Krystyna Gorzelańczyk

ZASTOSOWANIE METODY KORELACJI KANONICZNEJ W ANALIZIE KOSZTÓW PRZEDSIĘBIORSTW HANDLOWYCH

1. Uwagi wstępne

W badaniu współzależności pomiędzy zjawiskami społeczno-ekonomicznymi nie rzadko mamy do czynienia z układami zmiennych, o których chcielibyśmy wyrazić sąd o ich wzajemnych relacjach. Klasyczna analiza korelacji i regresji stwarza możliwość dochodzenia zależności pomiędzy jedną zmienną objaśnianą a zespołem zmiennych objaśniających, lecz w sytuacji, gdy zmienne te obciążone są przy tym zjawiskiem współliniowości, analiza okazuje się mało efektywna. Badacz zjawisk społeczno-gospodarczych zainteresowany jest często nie tylko kompleksową analizę współzależności zmiennych, lecz również uogólnionym wnioskowaniem o zmienności jednego zbioru zmiennych na podstawie innego zbioru. Po raz pierwszy zwrócił na ten fakt H. Hotelling[1] wprowadzając pojęcia: korelacji kanonicznych i zmiennych kanonicznych. Koncepcje jego zostały przedstawione w wielu klasycznych monografiach z zakresu wielowymiarowej analizy statystycznej[2], jak również doczekały się rozwinięcia w latach następnych[3].

Zastosowanie analizy kanonicznej, szczególnie w polskiej praktyce badawczej jest jednak ograniczone i analogicznie jak poniżej przedstawione posiada charakter eksperymentalny i poszukiwawczy[4]. Celem przedstawio-

---

[1] H. Hotelling (8), s. 139-142.

[2] T. Anderson (1), s. 288-306; P. Bertier, J. Bouroche (2), s.149-161; W. Colley, P. Lohnes (3), s. 31-59.

[3] H. Theil (13), s. 115; J. W. Hooper (7), s. 245-256; H. Gaensslen, W. Schubö (6), s. 165-199; J. Bortz (4) s. 746-762.

[4] S. Mejza (9) s. 253-274; Praca zbiorowa pod red. Z. Chojnickiego (10), ss. 127.

nego opracowania jest zaprezentowanie metody kanonicznej w odniesieniu do konkretnego przedmiotu badania.

Identyfikacja współzależności pomiędzy kosztami przedsiębiorstw handlowych i czynnikami generującymi ich zmienność jest wstępnym etapem pogłębionej analizy kosztów i obok cząstkowych dochodzeń związków w ramach poszczególnych rodzajów kosztów, interesującym wydaje się rozpoznanie uogólnionego związku między dwoma interesującymi układami zmiennych, z jednej strony kosztów reprezentowanych przez jednorodne grupy klasyfikacyjne, z drugiej czynników charakteryzujących działalność przedsiębiorstwa. Rozszerzenie analizy o takie syntetyczne spojrzenie umożliwia jednocześnie konstrukcję regresyjnych modeli kanonicznych dla każdej ze zmiennych wyjściowych, jak również ich redukcję do mniejszej liczby hipotetycznych wymiarów.

W przedstawionym opracowaniu metoda kanoniczna, pozwalająca dokonać obok kompleksowej również uogólnionej analizy współzależności została zaprezentowana na przykładzie przedsiębiorstw hurtowych branży spożywczej, których terytorialny zasięg obejmował rynek miejski całego kraju.

2. Metoda badania

Analiza kanoniczna zmierza do określenia związku współzależnościowego pomiędzy dwoma układami zmiennych. Zidentyfikowanie tej zależności możliwe jest poprzez zredukowanie wyjściowych zbiorów zmiennych do mniejszej liczby nieobserwowalnych ortogonalnych wymiarów, które będąc reprezentantami pierwotnych zmiennych pozwalają zmierzyć zależność pomiędzy interesującymi zbiorami cech.

Wykorzystując twierdzenia algebry liniowej, można znaleźć[5] dla obu zbiorów jednocześnie pewną liczbę kombinacji liniowych zmiennych należących do tych zbiorów, które nazywane są zmiennymi kanonicznymi. Obliczając korelacje pomiędzy liniowymi kombinacjami można określić parę najsilniej skorelowaną ze sobą, która jest parą pierwszych zmiennych kanonicznych, następnie parę słabiej skorelowaną, tj. parę drugich zmiennych, itd. Współczynniki korelacji pomiędzy otrzymanymi meta-cechami nazywa się kanonicznymi współczynnikami korelacji.

