

Mirosław Krzyśko^{*}, Łukasz Waszak^{**}

KERNEL FUNCTIONAL CANONICAL CORRELATION ANALYSIS

Abstract. In classical canonical correlation analysis we are interested in the relationship between two random vectors $Y = (Y_1, Y_2, \dots, Y_p)' \in \mathbb{R}^p$ and $X = (X_1, X_2, \dots, X_q)' \in \mathbb{R}^q$. We seek weight vectors $u \in \mathbb{R}^p$ and $v \in \mathbb{R}^q$ such that the linear combinations of vectors $U = u_1 Y_1 + \dots + u_p Y_p = u'Y$ and $V = v_1 X_1 + \dots + v_q X_q = v'X$ called canonical variables, are maximally correlated with each other. Canonical correlation methods for data representing functions or curves have received much attention in recent years. Such data, known in the literature as functional data (Ramsay and Silverman, 2005), has been the subject of much recent research interest. Examples of functional data can be found in several application domains, such as medicine, economics, meteorology and many others. Unfortunately, the multivariate data canonical correlation methods cannot be used directly for functional data, because of the problem of dimensionality and difficulty in taking into account the correlation and order of functional data. The problem of constructing canonical correlations and canonical variables for functional data was addressed by Leurgans et al. (1993), and further developments were made by Ramsay and Silverman (2005). In this paper we propose a new method of constructing canonical correlations and canonical variables for functional data.

Keywords: Functional data, Canonical correlation analysis, Kernel.

JEL: C15.

1. INTRODUCTION AND DATA CONVERSION

Let $\{y_{ij}, x_{ij}\}$ denote an observed value of a pair of statistical features Y and X on the i th individual at the j th time point, where $i = 1, 2, \dots, N, j = 1, 2, \dots, J$. Our data then consist of N triples $\{t_{ij}, y_{ij}, x_{ij}\}$, where $t_{ij} \in [0, T], i = 1, 2, \dots, N, j = 1, 2, \dots, J$. However, in many cases it is more convenient to use continuous functions of time, $y(t)$ and $x(t), t \in [0, T]$, i.e. functional data. In this case the discrete data $\{t_{ij}, y_{ij}, x_{ij}\}$ are transformed to functional data $\{y_i(t), x_i(t), t \in [0, T]\}$. The transformation process is identical for all functions

^{*} Department of Mathematics and Computer Science, Adam Mickiewicz University, mkrzyisko@amu.edu.pl

^{**} Department of Mathematics and Computer Science, Adam Mickiewicz University, lwaszak@amu.edu.pl

$y_i(t)$ and $x_i(t)$, hence our further considerations will concern a single function $x(t)$, $t \in [0, T]$.

Let x_j denote an observed value of statistical feature X at the j th time point t_j , where $j = 1, 2, \dots, J$. Discrete data $\{t_j, x_j, j = 1, 2, \dots, J\}$ can be smoothed by continuous functions $x(t)$, where $t \in [0, T]$ (Ramsay, Silverman 2005). Let us assume that the function $x(t)$ has the following representation

$$x(t) = \sum_{k=0}^K \alpha_k \phi_k(t), \quad t \in [0, T] \quad (1)$$

where $\{\phi_k\}$ are orthonormal basis functions, and the coefficients $\{\alpha_k\}$ are estimated by the least squares method. Let $\mathbf{x} = (x_1, x_2, \dots, x_J)'$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K)'$ and $\boldsymbol{\Phi}$ be a matrix of dimension $J \times (K + 1)$ containing the values $\phi_k(t_j)$, $k = 0, 1, \dots, K, j = 1, 2, \dots, J$. The coefficients $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K)'$ in the expression of (1) are estimated by the least squares method, that is, so as to minimize the function:

$$S(\boldsymbol{\alpha}) = (\mathbf{x} - \boldsymbol{\Phi}\boldsymbol{\alpha})'(\mathbf{x} - \boldsymbol{\Phi}\boldsymbol{\alpha}) \quad (2)$$

Differentiating $S(\boldsymbol{\alpha})$ with respect to the vector $\boldsymbol{\alpha}$, we obtain

$$\hat{\boldsymbol{\alpha}} = (\boldsymbol{\Phi}'\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}'\mathbf{x} \quad (3)$$

The degree of smoothness of the function $x(t)$ depends on the value K (a small value of K causes more smoothing of the curves). The optimum value for K is selected using the Bayesian information criterion BIC (see Shmueli 2010). In the case of N functions $x_i(t)$ one common value of K is chosen, as the modal value of the numbers K_1, K_2, \dots, K_N . In the statistical literature the data $\{x_i(t), t \in [0, T], i = 1, 2, \dots, N\}$ are called functional data.

