한국지도학회지 12권 2호 2012(37~46)

마할라노비스 거리를 이용한 다변량 공간 클러스터 분석

이몽현*

Multivariate Spatial Cluster Analysis Using Mahalanobis Distance

Monghyeon Lee*

요약 : 본 연구는 로컬 단위에서의 다변량 공간적 클러스터와 아웃라이어에 대한 분석에 대하여 논한다. 공간적 클러스터나 아웃라이어는 그 접근 방법이나 쓰임에 따라 다양한 정의를 내릴 수 있으나, 공간적 연관성을 기반으로 한다는 점에서는 근본적인 공통점이 있다. 그러나 현존하는 공간 연관성 척도들은 투입할 수 있는 변수의 수가 한정적이기 때문에 다변량 상황에서 공간적 연관성을 측정할 수 있는 방법에 대한 연구가 필요하다. 다변량 local 공간 연관성 척도의 개발을 위해, 본 연구에서는 두 집단간 분리 정도에 대한 측정이 가능한 마할라노비스 거리를 이용하였다. 마할라노비스 거리는 변수의 평균, 분산 그리고 변수간 공분산을 고려하여 계산이 된다. 본 연구에서 고안된 로컬 마할라노비스 거리는 해당 지역의 변수 벡터와 주변지역 변수의 평균 값 벡터를 통해 계산이 되며, 이를 수도권 지역 동읍면 단위에서 인구 전입/전출의 변수에 대해 적용하였다. 해당 단위에서의 공간적 변동은 카이제곱 *p*값 지도를 통해 확인할 수 있으며, 유의성 검정을 실시한 로컬 마할라노비스 거리 지도를 통해 인구 유출입 차원에서의 공간적 클러스터와 아웃라이어를 확인할 수 있다.

주요어 : 마할라노비스 거리, 공간 클러스터, 다변량 로컬 공간 연관성 척도

Abstract : This paper introduces an approach for analyzing multivariate spatial cluster/outlier in local scale. Even though spatial cluster/outlier has various definitions, the fundamental of spatial cluster/outlier is based on spatial association. Existing methods for measuring local spatial association had a limitation of applying multiple numbers of variables. Univariate local spatial association measures such as local Moran's I_i, local Geary's C_i and Getis and Ord's G_i^* are widely used, and bivariate local spatial association measures are already developed; Cross Moran and Lee's L_i. However, the measures are not used for measuring spatial association among three or more variables. This is a critical limitation when spatial variation with the complex multi-dimensional approaches is explained and described. The measure in this paper, multivariate local spatial association measure, is based on Mahalanobis Distance (MD) and it enables distinguishing spatial similarities and differences among multiple numbers of data sets simultaneously. MD considers variables' means, variances and co-variances and allows measuring the variables' distribution. It is the same concept as distance measuring with Euclidean Distance but improved. Significance of MD could be tested because it is following chi-square distribution when the variables are multi-normal. Local MD is applied to demographic variables, in- and out-migration in Seoul Metropolitan Area. The spatial variation of multivariables could be identified by chi-squared *p*-value map, and a local MD map is provided to show the detected spatial clusters or outliers at a given significance level.

Key Words : Mahalanobis Distance, Spatial cluster, Multivariate local spatial association measure

*텍사스대학교(댈러스) 박사과정(PhD Student, School of Economic, Political and Policy Sciences, University of Texas at Dallas, mong5523@naver.com)

I. Introduction

Spatial clusters are aggregations, real or perceived, of characteristics that are grouped together in space (Demattei *et al*., 2006). Spatial cluster has been one of the most important tasks for geographic knowledge discovery (Guo *et al*., 2002), required when spatial homogeneity/heterogeneity phenomena is modeled and provides a perspective of ESDA (exploratory spatial data anlaysis) (Jacquez, 2008). In modeling, spatial clusters help understanding patterns of spatial variation, and in the context of ESDA, they help identifying and describing spatial patterns (recognizing the pattern by visualization, describing the statistical significant for a location, scale and pattern in spatial statistics).

However, most existing methods, especially Spatial Association Measures, are only used for one or two geographical variables to detect spatial clusters even if geographical clusters are not explained only one or two variables. Here I describe a spatial association measure detecting multivariate local spatial cluster/outlier with Mahalanobis Distance (MD), implementation of the model into a commercial GIS package, and conduct the application to demographic variables.

