



Mirosław Krzyśko

The President Stanisław Wojciechowski State University of Applied Sciences in Kalisz,
Interfaculty Institute of Mathematics and Statistics, mkrzysko@amu.edu.pl

Wojciech Łukaszonek

The President Stanisław Wojciechowski State University of Applied Sciences in Kalisz,
Interfaculty Institute of Mathematics and Statistics, w.lukaszonek@g.pl

Waldemar Ratajczak

Adam Mickiewicz University in Poznań, Faculty of Geographical and Geological Sciences,
Institute of Socio-Economic Geography and Spatial Management, walrat@amu.edu.pl

Waldemar Wołyński

Adam Mickiewicz University in Poznań, Faculty of Mathematics and Computer Science,
Department of Probability Theory and Mathematical Statistics, wolynski@amu.edu.pl

Nonlinear Principal Component Analysis for Geographically Weighted Temporal-spatial Data

Abstract: Schölkopf, Smola and Müller (1998) have proposed a nonlinear principal component analysis (NPCA) for fixed vector data. In this paper, we propose an extension of the aforementioned analysis to temporal-spatial data and weighted temporal-spatial data. To illustrate the proposed theory, data describing the condition of state of higher education in 16 Polish voivodships in the years 2002–2016 are used.

Keywords: nonlinear principal component analysis, geographically weighted data, temporal-spatial data

JEL: C02, C31, C38

1. Introduction

Since the early 1970s the attention of statisticians, econometricians, geographers, and spatial economists has focused on features of geographical (spatial) data. The complexity of socio-economic space (S-ES), among others, is characterised by its heterogeneity and spatial dependence¹.

The features of S-ES have a great impact on the quality of data used in different analyses concerning important aspects of human activity. Heterogeneity and spatial dependence play a special role in cases in which statistical or econometrical models are built.

The data used in such cases have a spatial character because numerical values of variables depend on the location of object i characterised by a set of these variables.

For example, spatiality of the variables is implied by the fact that the other objects located around object i have some influence on the data collected at and for object i . This kind of interaction is well described by Tobler's first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970: 236).

Therefore, following Demšar et al. (2013: 107), we would like to state that: "Spatial data contain geographic as well as attribute information. Thus, whereas typical data sets only contain measurements of variables or attributes, spatial data sets are characterized by having a location associated with each measurement; that is, the geographic location within the basic three-dimensional framework of our physical world, where measurement was taken".

It has become clear that spatial data, used to build classical regression models, undoubtedly have an impact on the quality of parameter estimates, which in such case are biased and inefficient. Many papers have been published concerning these problems, for example: Cliff and Ord (1973), Casetti (1972), Anselin (1988), Swamy (1971).

Problems related to spatial data can also be solved with the use of other multivariate statistical methods. Indeed, Principal Components Analysis (PCA) was the first method taken into account by researchers. First Charlton et al. (2010) and then Demšar et al. (2013) proposed modifications of the classical PCA procedure by introducing a weight matrix to the PCA equation. Weighting is based on some kind of kernel.

Suppose that each object is characterised by $T \times p$ -size matrix X containing the p -values observed at T moments. This type of data is defined as temporal-spatial data. One of the possibilities is to transform matrix data into functional data

¹ Spatial heterogeneity refers to an uneven distribution of attribute, occurrence and relationship across a region. Spatial dependence alludes local non-independence of events that are near each other (Anselin, 2010).

and construct linear principal components for functional data (Górecki et al., 2018). However, this approach does not work in the case of geographically weighted temporal-spatial data. Hence, our goal is to construct nonlinear principal components for both temporal-spatial data as well as geographically weighted temporal-spatial data. The new methods proposed by us are an extension of Schölkopf, Smola and Müller (1998) method for fixed vector data. The temporal-spatial data is mapped nonlinearly into a Hilbert space and a centred kernel matrix is computed. Finally, nonlinear principal components are calculated by solving NPCA equations. In the case of geographically weighted temporal-spatial data, the observation matrix X_i is replaced by the $w_i X_i$ matrix, where w_i is the positive geographic weight associated with the i -th observation site, $i = 1, \dots, n$. The procedure just described is illustrated by the example in the following sections.