The orthonormal basis functions usually chosen are the Fourier system:

$$\phi_0(t) = 1, \phi_{2k-1} = \sqrt{\frac{2}{T}} \sin \frac{2\pi kt}{T}, \phi_{2k} = \sqrt{\frac{2}{T}} \cos \frac{2\pi kt}{T},$$

where $k = 1, 2, \dots, t \in [0, T]$, or the system of Legendre polynomials over the space $L_2([-1, 1])$:

$$\tilde{p}_k(x) = \sqrt{\frac{2k+1}{2}} p_k(x),$$

where

$$p_{k+1}(x) = \frac{1}{k+1} [(2k+1)xp_k(x) - kp_{k-1}(x)], \quad k \geq 1,$$

$$p_0(x) = 1, \quad p_1(x) = x.$$

In the functional case we are interested in the relationship between two stochastic processes $Y(t)$ and $X(t)$. We seek weight functions $u(t)$ and $v(t)$ such that $U = \int_0^T u(t)Y(t)dt$ and $V = \int_0^T v(t)X(t)dt$ are maximally correlated. The values U and V are called functional canonical variables (FCCA, see Krzyśko, Waszak 2013).

2. CONSTRUCTION OF FUNCTIONAL CANONICAL VARIABLES

Let us assume that we observe the realization of a two-dimensional random process (Y, X) where $Y(t) \in L_2([0, T])$ and $X(t) \in L_2([0, T])$. Here $L_2([0, T])$ is the space of square integrable functions on the interval $[0, T]$ equipped with the inner product

$$\langle u, v \rangle = \int_0^T u(t)v(t)dt.$$

We further assume that

$$E(\langle Z, Z \rangle) = E \left[\int_0^T Z^2(t)dt \right] < \infty,$$

for $Z = Y$ and $Z = X$.

Canonical correlation for finite-dimensional random vectors $\mathbf{Y} \in \mathbb{R}^{K_1+1}$, $\mathbf{X} \in \mathbb{R}^{K_2+1}$ and for stochastic processes $Y(t) \in L_2([0, T])$, $X(t) \in L_2([0, T])$ is defined as follows. Let $H_i = \mathbb{R}^{K_i+1}$ in the vector case and $H_i = L_2([0, T])$ in the functional case, $i = 1, 2$. Then the first canonical correlation ρ_1 and the associated vectors \mathbf{u}_1 and \mathbf{v}_1 or weight functions $u_1(t)$ and $v_1(t)$ are defined as follows:

$$\rho_1 = \sup_{u \in H_1, v \in H_2} Cov(\langle u, Y \rangle, \langle v, X \rangle) = Cov(\langle \mathbf{u}_1, Y \rangle, \langle \mathbf{v}_1, X \rangle), \quad (4)$$

where u and v are subject to the restrictions

$$Var(\langle u, Y \rangle) = 1, \quad Var(\langle v, X \rangle) = 1. \quad (5)$$

In general, the k th canonical correlation ρ_k and the associated \mathbf{u}_k and \mathbf{v}_k or weight functions $u_k(t)$ and $v_k(t)$ are defined as follows:

$$\rho_k = \sup_{u \in H_1, v \in H_2} \text{Cov}(\langle u, Y \rangle, \langle v, X \rangle) = \text{Cov}(\langle u_k, Y \rangle, \langle v_k, X \rangle), \quad (6)$$

where u and v are subject to the restrictions of (5), and the k th pair of canonical variables (U_k, V_k) is not correlated with the first $k-1$ pairs $\{(U_i, V_i), i = 1, 2, \dots, k-1\}$, where $U_k = \langle u_k, Y \rangle, V_k = \langle v_k, X \rangle$. The expression (ρ_k, u_k, v_k) will be called the k th canonical component.