Following sections provide a context with a review of local spatial statistics and then provides the background of multivariate statistics technique, MD. The process of the modification to implement spatial data sets into MD equation is followed by. Test of the modified equation and application to demographic variables are introduced after the modification process is explained. Finally, a summary of results and discussion of a research direction, limitation and conclusion of this paper are provided.

II. Review

For last several decades, many methods were developed to detect spatial clusters. Ripley (1977) developed point pattern analysis K-function that counts the number of points in specific distance, Kelsall and Diggle (1995) suggested Kernel intensity function to calculate the probability of expected events in a region using kernel function and Kulldorff (1997) developed SaTScan (Spatial scan statistics) to confirm a spatial group of high values using a pre-defined moving window.

Spatial Association Measure (SAM) is the most frequently one used for detecting spatial clusters of areal units. Global measures such as Moran's I (Moran, 1948) and Geary's C (Geary, 1954) provide statistics for spatial association of areal units over the entire research area, but local measures such as local Moran's I_i, local Geary's C_i (Anselin, 1995) and Getis-Ord's G_i^* (Ord and Getis, 1995) allow researchers to explore spatial variations in local scale. However, methods fore-mentioned, SAMs, are for univarate spatial association measures. In other words, these methods focus on the spatial clustering of observations in terms of a single variable (Lee, 2001a). As a step forward, Wartenberg (1985) has presented a bivariate spatial association method called Multivariate Spatial Correlation Analysis, based on Moran's I, principal components analysis, and eigenvector analysis. Lee (2001a) developed bivariate spatial association measure L_i to capture the relationship between two variables mainly using Pearson's r.

However, a multivariate spatial association measure has not been developed, yet. Even though Wartenberg (1985) tried to develop a multivariate spatial association measure, and Hwang *et al.* (2008) tried to explain the correlation of marine environmental factors by pairing variables then categorizing the results derived from bivariate spatial association measure Lee's L_i. It is impossible to use existing methods with three or more variables to measure spatial association when considering multidimensional approach spatially. For instance, analyzing demographic characteristics requires involvements of many variables such as fertility rate, mortality rate, composition of population, in-migration rate, out-migration rate and others simultaneously because one or two variables may not provide sufficient explanation of the complex relationship among those variables.

Ⅲ. Mahalanobis Distance (MD)

MD is a dimensionless measure of the distance in multivariate space from a point to another point and applicable to non-normally distributed variables (Knick and Rotenberry, 1998). In addition, it is a powerful method for distinguishing differences among groups with multivariate data set (Clark *et al.*, 1993; De Maesschalck *et al.*, 2000; Woodall *et al.*, 2003).

The method defines a degree of separation from groups containing multivariate data that have several correlated continuous variables and a pre-defined group variable. It also contains all appropriate standard specifications necessary for classifying objects into general groups, which is a structure analysis of distinction variables and statistical techniques related to typology (De Maesschalck *et al.*, 2000).

Moreover, MD is based on both the mean and variance of a variable and the covariance matrix of all variables, therefore it takes advantage of the covariance among variables. The region with a constant MD around the mean forms an ellipse in two-dimensional space (when only 2 variables are measured) or an ellipsoid or hyper ellipsoid when more variables are used.

Figure 1 is a comparison of Euclidian Distance (ED) and MD in two dimensional spaces. Point A and B have the same distance from the mean (μ) because MD considers variable's distribution. Therefore, the distribution of points of the same distance from the mean forms a circle in ED but it forms an ellipse in MD.

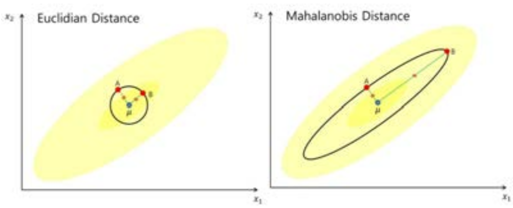


Fig. 1. Comparison of ED and MD

MD is calculated as

$$\mathrm{MD}_{ij} = \sqrt{(x_i - x_j)^T \Sigma^{-1} (x_i - x_j)},$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{1,2}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{2,n}\sigma_2\sigma_n \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{1,n}\sigma_1\sigma_n & \rho_{2,n}\sigma_2\sigma_n & \cdots & \sigma_n^2 \end{bmatrix}$$

