Rozwój gospodarczy województw Polski w latach 2010–2021 w świetle analizy danych funkcjonalnych

Marcin Szymkowiak[1,2], Mirosław Krzyśko[3], Waldemar Wołyński[4]

[1]Uniwersytet Ekonomiczny w Poznaniu
[2]Urząd Statystyczny w Poznaniu
[3] Akademia Kaliska im. Prezydenta Stanisława Wojciechowskiego
[4]Uniwersytet im. Adama Mickiewicza w Poznaniu

03.07.2023

1. Cel prezentacji
2. Rozwój gospodarczy
3. Dane funkcjonalne
4. Funkcjonalna analiza składowych głównych
5. Wybrane wyniki
6. Podsumowanie
7. Literatura

Cel prezentacji

- Głównym celem prezentacji jest ocena województw w Polsce z punktu widzenia ich rozwoju gospodarczego z wykorzystaniem aparatu jaki oferuje **analiza danych funkcjonalnych**.
- Na potrzeby realizacji tak postawionego celu autorzy wykorzystali **funkcjonalną analizę składowych głównych** i **analizę skupień**.

Rozwój gospodarczy

- **Rozwój gospodarczy** – długofalowy proces przemian dokonujących się w gospodarce.
- **Rozwój gospodarczy** wyraża się w ilościowym i jakościowym wzroście dóbr i usług konsumpcyjnych oraz produkcyjnych.
- Dotychczas nie wypracowano jednoznacznego i powszechnie stosowanego miernika **rozwoju gospodarczego**.
- **Rozwój gospodarczy** bywa najczęściej wąsko analizowany poprzez wzrost wolumenu produkcji mierzonego za pomocą PKB lub dochodu narodowego per capita i jego struktury.

Rozwój gospodarczy

Wśród czynników rozwoju regionalnego można wyodrębnić trzy grupy:

- **infrastrukturę techniczną** – obejmującą dostęp do urządzeń i sieci: wodno-kanalizacyjnej, energetycznej, gazowej, centralnego ogrzewania, telekomunikacyjnej, komputerowej, transportowej;
- **kapitał ludzki** – tworzony przez zasoby wykwalifikowanej i wykształconej siły roboczej, rozumiany także jako dostęp do oferty edukacyjnej;
- **czynniki miękkie** - w tym kapitał społeczny i innowacyjność – stworzenie warunków do budowania gospodarki opartej na wiedzy, dostęp do nowych technologii, itp.

Dane funkcjonalne

- W ostatnich latach można zaobserwować w obszarze statystyki coraz większe znaczenie reprezentacji danych za pomocą funkcji powierzchni lub krzywych.
- Szczególną rolę odgrywa **funkcjonalna analiza danych**, która zajmuje się analizą informacji zawartych w obiektach ciągłych, takich jak funkcje, powierzchnie, krzywe, etc.

Dane funkcjonalne

- Obiekty będące przedmiotem analizy określa się jako **dane funkcjonalne**.
- **Dane funkcjonalne** charakteryzują się pewną ciągłością, np. względem czasu. Ponieważ w praktyce dane o charakterze ciągłym występują stosunkowo rzadko zachodzi potrzeba zamiany danych obserwowanych w dyskretnych momentach czasu do danych funkcjonalnych o charakterze ciągłym.
- **Analiza danych funkcjonalnych** ma obecnie bardzo szerokie zastosowanie m.in. w chemometrii, klimatologii, biologii, medycynie oraz ekonomii.

Funkcjonalna analiza składowych głównych

- Z kolei za pomocą **funkcjonalnej analizy składowych głównych** (FPCA, ang. **Functional Principal Components Analysis**) możemy dokonać rzutu wysokowymiarowych danych na przestrzeń o dużo mniejszym wymiarze, jednocześnie zachowując maksymalnie dużo informacji.
- Główna idea metody polega na znalezieniu takich funkcji, których iloczyn skalarny z danymi pozwala otrzymać maksymalną zmienność.