Przyjmując, że $x = (x_1, x_2, \ldots, x_p)'$ i $y = (y_1, y_2, \ldots, y_q)'$ są dwoma wektorami kolumnowymi odpowiednio p i q zmiennych losowych, zmierza się w analizie do otrzymania zmiennych kanonicznych[6] u_i i v_i $i=1,2,\ldots,r$; $r = \min(p,q)$, będących liniowymi kombinacjami przestrzeni x i y o postaci:

$$u_i = \alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_p X_p \qquad (1)$$

$$v_i = \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_q Y_q \qquad (2)$$

---

[5] J. Bortz (4), s. 746-762; T. Anderson (2), s. 288-306.

[6] Tamże.

-gdzie:

u_i, v_i - zmienne kanoniczne,
$X_1, X_2, \ldots, X_p$ - zmienne objaśniające,
$Y_1, Y_2, \ldots, Y_q$ - zmienne objaśniane,
$\alpha_1, \alpha_2, \ldots, \alpha_p$ - ładunki zmiennej kanonicznej u_i w zmiennych objaśniających
$\beta_1, \beta_2, \ldots, \beta_q$ - ładunki zmiennej kanonicznej v_i w zmiennych objaśnianych.

Współczynniki $\alpha = \alpha_1, \alpha_2, \ldots, \beta_p$ oraz $\beta = \beta_1, \beta_2, \ldots, \beta_q$ znajduje się maksymalizując korelację pomiędzy dwoma odpowiednimi zmiennymi kanonicznymi. Korelację między parą zmiennych kanonicznych można przedstawić za pomocą wzoru:

$$\varrho_i = \frac{\cos u_i v_i}{\operatorname{var} u_i \operatorname{var} v_i} \tag{3}$$

Największy współczynnik korelacji kanonicznej nazywa się pierwszym współczynnikiem korelacji (ϱ_1), a otrzymane liniowe kombinacje w wyniku jego maksymalizacji pierwszymi zmiennymi kanonicznymi u_1 v_1. Maksymalizacja drugiego współczynnika korelacji ϱ_2 prowadzi do otrzymania drugich zmiennych kanonicznych u_2, v_2, nieskorelowanych z pierwszymi zmiennymi. Następne zmienne analogicznie nie są skorelowane z poprzednimi, lecz skorelowane między sobą. Przyjmując, że $q \leqslant p$ to otrzymamy q kanonicznych korelacji i q par zmiennych kanonicznych.

Niech macierz wariancji - kowariancji p + q zmiennych losowych będzie macierzą postaci:

$$\sum = \left[\begin{array}{c|c} \sum_{11} & \sum_{12} \\ \hline \sum_{21} & \sum_{22} \end{array}\right] \tag{4}$$

gdzie:

$\sum_{11}$ - macierz p·q wariancji-kowariancji zmiennych X,
$\sum_{22}$ - macierz q·q wariancji-kowariancji zmiennych Y,
$\sum_{12}$ - macierz p·q wariancji-kowariancji zmiennych X,Y,
$\sum_{21}$ - macierz transponowana.

Korelację pomiędzy zmiennymi kanonicznymi (1) można wyrazić:

$$\varrho_i = \frac{\operatorname{cov}(X\alpha, Y\beta)}{\sqrt{\operatorname{var}(X\alpha)\ \operatorname{var}(Y\beta)}} \tag{5}$$

lub

$$\varrho_i = \frac{\alpha' \sum_{12} \beta}{\sqrt{(\alpha' \sum_{11} \alpha)\ (\beta' \sum_{22} \beta)}} \tag{6}$$

Chcąc obliczyć ϱ_i z (6) należy maksymalizować ϱ_i ze względu na α i β. Przyjmując, że:

$$\mathrm{var}(u_i) = \alpha' \Sigma_{11} \alpha = 1 \tag{7}$$

$$\mathrm{var}(v_i) = \beta' \Sigma_{22} \beta = 1 \tag{8}$$

wtedy

$$\rho_i = \alpha' \Sigma_{12} \beta \tag{9}$$

Zadanie sprowadza się do zmaksymalizowania wyrażenia (9) przy warunkach (7), (8). Funkcja, która ma być zmaksymalizowana ma postać:

$$F(\alpha,\beta) = \alpha' \Sigma_{12} \beta - (\lambda/2)(\alpha' \Sigma_{11} \alpha - 1) - (\mu/2)(\beta' \Sigma_{22} \beta - 1) \tag{10}$$