where MD_{ij} is MD for object i to *j*, x_i is a variable vector for object *i*, x_j is a variables vector for object *j*, σ_n is a variance of "*n*" variable and $\rho_{n,m}\sigma_n\sigma_m$ is a co-variance of two variables *(n, m)*. MD's capability to handle multivariate data set provides broad usage extent of this method. In the field of multivariate calibration, MD used for different purposes, namely: for the detection of outliers (Lu *et al.*, 2003), the selection of calibration samples for, a large set of measurements (Shenk and Westerhaus, 1991), the wild animal habitat analysis (Clark *et al.*, 1993; Knick and Dyer, 1997; Knick and Rotenberry, 1998), the finding abnormal patient in medical (Woodall *et al.*, 2003) and chemistry (De Maesschalck *et al.*, 2000).

MD is occasionally converted to Chi-square *p*-values for analysis (Clark *et al.*, 1993; Knick and Dyer, 1997). If the variables in a data set are normally distributed, MD follows the Chi-square distribution with n-1 degrees of freedom (Krzanowski and Lai, 1988; Clark *et al.*, 1993; Penny, 1996; Knick and Dyer, 1997).

$$(x_i - x_j)^T \Sigma^{-1} (x_i - x_j) \sim \chi_p^2$$

Penny (1996) suggested an *F*-distribution is more appropriate than a Chi-square distribution when dealing with small sample sizes. The given equation is

$$(x_i - x_j)^T \Sigma^{-1} (x_i - x_j) \sim \frac{p(n-1)^2 F_{p,n-p-1}}{n(n-p-1+pF_{p,n-p-1})}$$

with the *n* observations in a *p*-variable data set.

Even though MD is a powerful method for

multivariate data analysis and frequently used for geographic models of wild animal habitat, it has not been considered as a measure of multivariate spatial association. Following chapter describes the converting process of multivariate data analysis, MD, to a local spatial association measure for detecting a multivariate local spatial homogeneity/heterogeneity.

Ⅳ. Development of an MD-based Cluster Analysis Method

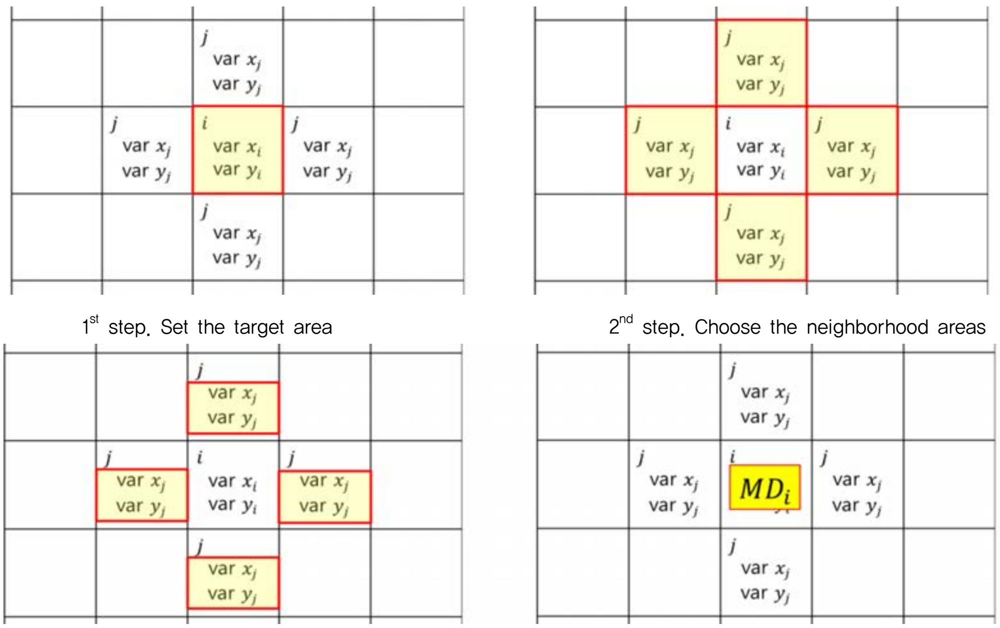
1. Modification of MD to multivariate local SAM

As with the existing Spatial Association Measures, MD, the multivariate local spatial association measure is calculated in mainly two parts. The first is identifying neighborhoods of location i, and the second is calculating MD. The process of the MD calculation is: first, set the target area (i) and this identification process applies to each location in an entire study area. Second, choose the neighborhood areas (j). Third, calculate the mean ($\bar{j}$) of each variables among defined neighborhood areas. Fourth, calculate the MD between the target area and neighborhood areas. There are a few methods for defining neighborhoods, and rook's contiguity method is used in this paper. Figure 2 shows the stepwise calculation of multivariate spatial association measure for two variables.

