between multiple spatial time-series. Naturally, once the asymptotic results for multivariate nonlinear time series models put forward in Chapter 3 and the penalization from Chapter 4 are understood, it is straightforward to apply these ideas together to the setting of multiple nonlinear spatial time-series. From a practical standpoint we, unfortunately, are still quite constrained by modern computing capacity to work with such complex descriptions of reality. Interesting linear dynamics between multiple spatial time series could still be modeled though, which admittedly already results in detailed dynamics at the observational level. In particular, the spatial spillover effects implied heterogeneous relationships at the local level, and the multiple variable setting thus allowed us to explore cause and effect between interrelated cross-sectional time series while taking into account that different cross-sectional variables themselves exhibit spatial feedback between observations that result in heterogeneous local impacts after shocks occur. We saw that models that do not factor in the cross-sectional dependences were likely to over-estimate the temporal effects and provided a generally poorer fit to the data that violated the martingale difference sequence assumption imposed on the score. Finally, the chapter explored the kernel trick from Chapter 4 as a mechanism to generate data-driven spatial weight matrices. The analysis showed that appropriate network structures could be estimated using Maximum Likelihood. This allowed generalizing the spatial dependencies discussed in this thesis and apply them to settings in which cross-sectional dependencies arise because of economic similarities or through other non-geographic channels.

Finally, in Chapter 7 we moved back to our starting discussion around estimators, and to the notion of correct specification specifically. Only this time, we approached the topic from a more general angle. We reconsidered the basic idea of inference and considered why flexible models, such as the ones introduced in this thesis, are desirable tools for inference in the first place. While the assumption of correct specification surfaced many times in parts of this thesis, it is easy to admit that this is possibly the most difficult assumption of all. In Chapter 4 and 6 we made use of different strategies to verify whether our estimated models provide an appropriate fit to the data. Nevertheless, when formulating empirical models we naturally abstract from reality and work with a description that is only an approximation to a complex reality. While mis-specification is often accepted in practice, it should not be a reason to opt for simple approximations merely because it is difficult to describe reality in fullness and easy to acknowledge that a simple model does not appropriately reflect that fullness. Particularly, when a result is taken as causal and representative of the real world, then that statement must reflect a belief that reality could be produced by a model that is reasonably similar to the estimated one. This means that if one is interested in making causal statements, then the estimated model used to build the arguments should at least be able to produce dynamics that we believe are relevant in the real world. In particular, the Stationarity and Ergodicity of the data introduced as an assumption in the introduction if this thesis must come from the model itself. If one is willing to verify all the stability conditions of the possibly complex analyzed dynamics, as we did in Chapter 3, then one must also be ensured that the empirical strategy that is followed inherently ensures that the estimator finds the correct causal structure. Critical here is that increasing model complexity leads to a higher number of parameters, hence an increased overall model uncertainty. We discussed approximation of causal structures in more detail and provided an argument that minimizing complexity penalized criteria such as the AIC, as we did in Chapters 4 and 6, is the right objective in empirical settings.

8.1 Final remarks

With the theory and methods introduced in this thesis, researchers can now estimate a wide range of flexible models that take into account possible heterogeneity in dependencies across time and space. While there are many thoroughly developed options for analysis of spatial time series data, there are still many possible other research methodologies left to cover. A few directions for future research are the following.