1/2 - stała

λ, μ - mnożniki Lagrange'a

Różniczkując funkcję $F(\alpha, \beta)$ względem składowych wektorów α i β oraz przyrównując pochodne do zera otrzymujemy:

$$\frac{\partial F}{\partial \alpha} = \Sigma_{12} \beta - \lambda \Sigma_{11} \alpha = 0 \tag{11}$$

$$\frac{\partial F}{\partial \beta} = \Sigma_{21} \alpha - \mu \Sigma_{22} \beta = 0 \tag{12}$$

Mnożąc wyrażenie (11) lewostronnie przez α, natomiast wyrażenie (12) lewostronnie przez β otrzymujemy:

$$\alpha' \Sigma_{12} \beta - \lambda \alpha' \Sigma_{11} \alpha = 0 \tag{13}$$

$$\beta' \Sigma_{12} \alpha - \mu \beta' \Sigma_{22} \beta = 0 \tag{14}$$

Ponieważ

$$\alpha' \Sigma_{11} \alpha = 1 \qquad \beta' \Sigma_{22} \beta = 1$$

to

$$\lambda = \mu = \alpha' \Sigma_{12} \beta' = \rho$$

Zatem wyrażenie (11), (12) można napisać w postaci:

$$-\rho \Sigma_{11} \alpha + \Sigma_{12} \beta = 0 \tag{15}$$

$$\Sigma_{21} \alpha - \rho \Sigma_{22} \beta = 0 \tag{16}$$

Z równania (16) można wyrazić zmienną za pomocą zmiennej następująco:

$$\beta = \rho^{-1} \Sigma_{22}^{-1} \Sigma_{21} \qquad \text{dla } \rho \neq 0 \tag{17}$$

Wstawiając do równania (15) mamy

$$-\rho \Sigma_{11} \alpha + \rho^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \alpha = 0 \tag{18}$$

oraz mnożąc lewostronnie przez Σ_{11}^{-1} otrzymujemy

$$\left(\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \rho^2 I \right) \alpha = 0 \tag{19}$$

Analogicznie postępując z równaniem (16) otrzymujemy kolejno:

$$\alpha = \varrho^{-1} \Sigma_{11}^{-1} \Sigma_{12} \beta \qquad \text{dla } \varrho \neq 0 \qquad (20)$$

$$\left(\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11} \Sigma_{12} - \varrho^2 I \right) \beta = 0 \qquad (21)$$

Aby istniały nietrywialne rozwiązania równań (19) i (21), macierze występujące w tych równaniach winny być osobliwe

$$\left| \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \varrho^2 I \right| = 0 \qquad (22)$$

$$\left| \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \varrho^2 I \right| = 0 \qquad (23)$$

Równanie (22) ma p pierwiastków, natomiast równanie (23) ma q pierwiastków. Niezerowe pierwiastki (22) i (23) są jednakowe, a ich liczba jest równa rzędowi macierzy

$$\Sigma_{12} \ (\text{rząd } \Sigma_{12} = r = \min (p,q))$$

Wektory charakterystyczne $\alpha = \alpha(\alpha_1 | \alpha_2 | \ldots | \alpha_p)$ oraz $\beta = (\beta_1 | \beta_2 | \ldots | \beta_q)$ odpowiadające pierwiastkom równań (22), (23) wyznaczamy z równań (18) i (19), przy czym między tymi wektorami zachodzą związki (17) oraz (20). Pierwiastki równania (23) to $\varrho_1^2 \geqslant \varrho_2^2 \ \ldots \geqslant \varrho_p^2$ a równania (22) to $\varrho_{11}^2 \geqslant \varrho_2^2 \geqslant \cdots \geqslant \varrho_q^2$. Wartości zaś $\varrho_1 > 0$, $\varrho_2 > 0$, ..., $\varrho_r > 0$ nazywa się korelacjami kanonicznymi. Wektory charakterystyczne α i β, których elementy są odpowiednimi wagami dla zmiennych pierwotnych pozwalają oszacować zmienne kanoniczne postaci $u_i = \alpha X$ i $v_i = \beta Y$.

