[FC-05-037] Spatial Autocorrelation

Abstract

The scientific term spatial autocorrelation describes Tobler's first law of geography: everything is related to everything else, but nearby things are more related than distant things. Spatial autocorrelation has a:

- past characterized by scientists' non-verbal awareness of it, followed by its formalization;
- present typified by its dissemination across numerous disciplines, its explication, its visualization, and its extension to non-normal data; and
- an anticipated future in which it becomes a standard in data analytic computer software packages, as well as a routinely considered feature of space-time data and in spatial optimization practice.

Positive spatial autocorrelation constitutes the focal point of its past and present; one expectation is that negative spatial autocorrelation will become a focal point of its future.

Keywords: Cliff and Ord, distance, econometrics, Moran Coefficient, neighborhoods, spatial statistics

Author & citation

Griffith, D. (2017). Spatial Autocorrelation. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2017 Edition), John P. Wilson (ed). DOI: 10.22224/gistbok/2017.4.13

Explanation

1. Introduction
2. Spatial Autocorrelation's Past
3. Spatial Autocorrelation's Present
4. Spatial Autocorrelation's Future

1. Introduction

Spatial autocorrelation (SA) is the correlation among georeferenced observations arising from their relative locations in geographic space. SA is everywhere! Although SA relates to a rather esoteric notion, namely correlated observations, it also may be understood in its relatively simple interpretation as a two-dimensional (e.g., map) pattern. Its most common everyday manifestation is moderate-to-strong and positive in nature and degree; a television screen or computer monitor image would not be discernible without positive SA. Many other real world examples exist, from artistic paintings, through magnetic sculptures as well as the concentration of households with similar socio-economic/demographic attributes in neighborhoods, to slider puzzles.

Recognition of SA as a scientific subject began in the early 1900s (Griffith, 2012a): in 1914, Student acknowledged that SA could influence the null statistical sampling distribution of the product moment correlation coefficient; in 1926, Yule acknowledged that SA could result in spurious variable correlation; and, in 1934, Stephan acknowledged that real world data exhibit SA rather than independent mixing. Accordingly, Fisher developed randomization techniques to neutralize SA effects in agricultural field trials. The concept of SA evolved from a conceptualization of correlation in the late 1800s, which then gave birth to temporal autocorrelation in the 1950s. The stage was set by 1950 for more formal academic treatments of SA. In that decade, Geary (1954), Moran (1950), and Whittle (1954) penned their respective famous pioneering articles, precipitating the establishment of a growing SA literature.

2. Spatial Autocorrelation's Past

As they allude to in their 2009 retrospective article, Cliff and Ord launched in the early 1970s the popularization of SA as an academic research topic. Getis (2007, p. 303) credits Golledge and Hubert for advancing a fundamental and more general understanding of SA in the late 1970s. He further notes that geostatistics also was developing in the 1960s, and that Paelinck coined the term spatial econometrics in 1974. Besag (1974) substantially augmented these efforts, coining the term auto-model to describe a wide range of probability models incorporating SA. Especially these three scholars helped move SA through the standard developmental stages of quantification, statistical distribution theory, and model specification. Cliff and Ord built upon the Moran Coefficient (MC) and Geary Ratio (GR) quantitative indices already existing. They established a statistical sampling distribution theory for these and other indices based upon both a random sampling and randomization (i.e., permutation) inferential basis. They also established the linear regression residual sampling distribution theory for the MC. Furthermore, they, as well as Besag, building on the work of Whittle and others, established spatial autoregressive model specifications. SA became a mainstream research topic in the tradition of normal curve theory, while presenting a number of different challenges preventing it from being extended to non-normal cases at that time.