Funkcjonalna analiza składowych głównych

- Pierwsza składowa wyjaśnia najwięcej zmienności, druga jest prostopadła do pierwszej i wyjaśnia maksymalnie dużo z tego, co zostało itd.
- Wkład każdej kolejnej składowej w ogólną zmienność początkowych zmiennych jest nie tylko znany, ale i coraz mniejszy, od pewnego momentu składowe będą wyjaśniać znikomy zakres zmienności.
- Większość zmienności zachowana jest zazwyczaj przez kilka pierwszych funkcjonalnych składowych głównych.

Funkcjonalna analiza składowych głównych

Załóżmy, że obserwujemy p-wymiarowy losowy proces z ciągłym parametrem czasu t: $\boldsymbol{X}(t) = (X_1(t), X_2(t), \ldots, X_p(t))^\top$ takim, że $\mathrm{E}(\boldsymbol{X}(t)) = \mathbf{0}$, oraz $\boldsymbol{X}(t) \in L_2^p(I)$, gdzie $L_2^p(I)$ jest przestrzenią Hilberta funkcji całkowalnych z kwadratem na przedziale I z iloczynem skalarnym postaci:

$$\langle \boldsymbol{u}(t), \boldsymbol{v}(t) \rangle = \int_I \boldsymbol{u}^\top(t)\boldsymbol{v}(t)dt. \tag{1}$$

Funkcjonalna analiza składowych głównych

Zakładamy ponadto, że k-ta składowa tego procesu $\boldsymbol{X}(t)$ jest reprezentowana przez skończoną liczbę ortonormalnych funkcji bazowych ϕ_b:

$$X_k(t) = \sum_{b=0}^{B_k} c_{kb}\phi_b(t), \;\; t \in I, \tag{2}$$

gdzie $\boldsymbol{c}_k = (c_{k0}, c_{k1}, \ldots, c_{kB_k})^\top$, są zmiennymi losowymi takimi, że $\mathrm{E}(c_{kb}) = 0$, $\mathrm{Var}(c_{kb}) < \infty$, $k = 1, 2, \ldots, p$, $b = 0, 1, \ldots, B_k$.

Funkcjonalna analiza składowych głównych

Niech $\boldsymbol{c} = (\boldsymbol{c}_1^\top, \boldsymbol{c}_2^\top, \dots, \boldsymbol{c}_p^\top)^\top$ oraz ortonormalna macierz $\boldsymbol{\Phi}(t)$

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} \boldsymbol{\phi}_1^\top(t) & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\phi}_2^\top(t) & \dots & \boldsymbol{0} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{\phi}_p^\top(t) \end{bmatrix}, \tag{3}$$

gdzie $\boldsymbol{\phi}_k(t) = (\phi_0(t), \phi_1(t), ..., \phi_{B_k}(t))^\top$, $k = 1, 2, \dots, p$.

W notacji macierzowej, proces $\boldsymbol{X}(t)$ można przedstawić w następujący sposób:

$$\boldsymbol{X}(t) = \boldsymbol{\Phi}(t)\boldsymbol{c},\ \ t \in I,\ \ \mathrm{E}(\boldsymbol{c}) = \boldsymbol{0},\ \ \mathrm{Var}(\boldsymbol{c}) = \boldsymbol{\Sigma}_{\boldsymbol{c}}, \tag{4}$$

$\boldsymbol{c} \in \mathbb{R}^{K+p}$, $\boldsymbol{\Phi}(t) \in \mathbb{R}^{p\times(K+p)}$, $K = B_1 + B_2 + \dots + B_p$.

Funkcjonalna analiza składowych głównych

Definiujemy funkcjonalną składową główną jako:

$$U = \langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle = \int_I \boldsymbol{u}^\top(t) \boldsymbol{X}(t) dt \tag{5}$$

mającą maksymalną wariancję dla wszystkich wektorów funkcji wagowych $\boldsymbol{u}(t) \in L_2^p(I)$ takich, że $\langle \boldsymbol{u}(t), \boldsymbol{u}(t) \rangle = 1$

Dalej zakładamy, że wektor funkcji wagowych $\boldsymbol{u}(t)$ oraz proces losowy $\boldsymbol{X}(t)$ należą do tej samej przestrzeni, tj. funkcja $\boldsymbol{u}(t)$ jest wyrażana jako:

$$\boldsymbol{u}(t) = \boldsymbol{\Phi}(t) \boldsymbol{u}, \tag{6}$$

gdzie $\boldsymbol{u} \in \mathbb{R}^{K+p}$, $K = B_1 + \cdots + B_p$.