The paper is organised as follows. Section 2 presents the basic ideas of Hilbert’s space. Section 3 describes the structure of the nonlinear principal components in the case of unweighted and weighted data. Section 4 illustrates the approaches presented in the paper on a real data set.

2. Hilbert spaces

A Hilbert space, named after David Hilbert, is an extension of a vector space. Regular vector spaces are sets of objects that are closed under a linear combination. That is, given a vector space \mathcal{X} , we have $u, v \in \mathcal{X} \Rightarrow \alpha u + \beta v \in \mathcal{X}$. While one normally thinks of these objects as finite dimensional vectors, they could potentially be infinite dimensional vectors, and as such should be treated as functions. A Hilbert space is a complete real vector space equipped with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$, with the following properties:

1. Symmetry: $\langle f, g \rangle_{\mathcal{H}} = \langle g, f \rangle_{\mathcal{H}}$,
2. Linearity: $\langle \alpha f_1 + \beta f_2, g \rangle_{\mathcal{H}} = \alpha \langle f_1, g \rangle_{\mathcal{H}} + \beta \langle f_2, g \rangle_{\mathcal{H}}$, for all real numbers α and β ,
3. Non-negativity: $\langle f, f \rangle_{\mathcal{H}} \geq 0$,
4. Zero: $\langle f, f \rangle_{\mathcal{H}} = 0 \Rightarrow f = 0$.

An example of an inner product might be:

$$\langle f, g \rangle_{\mathcal{H}} = \int f(x)g(x)dx.$$

Given this basic definition of a Hilbert space, we can now define a fundamental concept that is an operator. A linear operator C maps function f in one Hilbert

space to another function g in the same or another Hilbert space. Mathematically, this corresponds to:

$$Cf = g.$$

This operator has the following property:

$$C(\alpha f + \beta g) = \alpha Cf + \beta Cg.$$

Intuitively, one can think of functions as vectors and operators as matrices. In linear algebraic terms, a matrix typically projects a set of vectors to another set of vectors. Therefore, the effect of the operator is to transform a function in a Hilbert space to another function in another Hilbert space.

A function

$$k : \mathbb{R}^{T \times p} \times \mathbb{R}^{T \times p} \rightarrow \mathbb{R}$$

is called a kernel on $\mathbb{R}^{T \times p}$ if there exists a Hilbert space \mathcal{H} with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and a map

$$\phi : \mathbb{R}^{T \times p} \rightarrow \mathcal{H}$$

such that for all $\mathbf{X}, \mathbf{X}' \in \mathbb{R}^{T \times p}$ we have

$$k(\mathbf{X}, \mathbf{X}') = \langle \phi(\mathbf{X}), \phi(\mathbf{X}') \rangle_{\mathcal{H}}.$$

We call ϕ a feature map and \mathcal{H} a feature space of k .

A kernel function can be interpreted as a kind of similarity measure between the matrices \mathbf{X} and \mathbf{X}' . We say that the function k is a nonnegative definite kernel function if for any finite subset $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ of $\mathbb{R}^{T \times p}$ and any real numbers c_1, \dots, c_n

$$\sum_{i=1}^n \sum_{j=1}^n c_i c_j k(\mathbf{X}_i, \mathbf{X}_j) \geq 0.$$

This raises an interesting question: given a function of two variables $k(\mathbf{X}, \mathbf{X}')$, is there a function $\phi(\mathbf{X})$ such that $k(\mathbf{X}, \mathbf{X}') = \langle \phi(\mathbf{X}), \phi(\mathbf{X}') \rangle_{\mathcal{H}}$?

The answer is provided by Mercer's theorem (Mercer, 1909), which says, roughly, that if k is nonnegative definite, then such a ϕ exists.

For a given finite subset $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ of $\mathbb{R}^{T \times p}$ and a given kernel function k on $\mathbb{R}^{T \times p}$, the matrix $\mathbf{K} = (K_{ij})$ of size $n \times n$, where $K_{ij} = k(\mathbf{X}_i, \mathbf{X}_j)$, is called the Gram matrix or the kernel matrix of the kernel function k with respect to the set $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$.