As an awareness of SA emerged, many spatial scientists found it to be a somewhat confusing concept to fully understand. The literature offered little guidance here. In addition, geostatistics, developing in parallel to spatial autoregression, presented a much different conceptualization for both model specification and scientific visualization (Cressie, 1991). Meanwhile, Paelinck and Klaassen (1979) extended this notion to spatial economics, formally initiating the subdiscipline of spatial econometrics. Consequently, monographs by Goodchild (1986), Griffith (1987), and Odland (1988) appeared whose goal was to introduce spatial scientists to SA. The consensus was that SA could be interpreted in a number of different ways (Griffith, 1992).

SA interpretations include: self-correlation, map pattern, information content, spatial spillover, an indicator of areal unit demarcation appropriateness, a missing variable surrogate, a diagnostic tool, and a nuisance. Its literal definition—from its prefix auto—is self-correlation arising from attribute values' geographic context—from its adjective spatial: observations are correlated strictly due to their relative locational positions. SA relates to visualization; its most common interpretation is in terms of trends, gradients, or patterns across a map. Geostatistics exploits this property by emphasizing SA's spatial prediction/interpolation capabilities. Overwhelmingly, the most frequent nature and degree of SA encountered in empirical data to date is moderate-to-strong positive. In other words, a strong tendency exists for georeferenced values of some variable to cluster geographically such that nearby location concentrations occur of likewise values. Socio-economic/demographic tendencies have a propensity to be moderate; remotely sensed image tendencies have a propensity to be strong. Because it signifies duplicate/redundant information, SA also reflects information content in georeferenced data, vis-à-vis the standard information content for independent and identically distributed observations; this viewpoint relates to degrees of freedom as well as effective sample size. A remotely sensed image, for example, comprising hundreds of thousands of pixels may have an effective sample size of only a few hundred because of its latent SA. One of the most conspicuous impacts of positive SA is variance inflation, which means heavier tails for a normal random variable, and assorted frequency distribution distortions (e.g., skewness and kurtosis) for other random variables, such as binomial and Poisson. Consequently, increasing SA results in a bell-shaped frequency distribution that behaves similar to a t-distribution as its degrees of freedom change. Similarly, SA signifies a spillover of information from one location to another, resulting in redundant information being present in georeferenced data values, with this redundancy increasing as the degree of locational dependence increases—a notion analogous to that associated with classical correlation between two attribute variables. This duplicative information activates complications in the statistical analysis of georeferenced data that lie dormant in the statistical analysis of traditional data composed of independent observations, and that are similar to those encountered in time series analysis; these complications are exacerbated by the multidirectional, two-dimensional nature of spatial dependence (time series entail dependencies that are unidirectional along a single dimension).

Goodchild (1986) suggests that SA may be an outcome of a particular ill-conceived partitioning of a continuous surface into discrete areal units (i.e., polygons) for data aggregation purposes. Fine resolution contiguous georeferenced data can be aggregated in enormously many different ways, with the range of aggregations often capable of yielding almost any possible nature and degree of SA. This perspective relates to the modifiable areal unit problem (MAUP), and suggests that negative SA may be indicative of an inappropriate surface partitioning. It also suggests that, to mimic a random mixture, a surface partitioning could be gerrymandered in such a way that zero SA is its outcome. Meanwhile, spatial econometrics (Anselin 1988) views SA within the contexts of missing regression variables and diagnostic tools, especially for spatial heterogeneity. Griffith and Chun (2016) furnish eleven empirical examples illustrating the effectiveness of Moran eigenvector spatial filtering (MESF) to use SA to address the omitted variables problem for georeferenced data. In general, accounting for SA in a regression model tends to increase its variance accounted for capability by at least 5%-10%. Furthermore, Páez and Whalen (2010), among others, note that accounting for SA with an eigenvector spatial filter (ESF) tends to be effective in correcting regression residual non-normality, which also addresses issues of spatial and aspatial heterogeneity of variance.