Funkcjonalna analiza składowych głównych

Wówczas

$$\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle = \langle \boldsymbol{\Phi}(t)\boldsymbol{u}, \boldsymbol{\Phi}(t)\boldsymbol{c} \rangle = \boldsymbol{u}^\top \boldsymbol{c} \tag{7}$$

$$\mathrm{E}(\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle) = \boldsymbol{u}^\top \mathrm{E}(\boldsymbol{c}) = \boldsymbol{u}^\top \mathrm{E}(\boldsymbol{0}) = 0, \tag{8}$$

$$\mathrm{Var}(\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle) = \boldsymbol{u}^\top \boldsymbol{\Sigma}_c \boldsymbol{u}, \text{ gdzie } \boldsymbol{\Sigma}_c = \mathrm{E}(\boldsymbol{c}\boldsymbol{c}^\top). \tag{9}$$

Niech

$$\lambda_1 = \sup_{\boldsymbol{u}(t) \in L_2^p(I)} \mathrm{Var}(\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle) = \mathrm{Var}(\langle \boldsymbol{u}_1(t), \boldsymbol{X}(t) \rangle) = \boldsymbol{u}_1^\top \boldsymbol{\Sigma}_c \boldsymbol{u}_1, \tag{10}$$

gdzie

$$\langle \boldsymbol{u}_1(t), \boldsymbol{u}_1(t) \rangle = \boldsymbol{u}_1^\top \boldsymbol{u}_1 = 1. \tag{11}$$

U_1 jest pierwszą funkcjonalną składową główną, a wektor $\boldsymbol{u}_1(t)$ pierwszym wektorem funkcji wagowych.

Funkcjonalna analiza składowych głównych

Następnie poszukujemy drugiej funkcjonalnej składowej głównej

$$U_2 = \langle \boldsymbol{u}_2(t), \boldsymbol{X}(t) \rangle = \boldsymbol{u}_2^\top \boldsymbol{c} \tag{12}$$

maksymalizując wariancję:

$$\mathrm{Var}(\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle) = \boldsymbol{u}^\top \boldsymbol{\Sigma}_c \boldsymbol{u}, \tag{13}$$

w taki sposób, że

$$\langle \boldsymbol{u}_2(t), \boldsymbol{u}_2(t) \rangle = \boldsymbol{u}_2^\top \boldsymbol{u}_2 = 1 \tag{14}$$

i jest ona nieskorelowana z pierwszą funkcjonalną składową główną U_1, tj. spełniony jest warunek:

$$\langle \boldsymbol{u}_1(t), \boldsymbol{u}_2(t) \rangle = \boldsymbol{u}_1^\top \boldsymbol{u}_2 = 0. \tag{15}$$

Funkcjonalna analiza składowych głównych

W ogólności k-ta funkcjonalna składowa główna

$$U_k = \langle \boldsymbol{u}_k(t), \boldsymbol{X}(t) \rangle = \boldsymbol{u}_k^\top \boldsymbol{c} \tag{16}$$

spełnia następujące warunki:

$$\lambda_k = \sup_{\boldsymbol{u}(t) \in L_2^p(I)} \mathrm{Var}(\langle \boldsymbol{u}(t), \boldsymbol{X}(t) \rangle) = \boldsymbol{u}_k^\top \boldsymbol{\Sigma}_c \boldsymbol{u}_k, \tag{17}$$

$$\langle \boldsymbol{u}_{k1}(t), \boldsymbol{u}_{k2}(t) \rangle = \delta_{k_1,k_2}, \;\; k_1, k_2 = 1, 2, \ldots, k. \tag{18}$$

Parę $(\lambda_k, \boldsymbol{u}_k(t))$ nazywamy k-tym układem składowych głównych procesu stochastycznego $\boldsymbol{X}(t)$.