3. Nonlinear principal component analysis

3.1. Unweighted case

Let $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}$, $\mathbf{X}_i \in \mathbb{R}^{T \times p}$, $i = 1, \dots, n$, be a data set where p is the number of variables and T is the number of time-points in which these variables are observed.

Suppose we first map our data nonlinearly into a Hilbert space \mathcal{H} by $\phi: \mathbb{R}^{T \times p} \rightarrow \mathcal{H}$

Note that \mathcal{H} could have an arbitrarily large, possibly infinite, dimensionality.

In order to carry out linear principal component analysis in feature space \mathcal{H} , we have to find eigenvalues $\lambda \geq 0$ and eigenvector $u \in \mathcal{H}$ of the empirical covariance operator in \mathcal{H}

$$C = \frac{1}{n} \sum_{i=1}^n \langle \psi(\mathbf{X}_i), \psi(\mathbf{X}_i) \rangle_{\mathcal{H}} \tag{1}$$

satisfying

$$Cu = \lambda u, \tag{2}$$

where

$$\langle u, u \rangle_{\mathcal{H}} = 1$$

and

$$\psi(\mathbf{X}_i) = \phi(\mathbf{X}_i) - \bar{\phi}, \quad \bar{\phi} = \frac{1}{n} \sum_{i=1}^n \phi(\mathbf{X}_i), \quad i = 1, \dots, n$$

Since equation (2) can be written as $\frac{1}{n} \sum_{i=1}^n \psi(\mathbf{X}_i) \langle \psi(\mathbf{X}_i), u \rangle_{\mathcal{H}} = \lambda u$, the eigenvectors u lie in the span of $\psi(\mathbf{X}_1), \dots, \psi(\mathbf{X}_n)$. Hence there exist coefficients $\alpha_1, \dots, \alpha_n$ such that

$$u = \sum_{i=1}^n \alpha_i \psi(X_i). \tag{3}$$

Since the vector u lies in the span of $\psi(X_1), \dots, \psi(X_n)$, we can equivalently look for solutions of the n equations:

$$\langle \psi(X_i), Cu \rangle_{\mathcal{H}} = \lambda \langle \psi(X_i), u \rangle_{\mathcal{H}}. \tag{4}$$

Combining equations (1), (3) and (4) gives:

$$\tilde{K}\tilde{K}\alpha = n\lambda\tilde{K}\alpha, \tag{5}$$

where:

$$\alpha = (\alpha_1, \dots, \alpha_n)^T, \tilde{K} = \mathbf{H}\mathbf{K}\mathbf{H}, \mathbf{K} = (k(X_i, X_j)), \mathbf{H} = \mathbf{I}_n - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T, \mathbf{1}_n \in \mathbb{R}^{n \times 1}.$$

The idempotent matrix \mathbf{H} ($\mathbf{H}^2 = \mathbf{H}$) is called the centering matrix. Now every solution to

$$\tilde{K}\alpha = n\lambda\alpha \tag{6}$$

is also a solution of Equation (5), and it turns out that for our purposes it is sufficient to solve Equation (6). To see this, note that every solution of (5) can be written as $\alpha_{\mathcal{N}} + \alpha_{\perp}$, where $\alpha_{\mathcal{N}}$ lies in the null space \mathcal{N} of \tilde{K} and where α_{\perp} lies in the orthogonal subspace \mathcal{N}_{\perp} , then $\alpha_{\mathcal{N}}$ is also a solution to Equation (5), and $(\tilde{K}\alpha_{\mathcal{N}})$ is also a solution to Equation (6). Hence, the solutions to Equation (5), which are in \mathcal{N}_{\perp} , and the solutions to Equation (6) are in 1 – 1 correspondence. We can ignore solutions $\alpha_{\mathcal{N}} \in \mathcal{N}$ since to compute the projection of a given mapped sample X_j we only need to compute:

$$\langle \psi(X_j), u \rangle = \sum_{i=1}^n \alpha_i \langle \psi(X_j), \psi(X_i) \rangle = (\tilde{K}(\alpha_{\mathcal{N}} + \alpha_{\perp}))_j = (\tilde{K}\alpha_{\perp})_j.$$

Thus, we can find all relevant solutions to Equation (5) by taking all solutions to Equation (6) and pre-multiplying by \tilde{K} .