Analysis issues such as missing variables bias allude to the more general concern of specification error diagnostics. Cliff and Ord (1973, p. 106) were two of the first spatial analysts to demonstrate convincingly the use of SA for such diagnostic purposes. They show that the residual MC from a bivariate regression of population percentage on an arterial road network index is a function of model specification; more specifically, the original detection of SA is an artefact of model misspecification. Similar to viewing SA as, for example, a missing variables or functional form misspecification diagnostic is viewing it as a nuisance parameter. A number of spatial econometricians do not consider SA to be of immediate interest to them, but recognize that it must be accounted for in georeferenced data in order to estimate and test the regression parameters that are of interest to them. This perspective is in contradistinction to that of most spatial statisticians, many of whom view SA as a data feature supporting missing values imputation (e.g., interpolation a la kriging in geostatistics), among other useful things.

3. Spatial Autocorrelation's Present

The present state of SA includes its further explication, its extension to local statistics and non-normal random variables, and its visualization. Estimation of semivariogram and auto-normal model parameters, as well as the development of spatial econometrics, are hallmarks of the transition into this present-day situation, as is a preference shift to almost sole use of the MC index because of its superior statistical properties.

Going beyond the introductory monographs by Goodchild (1986), Griffith (1987), and Odland (1988), especially Getis (2007, 2010), Griffith (2005, 2009), and Griffith and Chun (2014) furnish more contemporary elucidations of SA. Getis (2007) points out that SA research has grown into a vibrant discipline that transcends geography, encompassing many other disciplines, including agriculture, climatology, criminology, ecology, economics, epidemiology, forestry, hydrology, geology, meteorology, planning, public health, sociology, soil science, and statistics. One of the present-day developments Getis (2007) highlights is the formulation of local SA statistics as well as principles of spatial filtering. The two most popular of these former statistics are the LISA (Anselin, 1995) and the Gi (Getis and Ord, 1992). They are employed extensively to identify hotspots and coldspots in geographic distributions. Griffith and Getis (2016) furnish a state-of-the-art literature overview of spatial filtering that highlights its scope beyond MESF.

Although Paelinck and Klaassen (1979) conceptualized spatial econometrics, Anselin (1988) was this subdiscipline’s most effective promoter (playing the same role as Cliff and Ord did for spatial statistics), in part because of his software developments, including GeoDA, supporting its implementation. LeSage and Pace (e.g., 2009) added Bayesian approaches to the toolbox for this methodology. Especially with the advent of Markov chain Monte Carlo (MCMC) techniques, followed by integrated nested Laplace approximation (ILNA) techniques, these procedures have become feasible numerically, and hence have grown considerably in popularity. More recently, Elhorst (2014) went one step further by augmenting this literature with properly devised spatial panel methodology, better supporting space-time data analyses.

A major shift in the description of SA in more recent decades is from its basic quantification to its scientific visualization. One class of this visualization exploits the scatterplot, and comprises the semivariogram plot (see Cressie, 1991) in geostatistics, and the Moran scatterplot (Anselin, 1995) in spatial autoregression analysis. In geostatistics, this graphical tool plots geographic variance (the vertical axis) versus distance between locations (the horizontal axis) working directly with a geographic covariance matrix. Figure 1a illustrates this visualization for this article’s associated image (the superimposed trend line is for a combined K-Bessell function-spherical model). The slope of the resulting curved pattern in the scatter of data points is a function of the degree of positive SA. A non-linear trend line (described by a semivariogram model) almost always is superimposed upon this scatterplot. The graphical tool employed in autoregression analysis plots the sum of surrounding (i.e., neighboring) z-scores (the vertical axis) versus z-scores (the horizontal axis) — the inverse of a geographic covariance matrix. Figure 1b illustrates this visualization for this article's associated image (MC ≈ 0.85; GR ≈ 0.10). A linear trend line frequently is superimposed on this scatterplot; its slope is defined by the corresponding MC.