Funkcjonalna analiza składowych głównych

Rozważmy teraz klasyczny problem składowych głównych dla losowego wektora $\boldsymbol{c}$. k-ta składowa główna takiego wektora spełnia warunek:

$$\gamma_k = \sup_{\boldsymbol{u} \in \mathbb{R}^{K+p}} \mathrm{Var}(\langle \boldsymbol{u}, \boldsymbol{c} \rangle) = \boldsymbol{u}_k^\top \boldsymbol{\Sigma}_c \boldsymbol{u}_k, \tag{19}$$

$$\langle \boldsymbol{u}_{k1}, \boldsymbol{u}_{k2} \rangle = \delta_{k_1,k_2}, \tag{20}$$

gdzie $k_1, k_2 = 1, 2, \ldots, k$, $K = B_1 + B_2 + \cdots + B_p$.
Parę $(\gamma_k, \boldsymbol{u}_k)$ nazywamy k-tą składową główną dla wektora $\boldsymbol{c}$.

Funkcjonalna analiza składowych głównych

Znalezienie k-tej składowej głównej dla wektora $\boldsymbol{c}$ jest równoważne ze znalezieniem wartości własnych i odpowiadającym im wektorów własnych macierzy kowariancji $\boldsymbol{\Sigma}_c$ $\boldsymbol{u}_{k1}^\top \boldsymbol{u}_{k2} = \delta_{k_1,k_2}$.

W związku z powyższym otrzymujemy, że k-ta składowa główna procesu losowego $\boldsymbol{X}(t)$ jest powiązana z k-tą składową główną wektora losowego $\boldsymbol{c}$ w następujący sposób:

$$\lambda_k = \gamma_k, \ \ \boldsymbol{u}_k(t) = \boldsymbol{\Phi}(t)\boldsymbol{u}_k, \ \ t \in I, \tag{21}$$

gdzie $k = 1, 2, \ldots, K + p$, $K = B_1 + B_2 + \cdots + B_p$.

Analiza składowych głównych dla losowego wektora $\boldsymbol{c}$ bazuje na macierzy kowariancji $\boldsymbol{\Sigma}_c$, która w praktyce zazwyczaj jest nieznana.

Funkcjonalna analiza składowych głównych

Estymujemy elementy macierzy kowariancji $\boldsymbol{\Sigma}_c$ na bazie n niezależnych realizacji wektora $\boldsymbol{x}_1(t), \boldsymbol{x}_2(t), \ldots, \boldsymbol{x}_n(t)$ procesu losowego $\boldsymbol{X}(t)$.

W większości przypadków dane wyrażane są jako dyskretne momenty w czasie.

Niech x_{ikj} oznacza wartość i-tej realizacji k-tej zmiennej w czasie t_j procesu $\boldsymbol{X}(t)$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, p$, $j = 1, 2, \ldots, J$, $t_j \in I$.

Funkcjonalna analiza składowych głównych

Niech

$$\boldsymbol{x}_{ik} = (x_{ik1}, x_{ik2}, \dots, x_{ikJ})^\top, \ \boldsymbol{c}_{ik} = (c_{ik0}, x_{ik1}, \dots, x_{ikB_k})^\top \tag{22}$$

oraz

$$\boldsymbol{\Psi}_k(t) = (\phi_b(t_j)) \tag{23}$$

będzie macierzą o wymiarach $J \times (B_k + 1)$ z elementami $\phi_b(t_j)$, $b = 0, 1, \dots, B_k$, $k = 1, 2, \dots, p$, $j = 1, 2, \dots, J$, $i = 1, 2, \dots, n$.

Funkcjonalna analiza składowych głównych

Wówczas estymator uzyskany metodą najmniejszych kwadratów wektora $\boldsymbol{c}_{ik}$ ma postać:

$$\hat{\boldsymbol{c}}_{ik} = (\boldsymbol{\Psi}_k^\top(t)\boldsymbol{\Psi}_k(t))^{-1}\boldsymbol{\Psi}_k^\top(t)\boldsymbol{x}_{ik}, \tag{24}$$

$i = 1, 2, \dots, n, \ k = 1, 2, \dots, p.$

Niech $\hat{\boldsymbol{c}}_i = (\boldsymbol{c}_{i1}^\top, \boldsymbol{c}_{i2}^\top, \dots, \boldsymbol{c}_{ip}^\top)^\top, \ i = 1, 2, \dots, n$, będzie wektorem scentrowanym oraz niech

$$\hat{\boldsymbol{C}} = (\hat{\boldsymbol{c}}_1, \hat{\boldsymbol{c}}_2, \dots, \hat{\boldsymbol{c}}_n). \tag{25}$$

