POLITECHNIKA ŁÓDZKA

WYDZIAŁ FIZYKI TECHNICZNEJ, INFORMATYKI I MATEMATYKI STOSOWANEJ

Kierunek: Matematyka
Specjalność: Matematyczne Metody Analizy Danych Biznesowych

WYBRANE ZASTOSOWANIE STATYSTYCZNYCH METOD PORZĄDKOWANIA DANYCH WIELOWYMIAROWYCH

Kamila Choja
Nr albumu: 204052

Praca licencjacka
napisana w Instytucie Matematyki Politechniki Łódzkiej

Promotor: dr, mgr inż. Piotr Kowalski

ŁÓDŹ, xxx 2017

Spis treści

Rozdział 1

Wstęp

Rozdział 2

Preliminaria

2.1 Notacja

- R^m - przestrzeń liniowa, wektorowa, jej elementy nazywamy zamiennie wektorami lub punktami
- $\mathcal{E}^n$ - przestrzeń euklidesowa n-wymiarowa
- $\mathrm{O} = \{\mathrm{O}_1, \mathrm{O}_2, ..., \mathrm{O}_n\}$ - zbiór obiektów przestrzennych
- $\mathrm{X} = \{\mathrm{X}_1, \mathrm{X}_2, ..., \mathrm{X}_m\}$ - zbiór zmiennych (cech)
- $\mathrm{T} = \{\mathrm{T}_1, \mathrm{T}_2, ..., \mathrm{T}_k\}$ - zbiór okresów (jednostek czasu)
- $\mathrm{OX} = \mathrm{O} \cdot \mathrm{X}$ - zbiór obiekto-zmiennych
- $\mathrm{OT} = \mathrm{O} \cdot \mathrm{T}$ - zbiór obiekto-okresów
- $\mathrm{XT} = \mathrm{X} \cdot \mathrm{T}$ - zbiór zmienno-okresów
- $\mathrm{OXT} = \mathrm{O} \cdot \mathrm{X} \cdot \mathrm{T}$ - zbiór obiekto-zmienno-okresów
- Ω - przestrzeń zdarzeń elementarnych
- ω - zdarzenie elementarne
- $\mathcal{F}$ - rodzina podzbiorów zbioru Ω
- $\mathrm{med}(X_j)$ - mediana cechy X_j
- ρ - relacja porządkująca

2.2 Słownik użytych pojęć

W pracy zostały wykorzystane następujące pojęcia, których wytłumaczenie znajduje się poniżej.

- Statystyka matematyczna [4, Rozdział 1]
  Statystyka matematyczna zajmuje się metodami wnioskowania o całej zbiorowości statystycznej na podstawie zbadania pewnej jej części zwanej próbką lub próbą.

- Cecha statystyczna [7, Rozdział 1]
  Cecha statystyczna jest to liczbowy opis przedmiotu dociekań tj. konkretnej dziedziny życia społeczno-gospodarczego.Służy ona do scharakteryzowania podmiotu badania.

- Macierz obserwacji ([7, Rozdział 2]
  Niech $m > 1$ oraz $n > 1$ będą liczbami naturalnymi. Macierzą obserwacji nazywamy macierz rozmiaru $n \times m$ postaci

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}$$

  gdzie:
  x_{ij} - zaobserwowana wartość j-tej cechy dla i-tego obiektu .

- Skala porządkowa [8, Rozdział 1.2]
  Skalą porządkową nazywana jest skala, pozwalająca na stwierdzeniu identyczności lub różnic porównywanych obiektów, a także na porównywanie wariantów zmiennych zaobserwowanych w obiektach. Nie pozwala ona określić odległości między obiektami. Umożliwia zliczanie obiektów uporządkowanych(liczby relacji równości, nierówności, większości i mniejszości).

- Skala przedziałowa [8, Rozdział 1.2]
  Skalą przedziałową nazywana jest skala, która w stosunku do skali porządkowej, pozwala obliczyć odległość między obiektami, dokonując pomiaru zmiennych za pomocą liczb rzeczywistych. Dla skali tej możliwe jest, obok operacji arytmetycznych dopuszczalnych dla skal o mniejszej mocy, także dodawanie i odejmowanie. Wartość zerowa na tej skali ma charakter umowny, co prowadzi do zachowania różnic między wartościami cechy, przy zmianie jednostek miary.