The written form of multivariate spatial association measure is

$$\mathrm{MD}_{ij} = \sqrt{\left(x_i - \frac{1}{n}\sum_j w_{ij}\, x_j\right)^T \mathbf{C}^{-1}\left(x_i - \frac{1}{n}\sum_j w_{ij}\, x_j\right)}$$

where x_i is a matrix of variables at location i, x_j is a matrix of variables at location j, n is the number of neighbors of location i, w_{ij} is a binary spatial weight



1st step. Set the target area

2nd step. Choose the neighborhood areas

3rd step. Calculate means of each variables among neighborhood area

4th step. Calculate MD between the target area and neighborhood areas

Fig. 2. Steps for calculating multivariate spatial association measure with two variables

matrix, and **C** is variance/co-variance matrix for every variable.

2. Demonstration of the Modified MD

In order to examine the spatial cluster and spatial outlier detection process of the multivariate spatial association measure with MD, two similar test patterns are tested in the model. The procedure is demonstrated with 400 units of square tessellations and each unit has two normally distributed variable values with the lower bound of approximately -0.1 and upper bound of 0.1.

These two test patterns are used in the demonstration because the patterns have clusters and outliers at the analogous locations. Homogeneous values in this pattern are in the edges of upper left and lower right, and other eight-parts represent the places of heterogeneous value. Figure 4 is the result of multivariate spatial association measure with MD for the test patterns.

The lowest value of local MD model is 0.001498 and the values from 0.001498 to 0.015 are detected as clusters which are significance at α = 0.1 level. The highest value is 26.249531 and the values from 0.015001 to 26.249531 are detected as outliers with the significance at α = 0.1 level. An application using census data is in the following section.

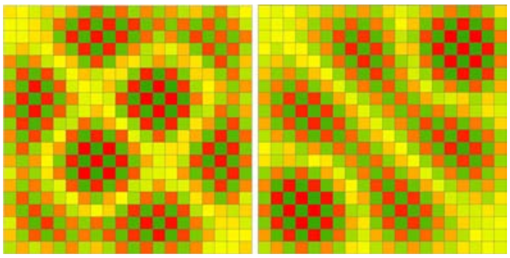


Fig. 3. Test pattern input variables

Ⅴ. Applications

1. Consideration of Variables

In this section, the applications of the multivariate local spatial association measure with MD are demonstrated using census data. The study area is the Seoul Metropolitan Area and the census data were extracted from the 2010 population and housing census from Statistics Korea. 1093 *dong-eup-myeon* level areal unit are analyzed for the Seoul Metropolitan Area. Two variables, IMR (in-migration rate: the number of in-migrants divided by the number of population) and OMR (out-migration rate: the number of out-migrants divided by the number of population), were the inputs for detecting multivariate local spatial clusters/outliers.

The most intuitive result would be the similar in- and out-migration relationship as depicted in figure 5 (Plane and Rogerson, 1994). The regions with the highest rate of in-migration have relatively lower rate of out-migration. The regions with high in-migration have

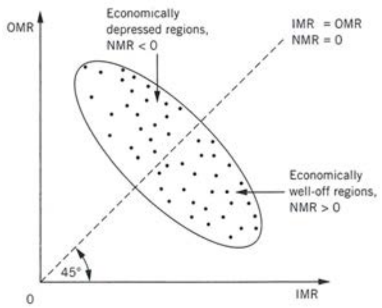


Fig. 4. The intuitive relationship between in– and out–migration rates

Source : Plane and Rogerson, 1994

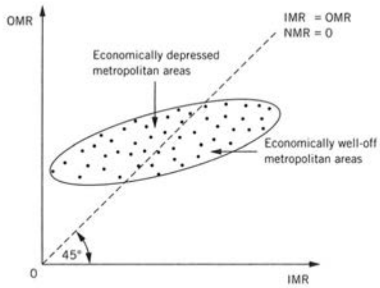


Fig. 5. The positive correlation perspective of relationship between in– and out– migration

Source : Plane and Rogerson, 1994

attractive economic conditions and provide a good environment for living; in such places, there is a little factor for people to leave. On the other hand, low in-migration rate region may refer having poor condition and environment for living causing high propensity to out-migration.