Wówczas

$$\hat{\boldsymbol{\Sigma}}_c = \frac{1}{n}\hat{\boldsymbol{C}}\hat{\boldsymbol{C}}^\top. \tag{26}$$

Funkcjonalna analiza składowych głównych

Niech $\hat{\gamma}_1 \geq \hat{\gamma}_2 \geq \cdots \geq \hat{\gamma}_s$ będą niezerowymi wartościami własnymi macierzy wariancji $\hat{\boldsymbol{\Sigma}}_c$ oraz niech $\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \ldots, \hat{\boldsymbol{u}}_s$ będzie odpowiadającym im wektorami własnymi, gdzie $s = \operatorname{rank}(\hat{\boldsymbol{\Sigma}}_c)$.

Współrzędne i-tej realizacji $\boldsymbol{x}_i(t)$ procesu $\boldsymbol{X}(t)$ na k-tej funkcjonalnej składowej głównej głównej są postaci

$$\hat{U}_{ik} = \langle \hat{\boldsymbol{u}}_k(t), \boldsymbol{x}_i(t) \rangle = \langle \boldsymbol{\Phi}(t)\hat{\boldsymbol{u}}_k, \boldsymbol{\Phi}(t)\hat{\boldsymbol{c}}_i \rangle = \hat{\boldsymbol{u}}_k^\top \hat{\boldsymbol{c}}_i, \tag{27}$$

dla $i = 1, 2, \ldots, n,\ k = 1, 2, \ldots, s$.

Funkcjonalna analiza składowych głównych

W przypadku klasycznej analizy składowych głównych wkład poszczególnych cech w konstrukcję danej składowej głównej jest wyznaczany na bazie modułów współczynników każdej cechy.

W przypadku funkcjonalnej analizy składowych głównych odpowiednikami współczynników cech są funkcje wagowe powiązane z danym procesem.

Funkcjonalna analiza składowych głównych

Funkcjonalna składowa główna U_j jest określona przez wektor wag funkcyjnych $\boldsymbol{u}_j(t) = (u_{j1}(t), u_{j2}(t), \ldots, u_{jp}(t))^\top$, $j = 1, 2, \ldots, s$.

Składowa $X_k(t)$ wektora $\boldsymbol{X}(t) = (X_1(t), X_2(t), \ldots, X_p(t))^\top$ procesu losowego ma największy wkład w konstrukcję funkcjonalnej składowej głównej U_j w momencie t jeśli:

$$|u_{jk}(t)| = \max_{1 \leq i \leq p} |u_{ji}(t)|, \ j = 1, 2, \ldots, s, \ \ k = 1, 2, \ldots, p. \qquad (28)$$

Funkcjonalna analiza składowych głównych

Niech A_{jk} oznacza pole powierzchni pod modułem składowej $u_{jk}(t)$ wektora funkcji wagowych, dla $t \in I$, i niech

$$A_{ik}^{\star} = \frac{A_{jk}}{\sum_{k=1}^{p} A_{jk}}, \; j = 1, 2, \ldots, s, \;\; k = 1, 2, \ldots, p. \qquad (29)$$

Składowa $X_k(t)$ wektora $\boldsymbol{X}(t) = (X_1(t), X_2(t), \ldots, X_p(t))^{\top}$ procesu losowego ma największy wkład w konstrukcję funkcjonalnej składowej głównej U_j, dla wszystkich $t \in I$, jeśli

$$A_{jk}^{\star} = \max_{1 \leq i \leq p} |A_{jk}^{\star}|, \; j = 1, 2, \ldots, s, \;\; k = 1, 2, \ldots, p. \qquad (30)$$

Założenia

- Wykorzystano dane opisujące województwa w Polsce z punktu widzenia determinant rozwoju gospodarczego za lata 2010–2021 (BDL): $p = 12$ zmiennych diagnostycznych.
- Dane poddano unitaryzacji zerowanej.
- Funkcje bazowe φ zostały określone na przedziale $I = [0, T] = [0, 12]$.
- Poszczególnym latom (punktom czasu) przypisano następujące wartości: $t_1 = 0.5(2010), t_2 = 1.5(2011), \ldots, t_{12} = 11.5(2021)$.
- Przyjęto, że $B_1 = \ldots = B_{12} = 4$, $K = B_1 + B_2 + \ldots, B_{12} = 48$ $K + p = 48 + 12 = 60$. Zatem $\boldsymbol{c} \in \mathbb{R}^{60}$ oraz $\boldsymbol{\Phi} \in \mathbb{R}^{12\times 60}$.