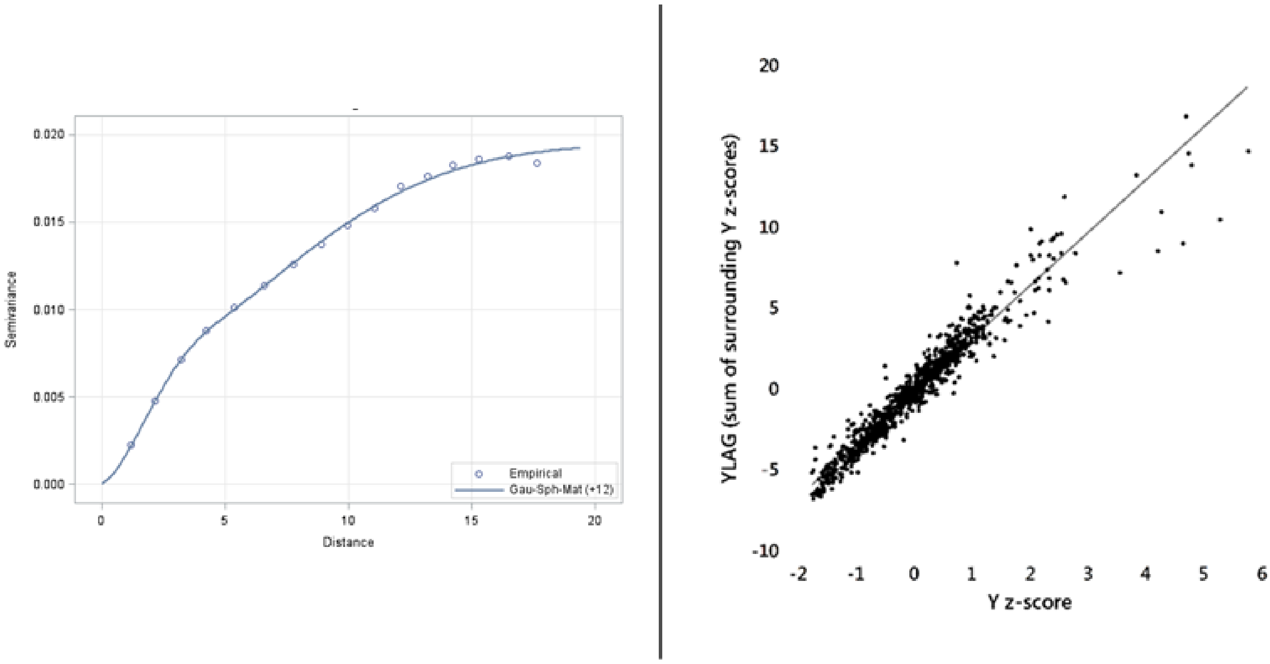

Figure 1. Left (a): Empirical semivariogram plot with a superimposed nonlinear trend line. Right (b): Moran scatterplot with a superimposed linear trend line.

A second class of scientific visualization exploits geographic maps, and is built on kriging and spatially varying coefficients. This former cartographic tool utilizes the redundant information represented by SA to interpolate a geographic distribution from locational information for a subset of locations (Cressie, 1991). In its simplest form, it imputes a few missing values in a geographic distribution. This latter cartographic tool utilizes SA to portray a map pattern of geographic dependencies. Spatially varying coefficients (SVCs) may be calculated with MESF, as statistical random effects (see Gelfand et al., 2003), or as other types of quantities such as geographically weighted regression (GWR) coefficients.

The focus shift of these latter procedures is to spatial heterogeneity—what Goodchild labels a second law of geography. GWR addresses this data feature by specifying smooth geographic variation in local, rather than constant global relationships between a response and covariate variables across a geographic landscape. GWR is a popular methodology in part because an ArcMap and other user-friendly software implementations of it exist. MESF can address this very same data feature, but solely through SA contained in georeferenced data. Its resulting relationship maps often portray fragmented, rather than smooth, map patterns. SVC specifications go beyond MESF by including both a spatially structured (i.e., ESF) and a spatially unstructured component, with the combination of these two constituting a random effects term in a model specification.