Considering the case where Y and X are stochastic processes, we will assume that Y and X can be represented by a finite number of orthonormal basis functions. For such finite-dimensional processes, functional canonical analysis is equivalent to ordinary canonical analysis for random vector coefficients. To demonstrate this fact, we adopt the following notation. Let

$$Y(t) = \sum_{k=0}^{K_1} \alpha_k \phi_k(t), \quad X(t) = \sum_{l=0}^{K_2} \beta_l \phi_l(t), \quad t \in [0, T]$$

where $\{\phi_k\}$ are the elements of a basis of the space $L_2([0, T])$ and $\{\alpha_k\}$ and $\{\beta_k\}$ are random variables with zero expected values and finite variances. This means that $E[Y(t)] = E[X(t)] = 0$. This does not lead to loss of generality, however, because the canonical components are determined on the basis of the covariances of the processes $Y(t)$ and $X(t)$, which do not depend on their means.

Let

$$\boldsymbol{\phi}_1(t) = (\phi_1(t), \phi_2(t), \dots, \phi_{K_1}(t))',$$

$$\boldsymbol{\phi}_2(t) = (\phi_1(t), \phi_2(t), \dots, \phi_{K_2}(t))',$$

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{K_1})',$$

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{K_2})',$$

where

$$E(\boldsymbol{\alpha}) = \mathbf{0}, \quad E(\boldsymbol{\beta}) = \mathbf{0},$$

$$\text{Var}(\boldsymbol{\alpha}) = \boldsymbol{\Sigma}_{11}, \quad \text{Var}(\boldsymbol{\beta}) = \boldsymbol{\Sigma}_{22}, \quad \text{Cov}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = E(\boldsymbol{\alpha}\boldsymbol{\beta}') = \boldsymbol{\Sigma}_{12}.$$

Then the processes $Y(t)$ and $X(t)$ can be written in the form

$$Y(t) = \boldsymbol{\alpha}' \boldsymbol{\phi}_1(t), \quad X(t) = \boldsymbol{\beta}' \boldsymbol{\phi}_2(t).$$

Theorem 1. *The k th canonical component $(\rho_k, \mathbf{u}_k, \mathbf{v}_k)$ of the pair of random vectors $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is related to the k th canonical component $(\gamma_k, u_k(t), v_k(t))$ of the pair of random vectors $(Y(t), X(t))$ by the equations*

$$\rho_k = \gamma_k, \quad u_k(t) = \mathbf{u}'_k \boldsymbol{\phi}_1(t), \quad v_k(t) = \mathbf{v}'_k \boldsymbol{\phi}_2(t).$$

The proof of the this theorem is given in Krzyśko and Waszak (2013).

The k th canonical component $(\rho_k, \mathbf{u}_k, \mathbf{v}_k)$ of the pair of random vectors $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is given by the equations:

$$(\mathbf{C}\mathbf{D} - \rho_k^2 \mathbf{I})\mathbf{u}_k = \mathbf{0}, \quad (\mathbf{D}\mathbf{C} - \rho_k^2 \mathbf{I})\mathbf{v}_k = \mathbf{0},$$

where $\mathbf{C} = \boldsymbol{\Sigma}_{11}^{-1} \boldsymbol{\Sigma}_{12}$ and $\mathbf{D} = \boldsymbol{\Sigma}_{22}^{-1} \boldsymbol{\Sigma}_{21}$. The matrices $\mathbf{C}\mathbf{D}$ and $\mathbf{D}\mathbf{C}$ have the same nonzero eigenvalues ρ_k^2 .

3. FUNCTIONAL CANONICAL VARIABLES FROM SAMPLE

Canonical correlation analysis for random vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is based on the matrices $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{22}$ and $\boldsymbol{\Sigma}_{12}$. In practice these matrices are unknown. We estimate them on the basis of N independent realizations of these vectors, forming the matrices

$$\hat{\mathbf{A}} = \begin{bmatrix} \hat{\alpha}_{10} & \hat{\alpha}_{11} & \dots & \hat{\alpha}_{1K_1} \\ \hat{\alpha}_{20} & \hat{\alpha}_{21} & \dots & \hat{\alpha}_{2K_1} \\ \dots & \dots & \dots & \dots \\ \hat{\alpha}_{N0} & \hat{\alpha}_{N1} & \dots & \hat{\alpha}_{NK_1} \end{bmatrix}, \quad \hat{\mathbf{B}} = \begin{bmatrix} \hat{\beta}_{10} & \hat{\beta}_{11} & \dots & \hat{\beta}_{1K_2} \\ \hat{\beta}_{20} & \hat{\beta}_{21} & \dots & \hat{\beta}_{2K_2} \\ \dots & \dots & \dots & \dots \\ \hat{\beta}_{N0} & \hat{\beta}_{N1} & \dots & \hat{\beta}_{NK_2} \end{bmatrix}$$