- Skala ilorazowa [8, Rozdział 1.2]
  Skalą ilorazową nazywana jest skala, która jest podobna do skali przedziałowej(odwołanie), z tym że występuje w niej zero bezwględne(zero ogranicza lewostronnie zakres tej skali). Powoduje to, że można na tej skali obok operacji dopuszczalnych na skalach słabszych dokonywać także dzielenia i mnożenia, a tym samym przedstawiać dowolną wartość cechy danego obiektu jako wielokrotność wartości cechy dla innego obiektu.

- Zmienna objaśniająca [3, Rozdział 1.1] Zmienną objaśniającą nazywamy zmienną w modelu statystycznym, która oddziałuje na zmienne objaśniane. Zmiennie objaśniającą oznaczamy jako $X_1, ..., X_k$, z kolei zmienne objaśniane jako Y.

- Stymulanta [8, Rozdział 1.5]
  Stymulantami nazywane są zmienne, których wysokie wartości badany w badanych obiektach są pożądane z punktu widzenia rozpatrywanego zjawiska.

- Destymulanta [8, Rozdział 1.5]
  Destymulantami nazywane są zmienne, których wysokie wartości badany w badanych obiektach są niepożadane z punktu widzenia rozpatrywanego zjawiska.

- Nominanta [8, Rozdział 1.5]
  Nominantami nazywane są zmienne, których odchylenia wartości w badanym obiekcie od wartości (lub przedziału wartości) uznawanych za najkorzystniejsze są niepożądane z punktu widzenia rozpatrywanego zjawiska.

- Transformacja normalizacyjna [8, Rozdział 1.5]
  Transformacją zmiennych diagnostycznych, mających na celu ujednolicenie ich jednostek pomiarowych, przy zastosowaniu zmiennych diagnostycznych, nazywana jest transformacją normalizacyjną. Można ją przeprowadzić na zmiennych, opisujących porównywane obiekty, mierzonych na skali przedziałowej lub ilorazowej.

  Ogólny wzór na przekształcenie normalizacyjne(Borys, 1978; Grabiński i in., 1989):

  $$z_{ij} = \left(\frac{x_{ij} - a}{b}\right)^p, i = 1, 2, ..., n; j = 1, 2, ..., m; b \neq 0, \tag{2.1}$$

  gdzie:
  z_{ij} - znormalizowana wartość j-tej zmiennej w i-tym obiekcie,
  a, b, p - parametry normalizacyjne.

- Średnia arytmetyczna z próby [7, Rozdział 2.2]
  Średnią arytmetyczną wartości cechy X_j nazywamy wartość

  $$\overline{x_j} = \frac{\sum_{i=1}^{n} x_{ij}}{n}$$

- Odchylenie standardowe z próby [7, Rozdział 2.2]
  Odchyleniem standardowym cechy X_j nazywamy wartość

  $$s_j = \sqrt{s_j^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij} - \overline{x_j})^2}$$

- Mediana [7, Rozdział 2.2]
  Medianę cechy X_j nazywamy wartość

$$\mathrm{med}(X_j) = y = \begin{cases} \frac{1}{2}\Big(x_{(\frac{n}{2})j} + x_{(\frac{n}{2}+1)j}\Big) & \text{jeśli n jest parzyste} \\ \\ x_{(\frac{n+1}{2})j} & \text{jeśli n jest nieparzyste} \end{cases}$$

- Przestrzeń euklidesowa n-wymiarowa $\mathcal{E}^n$ [6, Rozdział 9]
  Przestrzeń euklidesowa n-wymiarowa, jest przestrzenią metryczną przy zwykłej definicji odległości punktu $x = (x_1, x_2, ..., x_n)$ od punktu $y = (y_1, y_2, ..., y_n)$, danej wzorem Pitagorasa

$$|x - y| = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

  gdzie, x i y są ciągami złożonymi z n liczb rzeczywistych.

2.3 Podstawowe pojęcia rachunku prawdopodobieństwa

Definicja 2.3.1. *Ciało zbiorów [9, Rozdział 8.1]*
Rodzinę A podzbiorów zbioru X nazywamy ciałem zbiorów, jeżeli spełnia ona następujące warunku:

1. $\emptyset \in A$,
2. *jeżeli* $A \in \mathcal{A}$, *to* $X \ A \in \mathcal{A}$,
3. *jeżeli* $A \in \mathcal{A}$, *to* $A \cup B \in \mathcal{A}$.