First, the applications in thesis focused primarily on modeling conditional mean sequences, possibly with observation-driven nonlinear dynamics. The notions put forward in this work can easily be extended to higher moments. For example, the nonlinear dynamics explored in the context of the smooth transition spatial autoregressive model could be extended to allow for nonlinear cross-sectional dependence in multivariate GARCH models to allow instantaneous transmission of volatility spillovers in an asymmetric way. This is particularly relevant when one is interested in understanding risk by means of numerically calculating Value-at-Risk or Expected Shortfall for a collection of interrelated investments using stochastic simulations. Basic univariate threshold GARCH models have already been developed to incorporate simple regime switching behavior into volatility regressions, but the standard application is one of instantaneous switching between linear autoregressive regimes. Spatial GARCH models have also been developed to allow for linear instantaneous dependence in processes that share AR and GARCH parameters. The obvious drawback is that, while financial assets may exhibit feedback, particularly when markets crash or surge, they may be assumed to follow individual temporal dynamics. From that perspective, Generalzied Orthogonal GARCH is a useful model as it allows one to parameterize interactions in the conditional mean sequence using a VAR structure, while also allowing for volatility spillovers in a multivariate GARCH equation. The GO-GARCH spillovers are, however, not instantaneous. Instead, they lag over time. Given that these various models are already available, a generalization of multivariate GARCH, spatial GARCH and the threshold dynamics, seems within reach of the practitioner. The resulting nonlinear spatial dependence in conditional mean and conditional variance, together with VAR parameters, would provide a framework in which one can analyze shocks that travel through a system, both in regimes that are dominated by commonalities or idiosyncrasies.

Second, not all the world's phenomena can be described with continuous data. Future research may focus on extensions relevant to model categorical, ordinal and count data that are collected sequentially over time at possibly dependent locations. This may require assuming distributions of a different type than those assumed in the theory developed here. For example the Poisson distribution would be the starting point for basic count series, and a Poisson mixture like the negative binomial distribution could be the starting point to tackle zero-inflation. Mixture models that involve multiple distributions can also be used to combine both the characteristics of continuous process and those of count process jointly in a time series. For example, the jump-diffusion model combines continuous Brownian motion paths from Gaussian log returns with discontinuities, or jumps, that are drawn from a compound Poisson process. Generalization of jump-diffusion to the spatial time series setting may be interesting, but possibly they will have to wait until spatial multivariate volatility models are better understood. The development could be particularly challenging because jumps may occur simultaneously in a spatial time series, but the magnitude of jumps may differ over the cross-section while the assimilation of these jumps into the series may also happen partly in an idiosyncratic manner.

Third, the state-space framework, in particular the Kalman filter, has been extremely important in time series analysis and much work can be done to integrate the idea of cross-sectional nonlinearity and spatial dependence into this framework. This may be a particularly interesting direction for further advancement when one deals with processes that are only partially observed or measured with possible error. The smoothing framework could be particularly helpful to develop nonlinear interpolations for spatial time series that are intermittently observed. Ultimately, this seems to be an unavoidable problem for which tools will be needed. If we assume that local data gathering processes operate and report back information independently from one another, then logically it becomes likely that there will be local series in close proximity of one another that overlap mildly at best when one starts to track more regions in an economic system. A basic example would be a survey program in which households in different areas report back on local market prices whenever they buy goods. The challenge of constructing a continuous spatial time series will then have to deal with missing observations in space and time. While this seems an advanced application, the problems are relevant to key policy indicators that have been gathered for a long time already. Currently, typical large survey programs such as those carried out by institutions like the World Bank, carry on for weeks or possibly months. During that time, seasons and economic circumstances may change. While the surveys thus actually represent a partially complete spatial time series, key statistics are often derived from them in the form of a single complete cross-section of data. The standard approach that many follow is to simply ignore away temporal changes assuming that they are randomly distributed over the survey program, and use the surveys to construct a single figure relevant for, say, the year. Often, one can find footnotes in reports and papers acknowledging that the underlying micro-data may have been gathered at different times. Performing a proper spatial time series interpolation before collapsing the data to a certain point in time would likely result in much more accurate estimates.

As we continue to develop theory for those complex settings, our datasets continue to grow increasingly rich, and the advances in our capacity to compute continue to accelerate, we may be able to model real-world processes in an increasingly accurate manner. The models we may use to approximate complex realities then become increasingly complex as well. We must therefore never forget the foundation on which we built. While we may achieve a greater degree of idealization than was ever held possible, the elegance of simple models was that they dealt with a modest numbers of parameters to summarize a complex world in a clearly formulated, tractable, an generalized fashion. Sometimes this is enough.