where $\hat{\alpha}_{ik}$ and $\hat{\beta}_{il}$ are estimators obtained by the least squares method for the parameters α_{ik} and β_{il} in the representation

$$y_i(t) = \sum_{k=0}^{K_1} \alpha_{ik} \phi_k(t), \quad x_i(t) = \sum_{l=0}^{K_2} \beta_{il} \phi_l(t), \quad t \in [0, T],$$

of processes $Y(t)$ and $X(t)$.

Then the unknown matrices Σ_{11} , Σ_{22} and Σ_{12} have estimators of the form

$$\widehat{\Sigma}_{11} = \frac{1}{N} \widehat{A}' \widehat{A}, \quad \widehat{\Sigma}_{22} = \frac{1}{N} \widehat{B}' \widehat{B}, \quad \widehat{\Sigma}_{12} = \frac{1}{N} \widehat{A}' \widehat{B}.$$

Let $\widehat{C} = \widehat{\Sigma}_{11}^{-1} \widehat{\Sigma}_{12}$ and $\widehat{D} = \widehat{\Sigma}_{22}^{-1} \widehat{\Sigma}_{21}$, where $\widehat{\Sigma}_{21} = \widehat{\Sigma}_{12}'$. Then the matrices $\widehat{C}\widehat{D}$ and $\widehat{D}\widehat{C}$ have the same nonzero eigenvalues $\widehat{\rho}_k^2$, and their corresponding eigenvectors \widehat{u}_k and \widehat{v}_k are given by the equations:

$$(\widehat{C}\widehat{D} - \widehat{\rho}_k^2 I) \widehat{u}_k = 0, \quad (\widehat{D}\widehat{C} - \widehat{\rho}_k^2 I) \widehat{v}_k = 0,$$

where $1 \leq k \leq \min(K_1 + 1, K_2 + 1)$.

Having determined the eigenvectors \widehat{u}_k and \widehat{v}_k we determine the weight functions:

$$\widehat{u}_k(t) = \widehat{u}_k \phi_1(t), \quad \widehat{v}_k(t) = \widehat{v}_k \phi_2(t).$$

Hence, the coefficients of the projection of the i th realization $y_i(t)$ of process $Y(t)$ on the j th functional canonical variable are equal to

$$\widehat{U}_{ij} = \langle \widehat{u}_j(t), y_i(t) \rangle = \int_0^T \widehat{u}_j(t) y_i(t) dt = \sum_{k=0}^{K_1} \widehat{a}_{ik} \widehat{u}_{jk} = \widehat{\alpha}'_i \widehat{u}_j,$$

and the coefficients of the projection of the i th realization $x_i(t)$ of process $X(t)$ on the j th functional canonical variable are equal to

$$\widehat{V}_{ij} = \widehat{\beta}'_i \widehat{v}_j,$$

where $i = 1, 2, \dots, N, j = 1, 2, \dots, \min(K_1 + 1, K_2 + 1)$.

4 KERNEL VARIANT OF FUNCTIONAL CANONICAL CORRELATION ANALYSIS

The space \mathbb{R}^N of values of the random vectors $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_{N-1})'$ and $\beta = (\beta_0, \beta_1, \dots, \beta_{N-1})'$ is mapped into a reproducing kernel Hilbert space (with kernel k) by the nonlinear function $\Psi: \mathbb{R}^N \rightarrow H(k)$.

To make matrix multiplication possible, in the kernel variant the number of orthonormal basis functions (for both processes) has to be equal to N ($K_1 = K_2 = N - 1$). Then matrices \widehat{A} and \widehat{B} in a new space $H(k)$ have the forms:

$$\hat{\mathbf{A}}_\psi = \begin{bmatrix} \Psi(\hat{\alpha}_{10}) & \Psi(\hat{\alpha}_{11}) & \dots & \Psi(\hat{\alpha}_{1N-1}) \\ \Psi(\hat{\alpha}_{20}) & \Psi(\hat{\alpha}_{21}) & \dots & \Psi(\hat{\alpha}_{2N-1}) \\ \dots & \dots & \dots & \dots \\ \Psi(\hat{\alpha}_{N0}) & \Psi(\hat{\alpha}_{N1}) & \dots & \Psi(\hat{\alpha}_{NN-1}) \end{bmatrix},$$