Finally, another prominent current feature of SA is its extension to non-normal random variables (Griffith, 2010). This extension has been achieved with MCMC techniques, as well as with mathematical statistical treatments.

4. Spatial Autocorrelation's Future

Getis (2007) proposes that the future of SA is in terms of sophisticated user-friendly software, and the dissemination of the concept to, and its diffusion through, other disciplines. SA's future also holds conceptual developments. One of the most neglected topics in spatial statistics/econometrics is negative SA, which is ascribed to not only artefacts of surface partitionings, but also situations involving spatial competition. Previous results obtained with auto-normal models suggest that qualitative differences may exist between impacts from positive and negative SA. Model estimation, testing, and calculation of impact effects may differ between positive and negative SA situations, resulting in differences in interpretation, inference, and implications. More contemporary results indicate that geographical landscapes displaying near-zero SA actually contain mixtures of positive and negative SA. These mixture specifications may well become a popular description of georeferenced data, particularly when empirical landscapes involve both competitive and cooperative processes. This future is a far cry from negative SA simply indexing improper surface partitionings.

Another emerging trend is the interface of spatial and temporal autocorrelation. In some contexts, temporal autocorrelation dominates SA in most space-time datasets. Parmentier et al. (2017) show specific catastrophic geographic situations in which SA emerges as being dominant over temporal autocorrelation, but only for a short period of time. In other contexts, spatial analyses are being extended to space-time domains (e.g., Griffith, 2012b). SA offers many future opportunities to better understand it within the context of space-time data.

Yet another future theme is the interface between SA and spatial optimization. For example, location-allocation problems involve determining locations in discrete or continuous space, either a single location in its simplest form, or multiple locations in its more complicated form. Generally speaking, the desired solution is a set of locations that minimizes the sum of (weighted) costs, which is formulated as an objective function. The p-median problem is widely utilized as one of the basic location-allocation problems. It is classified as NP-hard, and finding an optimal solution for it is fraught with computational challenges, even when feasible solutions exist. Hence, heuristic methods to quickly find optimal or near-optimal solutions often are utilized. Meta-heuristics provide a general framework to strategically design heuristics to achieve an improved solution and computational efficiency. SA can contribute to finding an optimal or near optimal solution for these location-allocation problems by coming into play in two distinct ways: (1) it can be viewed as duplicate information latent in the weights used in the minimization process, and hence exploited when missing weight values occur in order to impute them (Griffith, 1997); and, (2) it can help inform the determination of an optimal solution (Griffith and Chun, 2015). One focus of this anticipated research is its ability to uncover tendencies for optimal locations to coincide with local SA hot/cold spots. Another is for SA to be helpful in solving other spatial optimization problems, such as the maximal covering location problem.

As a scientific topic, SA is here to stay. It merits considerable additional study in order for spatial scientists to achieve a better, more comprehensive, and more parsimonious understanding of it. It deserves to be promulgated. Its research agenda almost certainly contains investigations of positive and negative SA mixtures, of its relationships with spatial optimization, and of its interface with other forms of autocorrelation.

References

Anselin, L. (1988). Spatial econometrics: Methods and models. Dordrecht: Kluwer Academic Publishers.

Anselin, L. (1995). Local Indicators of Spatial Association -- LISA. Geographical Analysis, 27(2):93-115.

Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems (with discussions). Journal of the Royal Statistical Society: Series B, 36: 192–236.

Cliff, A., and Ord, J. (1973). Spatial Autocorrelation. London: Pion.

Cliff, A., and Ord, J.(2009). What were we thinking? Geographical Analysis, 41: 351-363.

Cressie, N. (1991). Statistics for Spatial Data. Wiley Series in Probability and Statistics. New York, NY: Wiley.

Elhorst, P. (2014). Spatial Econometrics: From Cross-Sectional Data to Spatial Panels. Berlin: Springer.