However, in reality, dynamic factors and complex relations influencing migration lead to a different observation result. Some relationships between two variables have been found contradictive in actual cases (Plane and Rogerson, 1994). Plane (1981) analyzed a correlation (Pearson's r) of 1975-1976, 1976-1977 and 1978-1979 U.S. inter-state migration that was 0.93, 0.91 and 0.84 respectively. The correlation coefficient is described in figure 6.

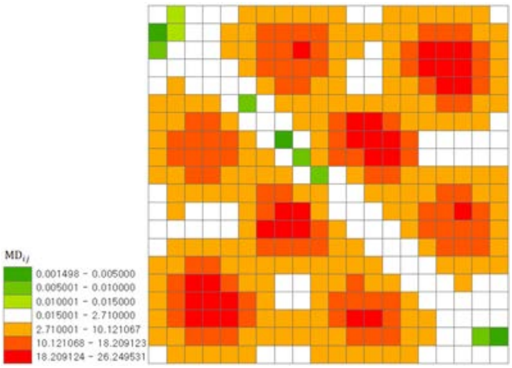


Fig. 6. Result of the test pattern

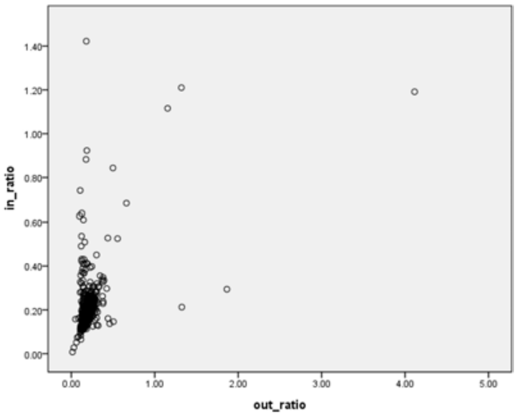


Fig. 7. The relationship between in- and out-migration in of the Seoul Metropolitan Area

Figure 7 represents the relationship between in- and out-migration of the Seoul Metropolitan Area in 2010. It clearly shows a positive correlation.

In order to explain spatial migration reasonably, it requires fundamental geographical inquiries (Lee, 2001c). It is the aggregation of spatially dependent events based on origins and the distances of origins and destinations. Therefore, values for in- and outmigration rate are interrelated to each other for every region. In other words, spatial migration is a very spatially dependent phenomenon. In addition, a measure of the spatial multivariate dependency also provides a perspective on Exploratory multivariate Spatial Data Analysis which has not been developed, yet.

2. Local MD of the Seoul Metropolitan Area

Figure 8 and 9 are input variables, in- and out-migration rate in the Seoul Metropolitan Area, for calculating MD_{ij}.

Table 1 shows the average out-migration has higher rate than in- migration rate but the value of the measurement univariate spatial association, Moran's I, are approximately the same. The result of global Moran's I value for each variable does not represent the global spatial association of the variables, but local Moran's I identifies univariate spatial cluster/outlier. Local Moran's I of in- and out- migration rates in Seoul are shown in figure 10 and 11.

Chi-square p-value of MD for each region in the Seoul Metropolitan Area is provided in figure 12. Chi-square p-value map shows multivariate spatial variation. Some regions in Seoul have low values of local MD in principle but the regions just outside of Seoul have high values and it means that they are spatially heterogeneous. Specifically, regions in Seoul, Seo-gu Incheon, Gimpo-si, Goyang-si, Paju-si, Gwangmyeong-si, Namyangju-si, Suwon-si, Hwaseong-si and Pocheon-si are heterogeneous regions and other regions in Incheon and Gyeonggi-do are homogeneous.

When everything is considered, all rural regions have low local MD referring that these regions have low spatial variation and most urban areas have high values and this observation proves that these regions have high spatial variations.