$$\hat{\mathbf{B}}_\psi = \begin{bmatrix} \Psi(\hat{\beta}_{10}) & \Psi(\hat{\beta}_{11}) & \dots & \Psi(\hat{\beta}_{1N-1}) \\ \Psi(\hat{\beta}_{20}) & \Psi(\hat{\beta}_{21}) & \dots & \Psi(\hat{\beta}_{2N-1}) \\ \dots & \dots & \dots & \dots \\ \Psi(\hat{\beta}_{N0}) & \Psi(\hat{\beta}_{N1}) & \dots & \Psi(\hat{\beta}_{NN-1}) \end{bmatrix}.$$

Using the kernel trick (Moore-Aronszajn theorem (1950)), i.e.

$$\langle \Psi(d_i), \Psi(d_j) \rangle \geq k(d_i, d_j)$$

we obtain:

$$\hat{\Sigma}_{11,\psi} = \frac{1}{N} \hat{\mathbf{A}}'_\psi \hat{\mathbf{A}}_\psi = [k(\hat{\alpha}_i, \hat{\alpha}_j)] = \mathbf{K}_{\alpha\alpha},$$

$$\hat{\Sigma}_{22,\psi} = \frac{1}{N} \hat{\mathbf{B}}'_\psi \hat{\mathbf{B}}_\psi = [k(\hat{\beta}_i, \hat{\beta}_j)] = \mathbf{K}_{\beta\beta},$$

$$\hat{\Sigma}_{12} = \frac{1}{N} \hat{\mathbf{A}}'_\psi \hat{\mathbf{B}}_\psi = [k(\hat{\alpha}_i, \hat{\beta}_j)] = \mathbf{K}_{\alpha\beta} = \mathbf{K}_{\beta\alpha}.$$

Then the matrices $\hat{\mathbf{C}}_\psi \hat{\mathbf{D}}_\psi$ and $\hat{\mathbf{D}}_\psi \hat{\mathbf{C}}_\psi$ have the form \mathbf{K}_1 and \mathbf{K}_2 , respectively

$$\mathbf{K}_1 = \mathbf{K}_{\alpha\alpha}^{-1} \mathbf{K}_{\alpha\beta} \mathbf{K}_{\beta\beta}^{-1} \mathbf{K}_{\beta\alpha},$$

$$\mathbf{K}_2 = \mathbf{K}_{\beta\beta}^{-1} \mathbf{K}_{\beta\alpha} \mathbf{K}_{\alpha\alpha}^{-1} \mathbf{K}_{\alpha\beta}.$$

The optimal ρ s are given by solving the following eigenproblem:

$$\mathbf{K} \mathbf{x} = \mathbf{0},$$

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_1 - \hat{\rho}_k^2 \mathbf{I}_{K_1+1} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_2 - \hat{\rho}_k^2 \mathbf{I}_{K_2+1} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \hat{\mathbf{u}}_{k,\psi} \\ \hat{\mathbf{v}}_{k,\psi} \end{bmatrix}.$$

The matrices $\mathbf{K}_{\alpha\alpha}$ and $\mathbf{K}_{\beta\beta}$ are non-negative definite and can include singular values. In practice (in order to solve this problem) we apply the idea used in ridge regression (regularization of the matrices), i.e. we replace $\mathbf{K}_{\alpha\alpha}$ with $\mathbf{K}_{\alpha\alpha} + \mu \mathbf{I}$ and $\mathbf{K}_{\beta\beta}$ with $\mathbf{K}_{\beta\beta} + \mu \mathbf{I}$, where μ is a very small constant (see Friedman 1989).

The kernel functions usually chosen are the polynomial kernel:

$$k(\mathbf{x}, \mathbf{y}) = (b + \mathbf{x}'\mathbf{y})^d,$$

or the Gaussian kernel

$$k(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{\sigma^2}\right).$$

5. EXAMPLE

The data considered come from the online database of the World Bank (<http://data.worldbank.org/>). For analysis, 13 countries were chosen ($N = 13$): Australia (AUS), Bangladesh (BGD), Brazil (BRA), Egypt (EGY), India (IND), Indonesia (IDN), Japan (JPN), Kenya (KEN), South Korea (KOR), Mexico (MEX), Pakistan (PAK), Switzerland (CHE), United States (USA). Each country was characterized by two features: Rate of growth in listed domestic companies (Y) and Broad money (% of GDP) (X). The features were recorded from 1991 to 2011 ($J=21$). The series were centered, and calculations were performed using the Fourier basis and polynomial kernel. Rate of growth in listed domestic companies (Y) is presented in Figure 5.1. and Broad money (% of GDP) (X) is presented in Figure 5.2.