Założenia

- W charakterze bazy ortonormalnej wykorzystano bazę Fouriera:

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

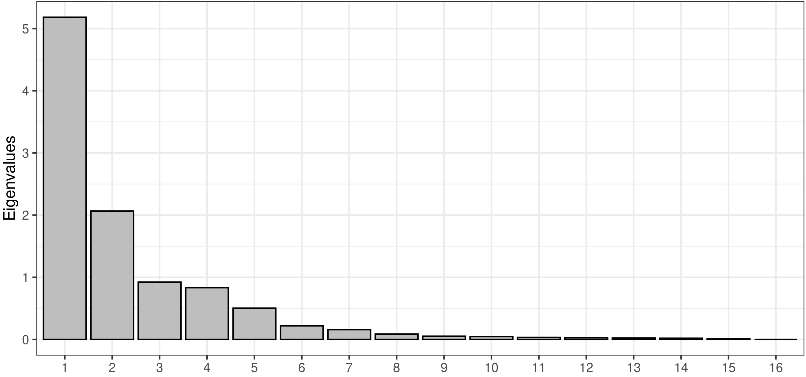
gdzie $k = 1, 2, \ldots, \ t \in [0, T]$

- Dokonano zamiany danych obserwowanych w dyskretnych momentach czasu do danych funkcjonalnych o charakterze ciągłym.
- Bazując na macierzy wariancji $\hat{\boldsymbol{\Sigma}}_c$, $s = 15$ funkcjonalnych składowych głównych zostało skonstruowanych, gdzie $s = \text{rank}(\hat{\boldsymbol{\Sigma}}_c)$.

Zmienne

Cecha	Opis
1	produkt krajowy brutto na 1 mieszkańca
2	przeciętne miesięczne wynagrodzenia brutto
3	liczba podmiotów gospodarczych na 1000 osób w wieku produkcyjnym
4	liczba nowo utworzonych miejsc pracy na 1000 osób w wieku produkcyjnym
5	nakłady inwestycyjne w przedsiębiorstwach na 1 mieszkańca
6	produkcja sprzedana przemysłu ogółem na 1 mieszkańca
7	dochody własne budżetów samorządów lokalnych per capita w zł
8	udział mieszkańców korzystających z wodociągu (w %)
9	udział mieszkańców korzystających z kanalizacji (w %)
10	udział mieszkańców korzystających z gazu (w %)
11	drogi o twardej nawierzchni w km na 100 km^2
12	nowe budynki mieszkalne na 1000 osób

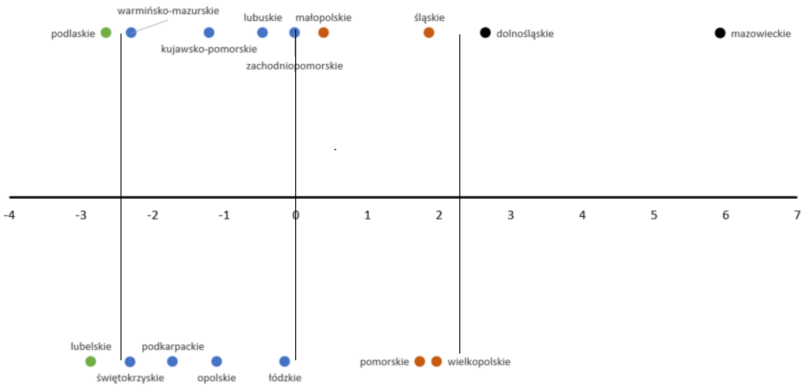
Wartości własne macierzy wariancji $\hat{\Sigma}_c$