Finally, to compute the projections we need to normalise the eigenvectors in \mathcal{H} to have unit length:

$$\langle u, u \rangle_{\mathcal{H}} = \sum_{i,j=1}^n \alpha_i \alpha_j \langle \psi(\mathbf{X}_i), \psi(\mathbf{X}_j) \rangle_{\mathcal{H}} = n\lambda \sum_{i=1}^n \alpha_i \alpha_i,$$

so the α 's must be normalised to have length $\frac{1}{\sqrt{n\lambda}}$. Since the eigenvalues of $\tilde{\mathbf{K}}$ are $n\lambda$, we can accomplish this by computing a given eigenvector of $\tilde{\mathbf{K}}$, normalising it to have length one, and then dividing by the square root of its eigenvalue. We summarise the nonlinear PCA algorithm schematically below.

Algorithm of Nonlinear Principal Component Analysis (Schematic)
Given: a data set $\{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n\}, \mathbf{X}_i \in \mathbb{R}^{T \times p}, i = 1, \dots, n,$
Compute the kernel matrix $\mathbf{K} = (K_{ij}), K_{ij} = k(\mathbf{X}_i, \mathbf{X}_j),$
Compute the centred kernel matrix $\tilde{\mathbf{K}} = \mathbf{H}\mathbf{K}\mathbf{H}, \mathbf{H} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$
Compute the eigenvectors α_i and eigenvalues λ_i of $\tilde{\mathbf{K}}, i = 1, \dots, n$
Choose the i -th eigenvector α_i along which you would like to project
Normalise α_i to have length $\frac{1}{\sqrt{n\lambda_i}}$
Then for sample $\mathbf{X}_i, i = 1, \dots, n,$ the value of the projection of $\phi(\mathbf{X}_i) \in \mathcal{H}$ along the j -th eigenvector \mathbf{u}_j of the covariance operator of the sample in \mathcal{H} is just
$\langle \psi(\mathbf{X}_i), \mathbf{u}_j \rangle_{\mathcal{H}} = \sum_{k=1}^n \alpha_k^j \langle \psi(\mathbf{X}_i), \psi(\mathbf{X}_k) \rangle = \sum_{k=1}^n \tilde{\mathbf{K}}_{ik} \alpha_k^j$

So far, we have not said anything about the kernel function selection. Throughout this paper, we use the Gaussian kernel

$$k(\mathbf{X}, \mathbf{X}') = \exp(-\gamma \|\mathbf{X} - \mathbf{X}'\|_F^2), \gamma > 0,$$

where

$$\|\mathbf{A}\|_F = \sqrt{\text{tr}(\mathbf{A}^T \mathbf{A})}$$

is the Frobenius norm. The constant $\gamma > 0$ is appropriately selected for the data $\mathbf{X}_1, \dots, \mathbf{X}_n,$ where $\mathbf{X}_i \in \mathbb{R}^{T \times p}, i = 1, \dots, n.$ We take into account the lower-triangu-

lar matrix which has its (i, j) th element given by $\|\mathbf{X}_i - \mathbf{X}_j\|_F^2$, $i, j = 1, \dots, n$. The value of γ was taken to be the reciprocal of the arithmetic mean of the elements of this matrix.

3.2. Weighted case

In this case, the observation matrix \mathbf{X}_i is replaced by the $w_i \mathbf{X}_i$ matrix, where w_i is the positive geographic weight associated with the i -th observation site, $i = 1, \dots, n$. Weights w_i are determined on the basis of a composed index d_i of transport accessibility of cities located in each region (Górniak, 2015) by the following formula:

$$w_i = \frac{d_i}{\sum_{j=1}^n d_j}, i = 1, 2, \dots, n.$$

Because spatial data may be dependent on the place of observation, in order to, at least partially, eliminate this relationship, the places of observation are assigned geographic weights. The rest of the theory remains unchanged.