3. Próba egzemplifikacji

Zbiór zmiennych objaśnianych tworzyły koszty przedsiębiorstw hurtowych sklasyfikowane wg następującego schematu:

1. Koszty osobowe w tys. zł (Y_1).
2. Koszty zaangażowanego kapitału w tys. zł (Y_2).
3. Koszty obrotu w tys. zł (Y_3).
4. Koszty utrzymania ruchomości i nieruchomości w tys. zł (Y_4).
5. Koszty pozostałe w tys. zł (Y_5).

W skład zbioru zmiennych objaśniających wchodziły zmienne charakteryzujące wielkość świadczonych usług i rozmiary zaangażowanych czynników wytwórczych przedsiębiorstwa, a mianowicie:

1. Wartość sprzedaży ogółem w mln zł (X_1).
2. Wartość sprzedaży magazynowej w mln zł (X_2).
3. Marża handlowa i prowizje w mln zł (X_3).
4. Liczba hurtowni (X_4).
5. Powierzchnia magazynowa ogółem w m^2 (X_5).
6. Powierzchnia składowa magazynów w m^2 (X_6).
7. Liczba zatrudnionych pracowników ogółem (X_7).
8. Zapas końcowy w tys. zł (X_8).

Tablica 1

Regresyjne modele kosztów
względem potencjalnych zmiennych objaśniających

b_i	Y'_1	Y'_2	Y'_3	Y'_4	Y'_5
b_1	0,1340 (0,2867)	0,5521 (0,7363)	1,2810 (0,7499)	0,4610 (0,8420)	1,1595 (0,2400)
b_2	-0,3597 (0,2687)	-0,5336 (0,6900)	-1,1366 (0,7028)	-0,0379 (0,7891)	-1,0635 (0,2250)
b_3	0,4236 (0,2155)	1,2015 (0,5535)	1,1157 (0,5637)	-0,1794 (0,6329)	0,0619 (0,1804)
b_4	-0,0346 (0,0303)	-0,0130 (0,0779)	-0,0322 (0,0794)	-0,0143 (0,0891)	0,0603 (0,0254)
b_5	0,0253 (0,2435)	-0,0085 (0,6253)	-1,5569 (0,6369)	0,2166 (0,7151)	-0,0666 (0,2039)
b_6	0,0148 (0,2649)	-0,2431 (0,6804)	1,3514 (0,6930)	0,4891 (0,7780)	0,0809 (0,2218)
b_7	0,7445 (0,1650)	0,3536 (0,4238)	0,2186 (0,4317)	0,1084 (0,4847)	0,8600 (0,1382)
b_8	0,0430 (0,0913)	-0,3386 (0,2346)	-0,2717 (0,2390)	-0,0430 (0,2683)	-0,0882 (0,0765)
S_y	0,0556	0,1427	0,1454	0,1632	0,0465
$\mathcal{R}^2$	0,002	0,010	0,010	0,014	0,002

Źródło: Obliczenia własne na podstawie danych zawartych w Banku Danych „Koszty uspołecznionych przedsiębiorstw handlowych" Instytutu Cybernetyki Ekonomicznej, Akademii Ekonomicznej w Poznaniu.

Analiza kanoniczna umożliwia[7] w pierwszej kolejności oszacowanie regresyjnych modeli opisowych dla każdej ze zmiennych objaśnianych (tabl. 1) o postaci:

$$Y_i = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n + \xi$$

[7] Statistical Analysis (12), s. 57-64.

BIBLIOTEKA AE Poznaniu

gdzie:

Y_i - zmienna objaśniana $(i = 1,2,\ldots,5)$

$b_1, b_2, \ldots, b_n$ - cząstkowe współczynniki regresji

$X_1, X_2, \ldots, X_n$ - zmienne objaśniające $(n = 8)$

ξ - składnik losowy

Wyjściowe zbiory zmiennych objaśnianych Y_i $i = 1,2, \ldots, 5$ oraz objaśniających X_i $i = 1,2,\ldots,8$ ulegają w metodzie kanonicznej transpozycji w wyniku której otrzymuje się $r = \min p,q$ par zmiennych kanonicznych.