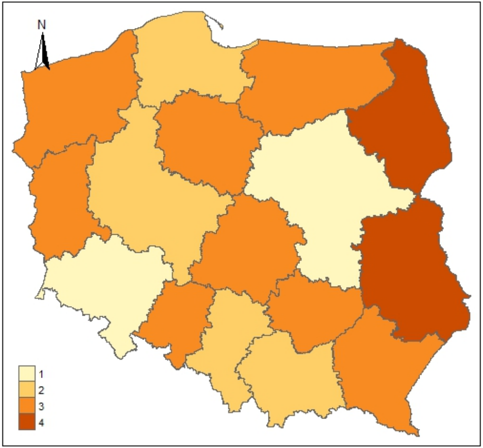
Miary jakości

FPCA	% zmienności	Skumulowany % zmienności
1	50.91	50.91
2	20.29	71.20
3	9.06	80.26

Rzut województw na 1 funkcjonalną składową główną



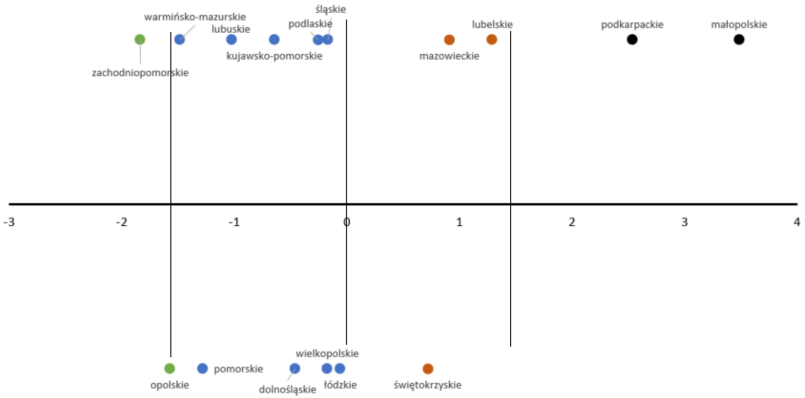
Skupienia województw podobnych – 1 FPCA



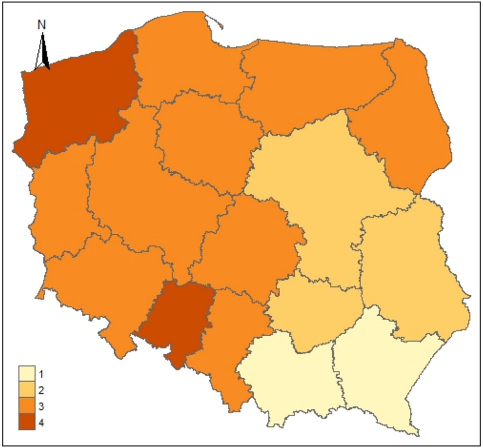
Wkład poszczególnych zmiennych w FPCA

Zmienna	1 FPCA		2 FPCA		3 FPCA	
	Area A_{jk}	Cum(%) $A^{\star}_{jk}$	Area A_{jk}	Cum (%) $A^{\star}_{jk}$	Area A_{jk}	Cum (%) $A^{\star}_{jk}$
1.	1.25	11.28	0.09	1.00	0.65	7.88
2.	1.15	10.39	0.22	2.53	0.66	8.02
3.	1.28	11.56	0.46	5.30	0.16	1.98
4.	1.04	9.42	0.38	4.33	0.32	3.92
5.	1.09	9.80	0.14	1.58	0.28	3.35
6.	1.47	13.28	0.28	3.26	0.12	1.43
7.	1.21	10.88	0.13	1.54	0.46	5.61
8.	0.33	2.99	2.41	27.71	0.38	4.62
9.	0.60	5.41	1.36	15.65	1.96	23.73
10.	0.59	5.33	0.74	8.55	2.51	30.40
11.	0.42	3.80	1.29	14.80	0.25	3.00
12.	0.65	5.88	1.20	13.77	0.50	6.06