Figure 13 shows local MD distribution in the Seoul Metropolitan Area at α=0.1 significance level. This result provides a multivariate spatial cluster or outlier with specific significance level. In this result, similar as the Chi-squared p-value map, spatial clusters are located sparsely at suburb of Gyeonggi-do, inner Seoul and Incheon, besides a large spatial outlier is located at south of Seoul.

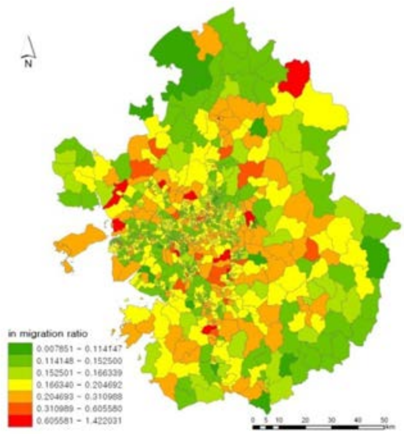


Fig. 8. In-migration rate in the Seoul Metropolitan Area

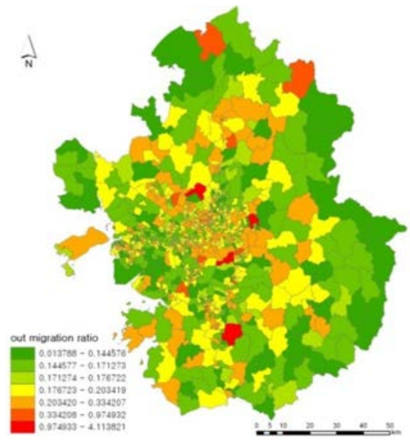


Fig. 9. Out-migration rate in the Seoul Metropolitan Area

Table 1. Statistics of in- and out-migration variables in the Seoul Metropolitan Area

SMA	In-migration rate	Out-migration rate
Mean	0.1889	0.1962
Standard Deviation	0.0968	0.1491
Global Moran's I	0.1148	0.0716