Tablica 2

Wektory charakterystyczne α
dla poszczególnych zmiennych kanonicznych u_i

α_i	u_1	u_2	u_3	u_4	u_5
α_1	0,1230	-0,7312	-1,0434	-0,2081	-0,2794
α_2	-0,4733	3,3848	-1,4529	10,3923	-12,7055
α_3	0,5744	-3,5325	0,5547	-5,0291	17,6307
α_4	2,2205	-11,0952	-8,5616	-7,0649	5,9841
α_5	-1,5659	5,7026	6,6954	9,0667	-1,8425
α_6	-0,3636	3,0187	-0,5954	0,5019	0,8298
α_7	0,4580	3,4527	-4,2416	-3,3551	-4,8572
α_8	0,0375	-0,1679	8,4668	-4,1978	-4,6534

Źródło: Por. tabl. 1.

Współczynniki α i β, będące elementami składowymi otrzymanych wektorów są ładunkami zmiennych hipotetycznych w poszczególnych zmiennych pierwotnych (tabl. 2, 3). Określają one postać zmiennych kanonicznych, nie można jednak nadać im sensownej interpretacji ekonomicznej.

Tablica 3

Współczynniki β dla poszczególnych zmiennych kanonicznych v_i

β_i	v_1	v_2	v_3	v_4	v_5
β_1	-1,3916	14,1877	5,3339	-0,3365	-1,1755
β_2	0,2512	-3,3797	3,9954	2,2846	-4,4308
β_3	0,2647	-1,9844	1,5072	-4,0847	3,9046
β_4	0,3582	-2,6708	0,5976	5,3396	2,6317
β_5	1,5213	-6,1274	-11,3367	-3,1728	-0,9049

Źródło: Por. tabl. 1.

Redukcja przestrzeni wielowymiarowej zmiennych X i Y do zbioru nowych ukrytych zmiennych daje możliwość wypowiedzenia się o współzależności pomiędzy zbiorami zmiennych oryginalnych poprzez określenie współzależności pomiędzy transponowanymi cechami. Otrzymane współczynniki korelacji ϱ_i (tabl. 4) dla pięciu par zmiennych wskazują na wysokie skorelowanie zmiennych. Weryfikacja istotności współczynników pozwoli na ocenę związku pomiędzy zbiorami interesujących zmiennych, jak również może być podstawą oceny wartości informacyjnej nowych wymiarów o zmiennych oryginalnych.

Tablica 4

Współczynniki korelacji kancnicznej
pomiędzy kolejnymi zmiennymi kanonicznymi

Para zmiennych	ϱ_i	χ^2 Bartletta	χ^2_α $\binom{(p-s)}{(q-s)}$
u_1v_1	0,9998	132,98	555,76
u_2v_2	0,9772	55,76	41,34
u_3v_3	0,9366	27,86	28,87
u_4v_4	0,7408	8,99	16,92
u_5v_5	0,4282	1,82	9,49
F Rao'a = 14,497 $>_{0,05} F_{40,20}$ = 1,84			

Źródło: Por. tabl. 1.

W pierwszej kolejności dokonano kompletnego testowania współczynników korelacji kanonicznej stawiając hipotezę:

$$H_0 : \varrho_1 = \varrho_2 = \dots = \varrho_r = 0$$

wobec hipotezy alternatywnej $H_1 : \varrho_1, \neq \varrho_2 \neq \dots \neq \varrho_r \neq 0$. W charakterze funkcji testowej wykorzystana zostaje Λ Wilksa, gdzie

$$\Lambda = \prod_{i=1}^{r} (1 - \varrho_i^2).$$

Posługując się aproksymacją Rao, obliczona zostaje statystyka F Rao'a[8],

$$F = \frac{1 - \Lambda^{1/s}}{\Lambda^{1/s}} \frac{df_2}{df_1}$$

$$df_1 = p \cdot q$$

$$df_2 = ms - \frac{pq}{2} + 1$$

[8] C.R. Rao, (8), ss. 255.

gdzie:

$m = N - 3/2 - (p + q)//2$

$$s = \sqrt{\frac{p^2 q^2 - 4}{p^2 + q^2 - 5}},$$

która posiada w przybliżeniu rozkład F z df_1 i df_2 stopniami swobody. Jeżeli $F > F_{df_1, df_2}$ odrzucamy hipotezę H_0. W przedstawionej analizie (tabl. 3) hipoteza H_0 zostaje odrzucona wobec tego stawiamy hipotezy cząstkowe poczynając od s ostatnich korelacji kanonicznych:

$$H_0 : \varrho_{r-s+1} = 0, \ \varrho_{r-s+2} = 0, \ \ldots, \ \varrho_r = 0 \qquad (r=\min p,q) \quad (s=1,2,\ldots,r)$$

wobec alternatywnych

$$H_1 : \varrho_{r-s+1} \neq 0, \ \varrho_{r-s+2} \neq 0, \ \ldots, \ \varrho_r \neq 0$$

Weryfikacja hipotez przeprowadzona jest sekwencyjnie przy użyciu testu χ^2 Bartletta[9]. Funkcja testowa ma postać:

$$\chi^2 = -\left(N - 3/2 - (p + q)/2\right) \ln \prod_{i=s+1}^{r} \left(1 - \varrho_1^2\right),$$

która przy prawdziwości H_0 ma w przybliżeniu rozkład χ^2 z (p-s)(q-s) stopniami swobody. Weryfikacja kolejnych współczynników korelacji kanonicznych (tabl. 4) pozwala zauważyć, że rozpatrywany zbiór kosztów reprezentowany przez pięć jednorodnych grup jest w wysokim stopniu wyjaśniony zmiennością rozmiarów działalności przedsiębiorstwa.

Otrzymane współczynniki korelacji mogą być podstawą do obliczenia średniego kwadratowego współczynnika wg wzoru[10]:

$$\bar{\varrho}_i^2 = \frac{1}{\min p,q} \sum \varrho_i^2$$

Jego wartość osiąga 0,7128, a więc zmienność zmiennych objaśnianych wyjaśniona jest w 71,28% zmiennością czynników kosztotwórczych.

Z analizy wynika również (tabl. 4), że dwie pierwsze pary zredukowanych zmiennych zawierają największy ładunek o zmiennych wyjściowych i mogą być wykorzystane w dalszej analizie.

4. Uwagi końcowe

Przedstawiona analiza współzależności pomiędzy dwoma zbiorami cech reprezentowanymi przez koszty i czynniki je określające na przykładzie przedsiębiorstw handlu hurtowego branży spożywczej, miała na celu nie tyle zaprezentowanie poznawczej strony osiągniętych rezultatów a zapoznanie z możliwościami zastosowanej metody, jak dotąd mało rozpowszechnio-

---

[9] H. Ahrens, J. Lauter, (1), s. 29 i nast.

[10] H. Theil, (13), s. 325.

nej w praktyce badawczej. Mimo, że analiza wielowymiarowa zyskuje coraz większe zainteresowanie[11] to wydaje się, że analizie kanonicznej poświęca się zbyt mało uwagi. Nie sposób jednak nie dostrzec jej widocznych zalet, jest ona bowiem uogólnieniem wielu metod statystycznych, umożliwiając syntetyczne spojrzenie na zbiory zmiennych zdefiniowane przez badacza. W prezentowanym opracowaniu był to przekrojowy układ: zmiennych objaśnianych prezentowany przez koszty handlowe; oraz potencjalnych zmiennych objaśniających charakteryzujących rozmiary działalności przedsiębiorstwa. W zależności od wyjściowych układów zmiennych analiza kanoniczna może być metodą dynamicznych badań strukturalno-przestrzennych. Pozwala jednocześnie określić współzależność pomiędzy zbiorami zmiennych, skonstruować regresyjne modele dla każdej z interesujących cech, jak również dokonać redukcji wyjściowych zmiennych.

Analiza kanoniczna w części odnoszącej się do transpozycji przestrzeni wielocechowej koresponduje z metodą głównych komponentów i analizą czynnikową. Można przypisać jej identyczne założenie jak w analizie składowych o tym, że zmienność danej populacji zawarta jest wewnątrz zmiennych stosowanych do ich zdefiniowania. Obie te metody umożliwiają redukcję przestrzeni wielowymiarowej do mniejszej liczby wymiarów, empirycznie nieobserwowalnych. O ile jednak pierwsza z nich dotyczy badania zmienności wewnątrz jednego zbioru cech i można powiedzieć, że jest metodą „wewnętrzną", to analiza kanoniczna dotyczy zależności pomiędzy jednym zbiorem i pewnym zewnętrznym do niego innym zbiorem i można nazwać ją metodą „zewnętrzną". Jeśli metoda głównych składowych nie jest metodą wolnej skali to metoda kanoniczna przeciwnie, nie jest zależna od pomiaru zmiennych i w tym momencie jest zbieżna z większością metod analizy czynnikowej. Z kolei jedna z nowszych, obok czynników alfa i czynników wyobrażalnych, mutacji analizy czynnikowej wykorzystuje idee analizy kanonicznej i nosi nazwę analizy czynnika kanonicznego.