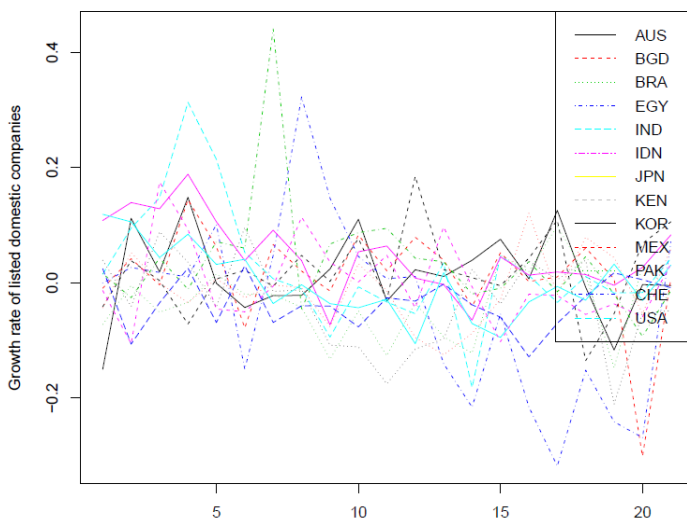
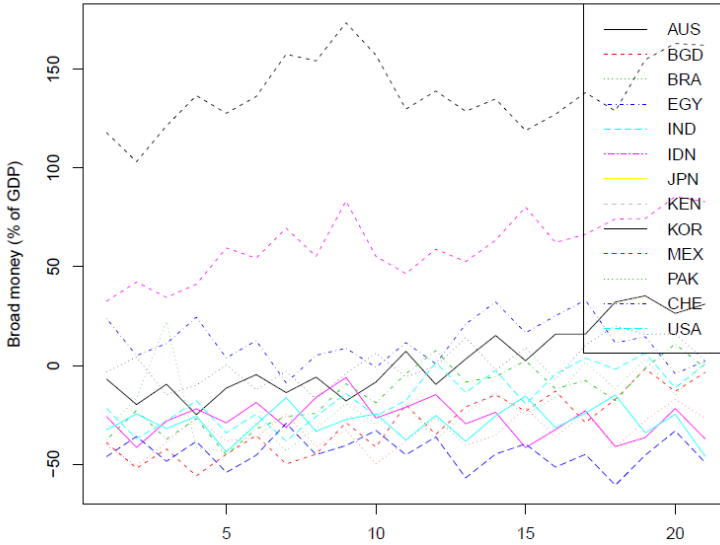


Figure 5.1. Time series of Rate of growth in listed domestic companies (Y)

Figure 5.2. Time series of Broad money (% of GDP) (X)

The two largest functional canonical correlation coefficients are given in Table 5.1, and the two largest kernel functional canonical correlation coefficients in Table 5.2.

Table 5.1. The two largest functional canonical correlation coefficients

ρ_1	ρ_2
0.9410	0.2204

Table 5.2. The two largest kernel functional canonical correlation coefficients

ρ_1	ρ_2
1.0000	0.0429

Finally, the projections of the 13 selected countries on the planes of functional and kernel functional canonical variables (U_1, V_1) are shown in Figure 5.3 and Figure 5.4.

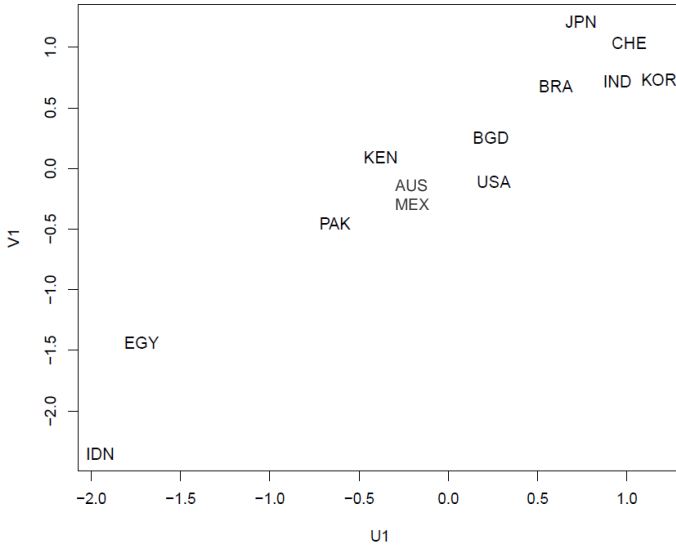


Figure 5.3. Projection of the 13 selected countries on the plane (U_1, V_1) – functional case