Definicja 2.3.2. *σ-algebra/ciało zbiorów/ zbiorów mierzalnych[9, Rozdział 8.1]*
Ciało zbiorów $\mathcal{A}$ nazywamy σ-ciałem zbiorów, jeżeli spełnia ona warunek dla dowolnych zbiorów $A_n \in \mathcal{A}, n \in \mathbb{N}$, mamy

$$\bigcup_{i=1}^{\infty} A_n \in \mathcal{A}.$$

Elementy σ-ciała $\mathcal{A}$ nazywamy zbiorami mierzalnymi.

Definicja 2.3.3. *Przestrzeń zdarzeń elementarnych [5, w oparciu o rozdział 1.1]*
Zbiór wszyskich możliwych wyników doświadczenia losowego nazywamy przestrzenią zdarzeń elementarnych i oznaczamy przez Ω. Elementy zbioru Ω nazywamy zdarzeniami elementarnymi i oznaczamy ω.

Definicja 2.3.4. *Zdarzenie losowe [5, w oparciu o rozdział 1.1]*
Zdarzeniem losowym (zdarzeniem) nazywamy każdy zbiór $A \in \mathcal{F}$, gdzie $\mathcal{F}$ jest rodziną podzbiorów Ω spełniającą następujące warunki:

1. $\Omega \in F$;
2. *Jeśli* $A \in \mathcal{F}$, *to* $A' \in F$, *gdzie* $A' = \Omega \backslash A$ *jest zdarzeniem przeciwnym do zdarzenia* A;
3. *Jeśli* $A_i \in F$, $i = 1, 2, ...,$*to* $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Rodzinę $\mathcal{F}$ spełniającą warunki 1 - 3 nazywamy σ-ciałem podzbiorów zbioru Ω

Definicja 2.3.5. *Przestrzeń propabilistyczna [5, w oparciu o rozdział 1.2]*
Przestrzenią propabilistyczną nazywamy uporządkowaną trójkę $(\Omega, \mathcal{F}, P)$, gdzie Ω jest zbiorem zdarzeń elementarnych, $\mathcal{F}$ jest σ-ciałem podzbiorów Ω, zaś P jest prawdopodobieństwem określonym na F.

Definicja 2.3.6. *Przestrzeń mierzalna w n-wymiarowej przestrzeni euklidesowej [1, Rozdział 1]*
Niech (Ω, F, P) oznacza przestrzeń propabilistyczną. Przestrzenią mierzalną w n-wymiarowej przestrzeni euklidesowej R^n, nazywamy uporządkowaną dwójkę (R^n, B^n), gdzie B^n jest σ-ciałem podzbiorów borelowskich tej przestrzeni, $n \geqslant 1$.

Definicja 2.3.7. *Prawdopodobieństwo [5, w oparciu o rozdział 1.1]*
Prawdopodobieństwem nazywamy dowolną funkcję P o wartościach rzeczywistych, określoną na σ-ciele zdarzeń $\mathcal{F} \subset 2^{\Omega}$, spełniającą warunki:

1. *$P(A) \geqslant 0$ dla każdego $A \in \mathcal{F}$*
2. *$P(\Omega) = 1$*
3. *Jeśli $A_i \in \mathcal{F}$, $i = 1, 2, ...$ oraz $A_i \cap A_j$ dla $i \neq j$, to*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

Definicja 2.3.8. *Zmienna losowa [5, Rozdział 2.1]*
Niech $(\Omega, \mathcal{F}, P)$ będzie dowolną przestrzenią propabilistyczną. Dowolną funkcję $X : \Omega \to \mathbb{R}$ nazywamy zmienną losową jednowymiarową, jeśli dla dowolnej liczby rzeczywistej x zbiór zdarzeń elementarnych ω, dla których spełniona jest nierówność $X(\omega) < x$ jest zdarzeniem, czyli

$$\{\omega : X(\omega) < x\} \in F \text{ dla każdego } x \in \mathbb{R}$$

Definicja 2.3.9. *Funkcja mierzalna [9, w oparciu o rozdział 8.2]*
Niech X będzie niepustym zbiorem, A σ-ciałem na X i $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. Funkcję $f : X \to \overline{\mathbb{R}}$ nazywamy mierzalną, jeżeli zbiór

$$\{x \in X : f(x) > a\}$$

jest mierzalny przy dowolnym $a \in \mathbb{R}$.