4. Numerical experiment

The Nonlinear Geographically Weighted PCA theory presented above has been applied in an analysis of higher education infrastructure and higher education students in the Polish regions (voivodships) in the period 2002–2016 ($T = 15$). The data used are taken from the Local Data Bank (<https://bdl.stat.gov.pl>). The Local Data Bank is Poland's largest database with data concerning economy, households, innovation, public finance, society, demographics and the environment. The full set of data contains $p = 7$ variables for $n = 16$ regions. For statistical unification, those data were unitarised (see Walesiak, 2014). They constitute $T \times p \mathbf{X}_i$ matrices $i = 1, \dots, n$.

Each region was characterised by seven-dimensional vector of features.

- X_1 – the number of universities per 1 million inhabitants,
- X_2 – the number of students per 1000 inhabitants,
- X_3 – the number of university graduates per 1000 inhabitants,
- X_4 – the number of academic teachers per 1000 inhabitants,
- X_5 – the number of professors per 100 000 inhabitants,
- X_6 – the number of post-graduate students per 10 000 inhabitants,
- X_7 – the number of doctoral students per 10 000 inhabitants.

The calculations were performed for two cases:

- 1) unweighted data and
- 2) weighted data.

On the graphs, the regions are denoted by numbers as follows: Dolnośląskie (1), Kujawsko-Pomorskie (2), Lubelskie (3), Lubuskie (4), Łódzkie (5), Małopolskie (6), Mazowieckie (7), Opolskie (8), Podkarpackie (9), Podlaskie (10), Pomorskie (11), Śląskie (12), Świętokrzyskie (13), Warmińsko-Mazurskie (14), Wielkopolskie (15), Zachodniopomorskie (16).

4.1. Unweighted case

In this case the data were not weighted, therefore the Nonlinear PCA was applied. The results were obtained by using the Frobenius norm. The first principal component U_1 explains 48.15% of the total variability, and the second principal component U_2 explains 23.21% of the total variability. Hence, together, two principal components explained 71.36% of the total variability. The regions were presented as points in the system based on the first two principal components (see Figure 1). A minimum spanning tree (dendrite) (Florek et al., 1951 and Kruskal, 1956) was constructed on these points. This indicates the division of the regions into homogeneous groups.

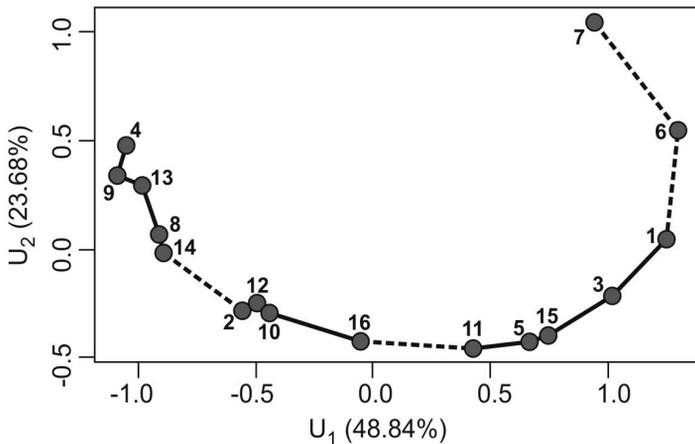


Figure 1. The location of the various regions (unweighted case)

Source: own elaboration

It is easy to identify five groups of regions: {Mazowieckie (7)}, {Małopolskie (6)}, {Dolnośląskie (1), Lubelskie (3), Wielkopolskie (15), Łódzkie (5), Pomorskie (11)}, {Zachodniopomorskie (16), Podlaskie (10), Śląskie (12), Kujawsko-Pomorskie (2)}, {Warmińsko-Mazurskie (14), Opolskie (8), Świętokrzyskie (13), Podkarpackie (9), Lubuskie (4)}.