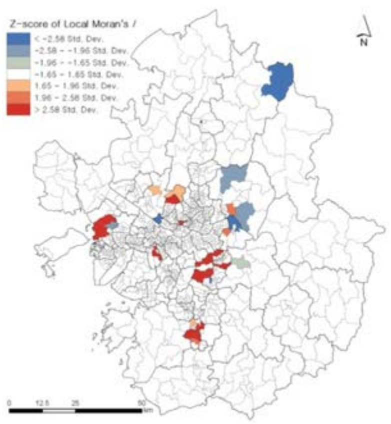


Fig. 10. In-migration rate in the Seoul Metropolitan Area

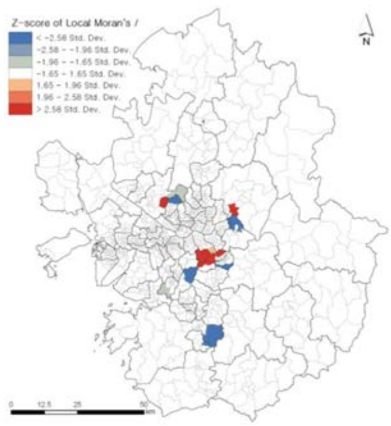


Fig. 11. Out-migration rate in the Seoul Metropolitan Area

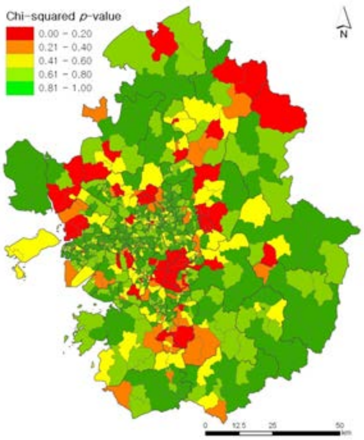


Fig. 12. In- migration rate in the Seoul Metropolitan Area

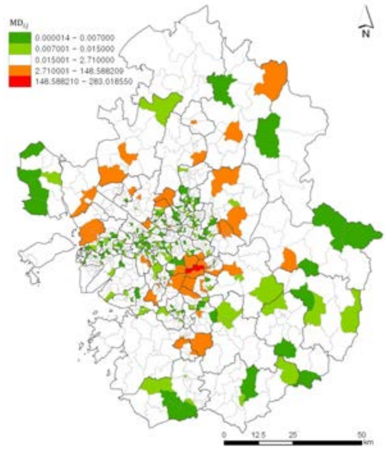


Fig. 13. Out- migration rate in the Seoul Metropolitan Area

VI. Conclusion

This paper developed a local multivariate spatial association measure (MD_{ij}) based on MD and demonstrated the usefulness of the measure with the test pattern and demographic variables, in- and out-migration of the Seoul Metropolitan Area in 2010. The measure provides a simple value for identifying spatial homogeneity/heterogeneity in complex multi-variate situation. Although only two geographical variables were used in this paper, the potential to expand the number of variables to three variables has revealed.

In the application, spatial variation identified the multiple spatial migration characteristics of each region through Chi-square *p*-value maps and spatial clusters and outliers were detected using local MD's significance test. Even though interpretations for each region maybe complicated, local MD provides a simple perspective for ESDA and also the usages could be utilized and transformed in accord with objective of researchers. For instance, the target region's MD could be calculated with global mean of variables, instead of using mean of neighbors (local MD), to detect global outliers. Another utilization is comparing two different (distant) regions using the MD's capability to consider variables of the two regions simultaneously.

There are some limitations in this research. Penny (1996) pointed out that if variables used to calculate MD have multi-collinearity, then the distance should be tested with *F*-test. Then, in order to eliminate multi-collinearity of variables, Principal Component Analysis should be applied ahead. Variable's multi-normality should also be assumed for using Chi-square significance test and alternative spatial weight matrix could be useful in this regard.

References

Aldstadt, J., 2010, Spatial clustering, in Fischer, M. and Getis, A. eds., *Handbook of Applied Spatial Analysis*, Berlin: Springer, 279-300.

Anselin, L., 1995, Local indicators of spatial association-LISA, *Geographical Analysis*, 27, 93-115.

Anselin, L., 1996, The Moran scatterplot as an ESDA tool to assess local intability in spatial association, in Fischer, M., Scholten, H., and Unwin, D. eds., *Spatial Analytical Perspectives in GIS*, London: Taylor & Francis, 111-125.

Anselin, L., Syabri, I., and Smirov, O., 2002, Visualizing multivariate spatial correlation with dynamically linked windows, in *New Tools for Spatial Data Analysis: Proceedings of the Specialist Meeting*, 51.

Clark, J.D., Dunn, J.E., and Smith, K.G., 1993, A multivariate model of female black bear habitat use for a geographic information system, *The Journal of Wildlife Management*, 57(3), 519-526.

Cliff, A.D. and Ord, J.K., 1973, *Spatial Autocorrelation*, London: Pion.

Demattei, C., Molinari, N., and Daures, J., 2006, Arbitrarily shaped multiple spatial cluster detection for case event data, *Computational Statistics & Data Analysis*, 51(8), 3931-3945.

De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D., 2000, The MD, *Chemometrics and Intelligent Laboratory Systems*, 50(1), 1-18.

Geary, R.C., 1954, The contiguity ratio and statistical mapping, *Incorporated Statistician*, 5, 115-145.

Getis, A. and Ord, J., 1992, The analysis of spatial association by distance statistics, *Geographical Analysis*, 24, 189-206.

Ord, J. and Getis, A., 1995, Local spatial autocorrelation statistics: distributional issues and an application, *Geographical Analysis*, 27(4), 286-306.

Guo, D., Peuquet, D., and Gahegan, M., 2002, Opening the black box: interactive hierarchical clustering for multivariate spatial patterns, *Association for Computing Machinery: the 10th ACM international symposium on Advances in geographic information systems*, 131-136.

Hwang, H.J., Choi, H.W., and Kim, T.R., 2008, Application of bivariate spatial association for the quantitative marine environment pattern analysis, *Journal of the Korean Association of Geographic Information Studies*, 11(1), 155-166 (in Korean).

Jacquez, G., 2008, Spatial cluster analysis, in Wilson, J. and Fortheringham, S. eds., *The Handbook of Geographical Information Science*, Malden, MA: Blackwell Publishing, 395-416.

Kelsall, J.E. and Diggle, P.J., 1995, Non-parametric estimation of spatial variation in relative risk, *Statistics in Medicine*, 14, 2335-2342.