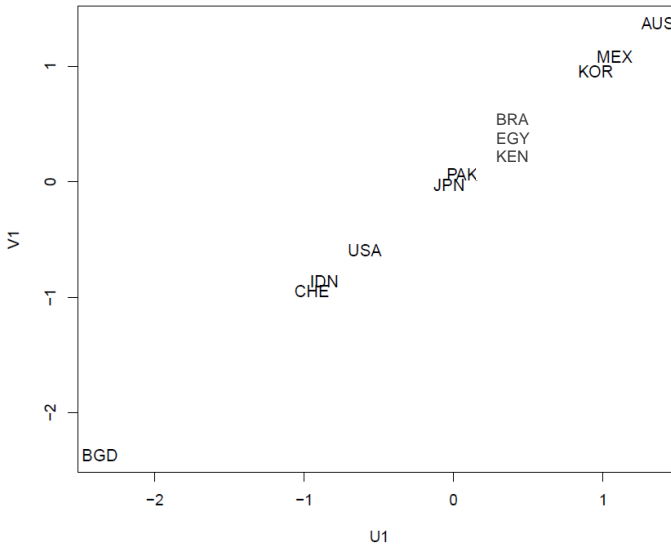


Figure 5.4. Projection of the 13 selected countries on the plane (U_1, V_1) – kernel functional case

The kernel variant, as compared with the classical functional method, gives a correlation coefficient close to 1, i.e. the two data sets can be presented as 100% correlated data sets (that is, the data projection in the kernel case lies on a straight line).

REFERENCES

- Aronszajn N. (1950), *Theory of reproducing kernels*, "Trans. Amer. Math. Soc." 68, p. 337–404.
- Friedman J. H. (1989), *Regularized Discriminant Analysis.*, "J. Amer. Statist. Assoc." 84, p. 165.
- Krzyśko M., Waszak Ł. (2013), *Canonical correlation analysis for functional data*, "Biometrical Letters".
- Leurgans S.E., Moyeed R.A., Silverman B.W. (1993), *Canonical correlation analysis when the data are curves*, "Journal of the Royal Statistical Society", Series B 55, p. 725–740.
- Ramsay J.O., Silverman B.W. (2005), *Functional Data Analysis*, Second Edition. Springer.
- Shmueli G. (2010), *To explain or to predict?*, "Statistical Science" 25(3), p. 289–310.
- The online database of the World Bank: <http://data.worldbank.org/>

Mirosław Krzyśko, Łukasz Waszak

JĄDROWA FUNKCJONALNA ANALIZA KORELACJI KANONICZNYCH

Streszczenie. W ostatnich latach wiele uwagi poświęca się analizie korelacji kanonicznych dla danych reprezentowanych przez funkcje lub krzywe. Takie dane są nazywane w literaturze danymi funkcjonalnymi (Ramsay i Silverman, 2005) i są obszarem wielu zainteresowań badawczych. Przykłady danych funkcjonalnych można znaleźć w wielu praktycznych zastosowaniach, takich jak medycyna, ekonomia, meteorologia i wiele innych. Niestety wielowymiarowe metody korelacji kanonicznych nie mogą być bezpośrednio zastosowane do danych funkcjonalnych, z uwagi na wymiar i trudności związane z uwzględnieniem korelacji i kolejności danych funkcjonalnych. Problem konstrukcji korelacji i zmiennych kanonicznych dla danych funkcjonalnych został zapoczątkowany przez Leurgansa i in. (1993), a następnie rozwinięty przez Ramsaya i Silvermana (2005). W tym artykule proponujemy nową metodę konstrukcji korelacji i zmiennych kanonicznych dla danych funkcjonalnych.

Słowa kluczowe: Dane funkcjonalne, Analiza korelacji kanonicznych, Jądro.

JEL: C15.