Obok wymienionych zalet przemawiających za stosowaniem prezentowanej metody, nie brak jej mankamentów, które przyczyniają się do tego, że są zwolennicy stosowania każdej z metod w niej zawartych samodzielnie i dochodzenie w ten sposób do rezultatów umożliwiających bardziej jednoznaczną interpretację. Można odnotować więc pewne słabości przedstawionej metody. Oszacowane regresyjne modele opisowe są tylko postaci a posteriori, ze znacznie uboższą zawartością parametrów struktury stochastycznej. Zasadniczą jej słabością, często niedostrzegalną jest ograniczona możliwość interpretacji ładunków kanonicznych. O ile ładunki komponentów oraz czynników posiadają bardziej jednoznaczną interpretację, to z uwagi na odmienne założenia tej metody, wagom zmiennych kanonicznych nie można przypisać wartości poznawczej. Pamiętać również należy, że związek pomiędzy zbiorami zmiennych określony jest w sposób uproszczony poprzez pomiar współzależności dla kombinacji liniowych tych cech. Uzupełnieniem metody kanonicznej, które pozwala częściowo zniwelować jej trudności interpretacyjne jest analiza redundacji, która będzie przedmiotem odrębnego opracowania.

---

[11] R. Zasępa, (14), ss. 179.

Literatura

[1] H. Ahrens, J. Läuter: Wielowymiarowa analiza wariancji, PWN, Warszawa 1979.
[2] T. Anderson: An introduction to multivariate statistical analysis, New York 1958.
[3] P. Bertier, J. Bouroche: Analyse des donnés multidimensionnelles, Paris 1975.
[4] J. Bortz: Statistik für Sozialwissenschaftler, Springer-Verlag, Berlin 1979.
[5] W. Cooley, P. Lohnes: Multivariate data analysis, New York 1971.
[6] H. Gaensslen, W. Schubö: Einfache und komplexe statistische Analyse, Ernst Reinhardt Verlag, München, Basel 1973.
[7] J. W. Hooper: Simultaneous equations and canonical correlation theory, Econometrica 1959 nr 27.
[8] H. Hotelling: Relations between two sets of variates, Biometrica 1936, nr 28.
[9] S. Mejza: Korelacje kanoniczne i ich zastosowanie w badaniach rolniczych, Materiały piątego Colloquium Metodologicznego z Agro-biometrii, Warszawa 1975.
[10] Praca zbiorowa pod red. Z. Chojnickiego: Analiza regresji w geografii, PWN, Warszawa-Poznań, 1980.
[11] C. R. Rao: Linear statistical inference and its applications, J. Wiley, New York, 1966.
[12] Statistical Analysis Mark 2, ICL, 1900 Series, Wrocławskie Zakłady Elektroniczne „Elwro", 1970.
[13] H. Theil: Zasady ekonometrii, PWN, Warszawa, 1979.
[14] R. Zasępa: Elementy wielowymiarowej analizy statystycznej, SGPiS, Warszawa 1981.

APPLICATION OF THE METHOD OF CANONICAL CORRELATION

Summary

The paper presents hitherto not widely used in practice, method of canonical analysis which is a generalized form of the analysis of correlation and regression. On the example of the state re-tail food enterprises, apart from a complex analysis, a synthetic analysis of the interdependencies between the costs of the stadied enterprises and the factors which characterize their variability was presented. Initial set of egzogenic variables, represented by five relatively homogenous costs' complexes and the set of endogenic variables which include variables characteristic for the volume of the services rendered and volume of produc-

tive factors, were subjected to simultaneous linear transformation obtaining sequentially pairs of hypothetical canonical variables. These variables are the basis for determining generalized interdependence relation verified by means of F Rao test and χ^2 Bartlett's.

As a result, the presented method corresponding to principle components method and factor analysis method, when supplemented by redundance analysis can display considerable advantages in identifying cause-effect relations. This requires further empirical verifications.