A closer scrutiny of Figure 1 reveals that, compared to other regions, the Lubelskie Region (3) is positioned too high, and the Śląskie Region (12) is positioned too low. The positions of these two regions appear to be inappropriate.

4.2. Weighted case

The values of the composed index d_i of transport accessibility of cities located in each region (Górniak, 2015) are given in Table 1.

Table 1. Composed index d_i of transport accessibility (CITA)

No.	Voivodship	CITA values
1	Dolnośląskie	285
2	Kujawsko-Pomorskie	327
3	Lubelskie	108
4	Lubuskie	88
5	Łódzkie	254
6	Małopolskie	345
7	Mazowieckie	367
8	Opolskie	91
9	Podkarpackie	131
10	Podlaskie	144
11	Pomorskie	221
12	Śląskie	418
13	Świętokrzyskie	189
14	Warmińsko-Mazurskie	168
15	Wielkopolskie	238
16	Zachodniopomorskie	93

Source: Górniak, 2015

The weights w_i are given in Table 2.

Table 2. The weights w_i determined on the basis of the composed index d_i

w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8
0.0822	0.0943	0.0321	0.0254	0.0733	0.0995	0.1059	0.0262
w_9	w_{10}	w_{11}	w_{12}	w_{13}	w_{14}	w_{15}	w_{16}
0.0378	0.0415	0.0637	0.1206	0.0545	0.0485	0.0686	0.0268

Source: own elaboration

The matrix X was first multiplied by the values of weights w_i , and then principal components were calculated by taking into account the Frobenius norm.

In this case, two principal components explained 80.94% of the total variability. As before, we obtained five groups of regions – according to the number of professors employed and active students in higher education (see Figure 2). These five groups are as follows: {Mazowieckie (7)}, {Małopolskie (6)}, {Dolnośląskie (1)}, {Wielkopolskie (15), Łódzkie (5), Śląskie (12), Kujawsko-Pomorskie (2), Pomorskie (11)}, {Lubel-

skie (3), Podlaskie (10), Zachodniopomorskie (16), Warmińsko-Mazurskie (14), Świętokrzyskie (13), Opolskie (8), Podkarpackie (9), Lubuskie (4}).

A careful analysis of particular groups confirms that the Nonlinear Geographically Weighted PCA procedure provided more adequate results than the Nonlinear PCA.

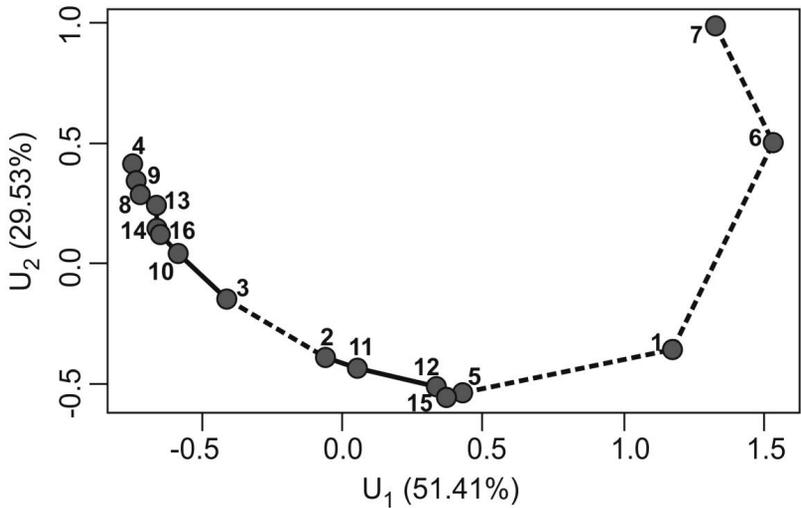


Figure 2. The location of the various regions (weighted case)
 Source: own elaboration

For the calculations, we used R (R Core Team, 2017) software. The R source code is available on request from the fourth co-author.