Rzut województw na 2 funkcjonalną składową główną



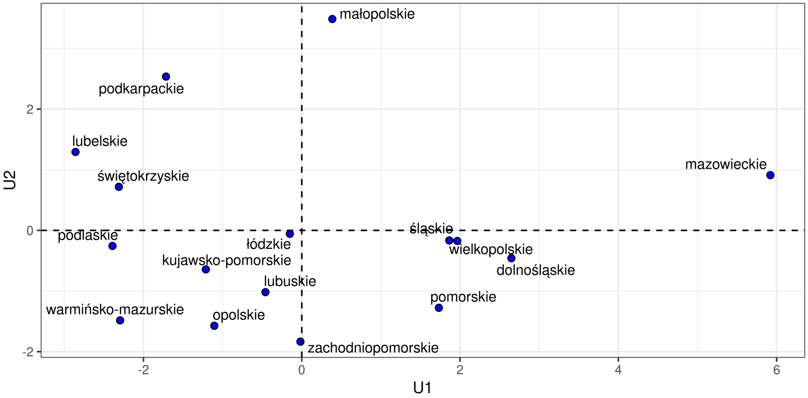
Skupienia województw podobnych – 2 FPCA



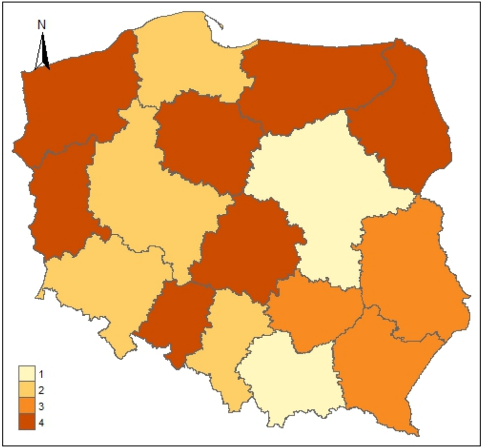
Wkład poszczególnych zmiennych w FPCA

Zmienna	1 FPCA		2 FPCA		3 FPCA	
	Area A_{jk}	Cum(%) $A^{\star}_{jk}$	Area A_{jk}	Cum (%) $A^{\star}_{jk}$	Area A_{jk}	Cum (%) $A^{\star}_{jk}$
1.	1.25	11.28	0.09	1.00	0.65	7.88
2.	1.15	10.39	0.22	2.53	0.66	8.02
3.	1.28	11.56	0.46	5.30	0.16	1.98
4.	1.04	9.42	0.38	4.33	0.32	3.92
5.	1.09	9.80	0.14	1.58	0.28	3.35
6.	1.47	13.28	0.28	3.26	0.12	1.43
7.	1.21	10.88	0.13	1.54	0.46	5.61
8.	0.33	2.99	2.41	27.71	0.38	4.62
9.	0.60	5.41	1.36	15.65	1.96	23.73
10.	0.59	5.33	0.74	8.55	2.51	30.40
11.	0.42	3.80	1.29	14.80	0.25	3.00
12.	0.65	5.88	1.20	13.77	0.50	6.06

Rzut województw na 1 i 2 funkcjonalną składową główną



Skupienia województw podobnych – 1 i 2 FPCA



Podsumowanie

- Zaproponowano wykorzystanie funkcjonalnej analizy składowych głównych w ocenie rozwoju gospodarczego województw w Polsce za lata 2010–2021.
- Uzyskane wyniki są bardzo ważne z punktu widzenia prowadzenia odpowiedniej polityki gospodarczej.

Literatura

Goswami, S., Das, A. K., Chakrabarti, A., Chakraborty, B. (2017), *A feature cluster taxonomy based feature selection technique*. Expert Systems with Applications, 79, 76–89.

Miao, J., Ping, Y., Chen, Z., Jin, X. B., Li, P., Niu, L. (2021), *Unsupervised feature selection by non-convex regularized self-representation*. Expert Systems with Applications, 173, 114643.

Schmutz, A., Jacques, J., Bouveyron, C., Cheze, L., Martin, P. (2020), *Clustering multivariate functional data in group-specific functional subspaces*. Computational Statistics, 1–31.

Literatura

Wu, R., Wang, B., Xu, A. (2021), *Functional data clustering using principal curve methods*. Communications in Statistics-Theory and Methods, 1–34.

Dziękujemy za uwagę!