Definicja 2.3.10. *Wektor losowy n-wymiarowy [1, Rozdział 1]*
Wektorem losowym n-wymiarowym nazywamy funkcję $X : \Omega \to \mathbb{R}^n$ mierzalną względem σ-ciała $\mathcal{F}$ ($\mathcal{F}$-mierzalną), tzn. taką, że $X^{-1}(B) \in \mathcal{F}$ dla każdego $B \in \mathcal{F}$.

Definicja 2.3.11. *Wartość oczekiwana [5, Rozdział 2.6]*
Niech X będzie zmienną losową typu dyskretnego lub ciągłego. Wartością oczekiwaną zmiennej losowej X nazywamy

$$E(X) = \begin{cases} \sum_{i=1}^{n} x_i p_i & \text{jeśli zmienna jest typu dyskretnego} \\ \\ \int_{-\infty}^{\infty} x f(x) dx & \text{jeśli zmienna jest typu ciągłego} \end{cases}$$

Definicja 2.3.12. *Wartość oczekiwana macierzy losowej X [1, Rozdział 1.3]*
Wartością oczekiwaną macierzy losowej X nazywamy macierz postaci

$$E(X) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & ... & E(X_{1r}) \\ E(X_{21}) & E(X_{22}) & ... & E(X_{2r}) \\ ... & ... & ... & ... \\ E(X_{n1}) & E(X_{n2}) & ... & E(X_{nr}) \end{bmatrix}$$

przy założeniu, że wszystkie wartości oczekiwane $E(X_{ij})$, $i = 1, 2, ..., n, j = 1, 2, ..., r$, istnieją.

Definicja 2.3.13. *Macierz kowariancji n-wymiarowego wektora losowego X [1, Rozdział 1]*
Macierzą kowariancji n-wymiarowego wektora losowego X nazywamy macierz

$$\sum = E\{[X - E(X)][X - E(X)]'\}$$

Definicja 2.3.14. *Kowariancja [1, Rozdział 1]*
Niech X_i i X_j będą zmiennymi losowymi, $\sum$ będzie macierzą kowariancji n-wymiarowego wektora losowego X. Kowariancją zmiennych losowych X_i i X_j, nazywamy

$$cov(X_i, X_j) = \sigma_{ij} = E\{[X_i - E(X_i)][X_j - E(X_j)]\},\ i, j = 1, 2, ..., n$$

gdzie σ_{ij} jest elementem macierzy kowariancji n-wymiarowego wektora losowego X.

Definicja 2.3.15. *Współczynnik Pearsona [7, Rozdział 2.2]*
Współczynnik Pearsona oznaczamy:

$$r_{jk} = \frac{cov(X_j, X_k)}{s_j s_k}$$

gdzie:
$cov(X_j, X_k)$ *- kowariancja cech X_j i X_k .*

Definicja 2.3.16. *Macierz korelacji par zmiennych [7, Rozdział 2.2]*
Macierzą korelacji par zmiennych, nazywamy macierz postaci:

$$R = \begin{bmatrix} 1 & r_{12} & ... & r_{1m} \\ r_{21} & 1 & ... & r_{2m} \\ ... & ... & ... & ... \\ r_{m1} & r_{m2} & ... & 1 \end{bmatrix}$$

gdzie:
r_{jk} *- współczynnik korelacji liniowej Pearsona j-tej i k-tej cechy (czyli X_j oraz X_k) .*

2.4 Relacja porządkująca

W niniejszej pracy skupiamy się na zagadnieniu porządkowania danych wielowymiarowych. Koniecznym jest zatem przywołanie odpowiednich sformułowań dotyczących matematycznej definicji porządku. Najbardziej podstawowym pojęciem jest relacja porządku, którą teraz definiujemy

Definicja 2.4.1. *Relacja porządkująca [6, Rozdział 1]*
Niech dana relacja ρ, którą oznaczać będziemy przez $\leqslant$, będzie określona dla elementów ustalonego zbioru X. Mówimy, że relacja $\leqslant$ jest relacją porządkującą, jeśli spełnione są warunki:

1. *$x \leqslant x$ dla każdego x (zwrotność),*
2. *jeśli $x \leqslant y$ i $y \leqslant x$, to $x = y$,*
3. *jeśli $x \leqslant y$ i $y \leqslant z$, to $x \leqslant z$ (przechodniość).*

Definicja 2.4.2. *Relacja liniowo porządkująca (liniowy porządek) [2, Rozdział 2]*
Niech dany będzie zbiór X. Relację $\leqslant$ porządkującą zbiór X, nazywamy relacją liniowo porządkującą lub porządkiem liniowym, gdy dla dowolnych x, $y \in X$ spełnia ona następujący warunek liniowości:

$$x \leqslant y \text{ lub } y \leqslant x$$

Parę $(X, \leqslant)$ nazywamy zbiorem liniowo uporządkowanym lub łańcuchem.

Definicja 2.4.3. *Dobry porządek [2, Rozdział 2]*
Niech dany będzie zbiór X. Relację $\leqslant$ porządkującą zbiór X, nazywamy dobrym porządkiem na zbiorze X, gdy w każdym niepustym podzbiorze zbioru X istnieje element najmniejszy względem relacji $\leqslant$. Jeśli relacja $\leqslant$ na zbiorze X jest dobrym porządkiem, to mówimy, że para $(X, \leqslant)$ jest zbiorem dobrze uporządkowanym.

Definicja 2.4.4. *Ograniczenie górne [2, Rozdział 2]*
Niech $A \subseteq X$, gdzie $(X, \leqslant)$ jest zbiorem uporządkowanym. Element $x \in X$ nazywamy ograniczeniem górnym zbioru A względem relacji $\leqslant$, gdy $a \leqslant x$ dla każdego $a \in A$.

Definicja 2.4.5. *Ograniczenie dolne [2, Rozdział 2]*
Niech $A \subseteq X$, gdzie $(X, \leqslant)$ jest zbiorem uporządkowanym. Element $y \in X$ nazywamy ograniczeniem dolnym zbioru A względem relacji $\leqslant$, gdy $a \leqslant x$ dla każdego $a \in A$.

Definicja 2.4.6. *Zbiór ograniczony z góry, zbiór ograniczony z dołu, zbiór ograniczony [2, Rozdział 2]*
Niech $A \subseteq X$, gdzie $(X, \leqslant)$ jest zbiorem uporządkowanym. Zbiór nazywamy ograniczonym z góry (ograniczonym z dołu), jeśli ma on ograniczenie górne (dolne).
Zbiór ograniczony z dołu i z góry nazywamy ograniczonym.

Definicja 2.4.7. *Kres górny [2, Rozdział 2]*
Niech $A \subseteq X$, gdzie $(X, \leqslant)$ jest zbiorem uporządkowanym. Jeśli zbiór A jest ograniczony z góry i wśród ograniczeń górnych zbioru A istnienie element najmniejszy x_0, to element ten nazywamy kresem górnym zbioru A i oznaczamy symbolem $supA$. Tak więc $x_0 = supA$, gdy spełnione są następujące warunki:

1. *$a \leqslant x_0$ dla każdego $a \in A$,*
2. *jeśli $a \leqslant x$ dla każdego $a \in A$, to $x_0 \leqslant x$.*

Definicja 2.4.8. *Kres dolnym [2, Rozdział 2]*
Niech $A \subseteq X$, gdzie $(X, \leqslant)$ jest zbiorem uporządkowanym. Jeśli zbiór A jest ograniczony z dołu i wśród ograniczeń dolnych zbioru A istnienie element największy x_0, to element ten nazywamy kresem dolnym zbioru A i oznaczamy symbolem $infA$. Tak więc $x_0 = infA$, gdy spełnione są następujące warunki:

1. *$y_0 \leqslant a$ dla każdego $a \in A$,*
2. *jeśli $y \leqslant a$ dla każdego $a \in A$, to $y \leqslant y_0$.*

Przechodząc do dalszej dalszych pojęć związanych z porządkowaniem, przytoczę opis obiektu wzorcowego, a następnie podam jego formalną definicję.