References

- Anselin L. (1988), *Spatial econometrics: methods and models*, Kluwer Academic Publishers, Dordrecht.
- Anselin L. (2010), *Thirty years of spatial econometrics*, "Regional Science", no. 89(1), pp. 3–25.
- Casetti E. (1972), *Generating Models by the Expansion Method: Applications to Geographical Research*, "Geographical Analysis", no. 4(1), pp. 81–89.
- Charlton M., Brundson C., Demšar U., Harris P., Fotheringham A.S. (2010), *Principal components analysis: From global to local*, paper presented at the 13th AGILE International Conference on Geographic Information Science, Guimarães, Portugal.
- Cliff A.D., Ord J.K. (1973), *Spatial autocorrelation*, Pion, London.
- Demšar U., Harris P., Brundson C., Fotheringham A.S., McLoone S. (2013), *Principal Component Analysis on Spatial Data: An overview*, "Annals of the Association of American Geographers", no. 103(1), pp. 106–128.
- Florek K., Łukaszewicz J., Perkal J., Steinhaus H., Zubrzycki S. (1951), *Sur la liaison et la division des points d'un ensemble fini*, "Colloquium Mathematicum", no. 2, pp. 282–285.
- Górecki T., Krzyśko M., Waszak Ł., Wołyński W. (2018), *Selected statistical methods of data analysis for multivariate functional data*, "Statistical Papers", no. 59, pp. 153–182.
- Górniak J. (2015), *Identification of transport accessibility of Polish cities based on their transport infrastructures*, "Studia Ekonomiczne. Zeszyty Naukowe UE w Katowicach", no. 249, pp. 145–154.
- Kruskal J.B. (1956), *On the shortest spanning subtree of a graph and the traveling salesman problem*, "Proceedings of the American Mathematical Society", no. 7(1), pp. 48–50.
- Mercer J. (1909), *Functions of positive and negative type and their connection with the theory of integral equations*, "Philosophical Transactions of the Royal Society of London", Series A, no. 209, pp. 415–446.
- R Core Team (2017), *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, <https://www.R-project.org/> [accessed: 8.05.2018].
- Schölkopf B., Smola A., Müller K.R. (1998), *Nonlinear component analysis as a kernel eigenvalue problem*, "Neural Computation", no. 10, pp. 1299–1319.
- Swamy P.A.V. (1971), *Statistical inference in random coefficient regression models*, Springer, Berlin.
- Tobler W.R. (1970), *A computer movie simulating urban growth in the Detroit region*, "Economic Geography", no. 46(2), pp. 234–248.
- Walesiak M. (2014), *Data normalization in multivariate data analysis. An overview and properties*, "Przegląd Statystyczny", no. 61(4), pp. 363–372.

Analiza nieliniowych składowych głównych dla danych czasowo-przestrzennych geograficznie ważonych

Streszczenie: Schölkopf, Smola i Müller (1998) zaproponowali analizę nieliniowych składowych głównych (NPCA) dla ustalonych danych wektorowych. Niniejszy artykuł zawiera rozszerzenie tej metody na dane czasowo-przestrzenne oraz czasowo-przestrzenne geograficznie ważne. Każdy obiekt jest scharakteryzowany za pomocą macierzy X_i rozmiaru $T \times p$, zawierającej wartości p cech zaobserwowanych w T momentach czasowych, $i = 1, \dots, n$. Macierze te są przekształcane nieliniowo do przestrzeni Hilberta i budowana jest scentrowana macierz jądrowa. Ostatecznie macierz ta jest podstawą konstrukcji nieliniowych składowych głównych. W przypadku danych geograficznie ważonych macierz X_i zostaje zastąpiona macierzą $w_i X_i$, gdzie w_i jest dodatnią wagą geograficzną związaną z i -tym miejscem obserwacji, $i = 1, \dots, n$. Teoria zilustrowana jest przykładem dotyczącym stanu szkolnictwa wyższego w 16 polskich województwach, notowanego w latach 2002–2016.

Słowa kluczowe: nieliniowa analiza składowych głównych, dane geograficznie ważne, dane czasowo-przestrzenne

JEL: C02, C31, C38

	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p>
	<p>Received: 2018-01-13; verified: 2018-04-03. Accepted: 2018-06-26</p>