Knick, S. and Dyer, D.L., 1997, Distribution of black-tailed jackrabbit habitat determined by GIS in southwestern Idaho, *The Journal of Wildlife Management*, 61(1), 75-85.

Knick, S. and Rotenberry, J., 1998, Limitations to mapping habitat use areas in changing landscapes using the MD statistic, *Journal of Agricultural, Biological, and Environmental Statistics*, 3, 311-322.

Krzanowski, W. and Lai, Y., 1988, A criterion for determining the number of groups in a data set using sum-of-squares clustering, *Biometrics*, 44, 23-34.

Kulldorff, M., 1997, A spatial scan statistic, *Communications in Statistics- Theory and Methods*, 26(6), 1481–1496.

Lee, S.-I., 2001a, Developing a bivariate spatial association measure: an integration of Pearson's r and Moran's *I*, *Journal of Geographical Systems*, 3, 36-385.

Lee, S.-I., 2001b, Spatial Association Measures for an ESDA-GIS Framework: Developments, Significance Tests, and Applications to Spatio-Temporal Income Dynamics of U.S. Labor Market Areas, 1969-1999, Ph.D. Dissertation, The Ohio State University.

Lee, S.-I., 2001c, A spatial statistical approach to migration studies, *Journal of the Korean Association of Regional Geographers*, 7(3), 107-120.

Lee, S.-I., Cho, D.H., Son, H.K., and Chae, M.O., 2010, A GIS-based method for delineating spatial clusters: a modified AMOEBA technique, *Journal of the Korean Geographical Society*, 45(4), 502-520 (in Korean).

Lu, C.T., Chen, D., and Kou, Y., 2003, Detecting spatial outliers with multiple attributes, *IEEE International Conference on Tools with Artificial Intelligence*, 15, 122-128.

Miller, H., 2004, Tobler's first law and spatial analysis, *Annals of the Association of American Geographers*, 94(2), 284-289.

Moran, P.A.P., 1948, The interpretation of statistical maps, *Journal of the Royal Statistical Society Series B*, 10, 243-251.

Penny, K.I., 1996, Appropriate critical values when testing for a single multivariate outlier by using the MD, *Journal of the Royal Statistical Society Series C*, 45(1), 73-81.

Plane, D., 1981, Estimation of place-to-place migration flows from net migration totals: a minimum information approach, *International Regional Science Review*, 6, 33-51.

Plane, D. and Rogerson, P., 1994, *The Geographical Analysis of Population: With Applications to Planning and Business*, NY: John Wiley & Sons.

Ripley, B.D., 1977, Modeling spatial patterns (with discussion), *Journal of the Royal Statistical Society Series B*, 39, 172-212.

Shenk, J. and Westerhaus, M., 1991, Population definition, sample selection, and calibration procedures for near-ingrared reflectance spectroscope, *Crop Science*, 31, 469-474.

Szalay, A.S., Budavári, T., Connolly, A., Gray, J., Matsubara, T., Pope, A., and Szapudi, I., 2002, Spatial clustering of galaxies in large datasets, *SIPE Astronomy Telescopes and Instruments*, 4847(1), 22-28.

Vinson, T. and Baldry, E., 1999, The spatial clustering of child maltreatment: are micro-social environments involved?, *Australian Institute of Criminology*, 119, 1-6.

Wartenberg, D., 1985, Multivariate spatial correlation: a method for exploratory geographical analysis, *Geographical Analysis*, 17, 263-283.

Wheeler, D., 2007, A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996-2003, *International Journal of Health Geographics*, 6(13), 1-16.

Woodall, W.G., Koudelik, R., Tsui, K., Kim, S.B., Stoumbos, Z.G., and Carvounis, C.P., 2003, A review and analysis of the Mahalanobis-Taguchi system, *Technometrics*, 45(1), 1-15.

Correspondence : Monghyeon Lee, 800 W. Campbell Rd, Richardson, Texas 75080-3021, USA, School of Economic, Political and Policy Sciences, University of Texas at Dallas (Email: mong5523@naver.com)

교신 : 이몽현, 800 W. Campbell Rd, Richardson, Texas 75080-3021, USA, School of Economic, Political and Policy Sciences, 텍사스대학교(댈러스) (이메일: mong5523@naver.com)

Received: August 2, 2012
Revised: August 12, 2012
Accepted: August 14, 2012