Stwierdzenie w oparciu o [8, Rozdział 2.2]
Obiektem wzorcowym nazywany jest obiekt modelowy o pożądanych wartościach zmiennych wejściowych.

Definicja 2.4.9. *Obiekt wzorcowy [7, Rozdział 2.1]*
Obiektem wzorcowym, nazywamy obiekt powstały na podstawie macierzy wystandaryzowanych zmiennych wejściowych. Współrzędnymi obiektu są:

$$O_0 = [z_{oj}], j = 1, 2, ..., m.$$

gdzie:
Współrzędne obiektu wzorcowego obliczamy na podstawie następującego wzoru:

$$z_{oj} = max_i$$

$$z_{oj} = \begin{cases} \max\limits_i \left\{ z_{ij} \right\} & dla\ z_j^S \\ \\ \min\limits_i \left\{ z_{ij} \right\} & dla\ z_j^D \end{cases} \tag{2.2}$$

gdzie:
$j = 1, 2, ..., m; i = 1, 2, ..., n.$

Definicja 2.4.10. *Funkcja kryterium dobroci uporządkowania [8, Rozdział 2.2]*
Funkcją kryterium dobroci uporządkowania nazywamy funkcję:

$$F^2 = \sum_{i'=1}^{n-1} i' \sum_{i=1}^{n-i'} d_{ii'}$$

gdzie:
$d_{i,i'}$ - *odległość euklidesowa między i-tym i i'-tym obiektem .*

Rozdział 3

Metody porządkowania

W oparciu o [8, Rozdział 2], metody porządkowania zbioru obiektów można podzielić na metody porządkowania liniowego i metody porządkowania nieliniowego. Obie grupy metod mogą stanowić punkt wyjścia do grupowania obiektów.
Metody porządkowania liniowego pozwalają na ustalenie hierarchii obiektów ze względu na określone kryterium. Problematyka związana z grupowaniem obiektów ma tutaj znaczenie drugoplanowe. Natomiast stosowanie metod porządkowania nieliniowego nie pozwala na ustalenie hierarchii obiektów, lecz wyłącznie wskazanie dla każdego z tych obiektów podobnych ze względu na wartości opisujących je zmiennych. Powoduje to, że porządkowanie nieliniowe stanowi przede wszystkim etap wstępny do grupowania obiektów.

Bibliografia

[1] Jarosław Bartoszewicz. *Wykłady ze statystyki matematycznej*. Państwowe Wydawnictwo Naukowe, Warszawa, 1996.

[2] Aleksander Błaszczyk and Sławomir Turek. *TEORIA MNOGOŚCI*. Państwowe Wydawnictwo Naukowe, Warszawa, 2007.

[3] Tadeusz Grabiński, Stanisław Wydmus, and Aleksander Zelias. *Metody doboru zmiennych w modelach ekonometrycznych*. Państwowe Wydawnictwo Naukowe, Warszawa, 1982.

[4] Jerzy Greń. *STATYSTYKA MATEMATYCZNA MODELE I ZADANIA*. Państwowe Wydawnictwo Naukowe, Warszawa, 1984.

[5] W. Krysicki, J. Bartos, W. Dyczka, K. Królikowska, and M. Wasilewski. *RACHUNEK PRAWDOPODOBIEŃSTWA I STATYSTYKA MATEMATYCZNA W ZADANIACH: CZĘŚĆ I. Rachunek prawdopodobieństwa*. Państwowe Wydawnictwo Naukowe, Warszawa, 1999.

[6] Kazimierz Kuratowski. *WSTĘP DO TEORII MNOGOŚCI I TOPOLOGII*. Państwowe Wydawnictwo Naukowe, Warszawa, 2004.

[7] Andrzej Młodak. *ANALIZA TAKSONOMICZNA W STATYSTYCE REGIONALNEJ*. Centrum Doradztwa i Informacji Difin, Warszawa, 2006.

[8] Tomasz Panek and Jan Karol Zwierzchowski. *Statystyczne metody wielowymiarowej analizy porównawczej: teoria i zastosowania*. Oficyna Wydawnicza, Szkoła Główna Handlowa, Warszawa, 2013.

[9] Ryszard Rudnicki. *WYKŁADY Z ANALIZY MATEMATYCZNEJ*. Państwowe Wydawnictwo Naukowe, Warszawa, 2006.