Geary, R. C. (1954). The Contiguity Ratio and Statistical Mapping. The Incorporated Statistician, 5(3), 115-146.

Gelfand, A. E., Kim, H. J., Sirmans, C. F., & Banerjee, S. (2003). Spatial modeling with spatially varying coefficient processes. Journal of the American Statistical Association, 98(462), 387-396.

Getis, A. (2007). Reflections on spatial autocorrelation. Regional Science and Urban Economics 37(4): 491-496.

Getis, A. (2010). Spatial autocorrelation. In M. Fischer and A. Getis (eds.), Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications, pp. 255-278. Berlin: Springer-Verlag.

Getis, A., & Ord, J.K. (1992). The analysis of spatial association by use of distance statistics. Geographical Analysis, 24(3), 189-206.

Goodchild, M. F. (1986). Spatial Autocorrelation. CATMOD #47. Norwich, UK: GeoBooks.

Griffith, D. A. (1987). Spatial Autocorrelation. Washington, DC: Association of American Geographers.

Griffith, D. A. (1992). What is Spatial Autocorrelation? Reflections on the Past 25 Years of Spatial Statistics. L'Espace Géographique, 21: 265–280.

Griffith, D. A. (1997). Using estimated missing spatial data in obtaining single facility location-allocation solutions. L'Espace Géographique, 26: 173-182.

Griffith, D. A. (2005). Spatial autocorrelation. In K. Kempf-Leonard (ed.), Encyclopedia of

Social Measurement, Vol. 3. San Diego, CA: Academic Press (Elsevier), pp. 581-590.

Griffith, D. A. (2009). Methods: spatial autocorrelation. In R. Kitchin and N. Thrift (eds.), International Encyclopedia of Human Geography. NY: Elsevier, pp. 396-402.

Griffith, D. A. (2010). The Moran Coefficient for non-normal data. Journal of Statistical Planning and Inference, 140: 2980-2990.

Griffith, D. A. (2012a). Spatial statistics: a quantitative geographer's perspective. Spatial Statistics, 1: 3–15.

Griffith, D. A. (2012b). Space, time, and space-time eigenvector filter specifications that account for autocorrelation. Estadística Española, 54 (#177), 7-34.

Griffith, D. A. and Chun, Y. (2014). Spatial autocorrelation and spatial filtering. In M. Fischer and P. Nijkamp (eds.), Handbook of Regional Science. Berlin: Springer-Verlag, pp. 1477-1507.

Griffith, D. A. and Getis, A. (2016). Spatial filtering. In Shekhar, S., H. Xiong, and X. Zhou, Encyclopedia of GIS (2nd ed.). Cham, Switzerland: Springer, pp. 1-14.

Griffith, D. A., and Chun, Y. (2016). Evaluating eigenvector spatial filter corrections for omitted georeferenced variables. Econometrics, 4: 29.

LeSage, J. P. and Pace, R. K. (2009). Introduction to Spatial Econometrics. Chapman and Hall/CRC Press.

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. Biometrika, 37: 17–23.

Odland, J. (1988). Spatial Autocorrelation. Thousand Oaks, CA: SAGE Publishers.

Paelinck, J. H. P., and Klaassen, L.H. (1979). Spatial Econometrics. Saxon House Farnborough.

Páez, A., and Whalen, K. (2010). Enjoyment of commute: a comparison of different transportation modes. Transportation Research Part A: Policy and Practice, 44: 537–549.

Parmentier, B., Millones, M., Griffith, D., Hamilton, S., Chun, Y. and McFall, S. (2017). When space beats time: a proof of concept with Hurricane Dean. In D. Griffith, Y. Chun, and D. Dean (eds.), Advances in Geocomputation: Geocomputation 2015—The 13th International Conference. Berlin: Springer, pp. 207-215.

Whittle, P. (1954). On Stationary Processes in the Plane. Biometrika, 41 (3/4), 434-449.