ZESZYTY PROBLEMOWE POSTĘPÓW NAUK ROLNICZYCH 1983 z. 290

PROPOZYCJA ADAPTACJI ANALIZY KANONICZNEJ DO OCENY RODÓW ZIEMNIAKA

Maria Parlińska, Leokadia Ubysz-Borucka

Instytut Zastosowań Matematyki i Statystyki SGGW-AR w Warszawie

Podczas selekcjonowania rodów ziemniaka hodowcy oceniają materiał hodowlany pod względem wielu cech. Na ogół w doświadczeniach porównawczych z rodami uwzględnia się opracowanie każdej ważniejszej cechy osobno, a następnie dokonuje się syntezy w celu uzyskania kompleksowej oceny poszczególnych rodów. Potrzebę przeprowadzania wielocechowej oceny odmian rodów ziemniaka przed ich wprowadzaniem do uprawy podkreślają w swoich publikacjach Nawrocki [11], Simmonds [14], Dmochowski [4] i inni.

W początkowej fazie cyklu hodowlanego są również przydatne metody umożliwiające porównanie i ocenę populacji wielocechowych. Ocena taka ułatwia hodowcy przeprowadzenie selekcji materiałów hodowlanych.

Jedną z metod analizy wielowymiarowych zmiennych jest analiza kanoniczna, niezwykle przydatna w badaniach rolniczych, gdzie często o zmienności kilku cech należy wnioskować na podstawie innego układu zmiennych.

Zagadnienie to polega na ustaleniu zależności między dwoma wektorami zmiennych X i Y. Składowe tych wektorów są na ogół skorelowane, dlatego też analiza każdej cechy osobno może być mało efektywna.

Badacz dąży do otrzymania pewnych optymalnych kombinacji liniowych dla obu układów zmiennych i skupia uwagę tylko na tych kombinacjach, które mają decydujące znaczenie w dalszych badaniach. Ogólnie analiza kanoniczna polega na przekształcaniu układów oryginalnych zmiennych losowych w nowe kombinacje, wewnątrz których zmienne losowe nie są skorelowane, natomiast korelacja między zmiennymi losowymi z różnych kombinacji jest maksymalna. Konstruowanie kombinacji o największym ich wzajemnym skorelowaniu prowadzi do pewnej redukcji zmiennych przy jednoczesnym zachowaniu możliwie jak największej informacji o reakcjach zachodzących między wyróżnionymi układami zmiennych. Jako miarę związku między dwoma zbiorami zmiennych przyjmuje się współczynnik korelacji kanonicznej, który określa stopień zależności.liniowej między dwiema funkcjami liniowymi obu zbiorów zmiennych.

Metoda analizy kanonicznej może mieć zastosowanie w następujących szczególnych przypadkach:

- pewne zmienne (cechy) są łatwiej obserwowalne od innych;
- wartości pewnych zmiennych są uzyskiwane przez pomiary niszczące; zachodzi więc konieczność ograniczenia się tylko do tych cech, których pomiary niszczące powodują najmniejsze straty.

Pewnym problemem analizy kanonicznej jest interpretacja zmiennych kanonicznych, które są liniowymi kombinacjami zmiennych z poszczególnych zbiorów i dlatego w większości przypadków są sztuczne, a zatem trudno im nadać jakiś sens fizyczny.

Po raz pierwszy analizę kanoniczną zastosował Hotelling [6] w psychologii, badając związek zachodzący między zbiorem charakterystyk umysłowych a zbiorem pomiarów fizycznych. Później stosowało ją wielu uczonych, w takich dziedzinach nauk jak np. w ekonomii do szacowania relacji podaży i popytu w czasie, antropologii, botanice i innych.

W Polsce metodę tę wykorzystał Mejza [10] do oceny okrywy włosowej zwierząt futerkowych. Również Krzyśko i Ratajczak [9] posłużyli się tą metodą podczas ustalania zależności między urbanizacją a uprzemysłowieniem w układzie gmin województwa poznańskiego.

Celem niniejszej pracy jest przedstawienie w ogólnych zarysach teorii analizy kanonicznej dla dwóch zbiorów zmiennych.

ZMIENNE KANONICZNE I KORELACJA KANONICZNA

Niech $x = (x_1, x_2, \ldots x_p)'$ i $y = (y_1, y_2, \ldots y_g)'$ będą wektorami kolumnowymi p i g zmiennych losowych, gdzie $x_1, x_2, \ldots x_p$ - zmienne objaśniające i $y_1, y_2, \ldots y_g$ - zmienne objaśniane. Macierz kowariancji tych (p + g) zmiennych losowych będzie miała postać:

$$\begin{array}{cc} \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right] & \begin{array}{c} p \\ g \end{array} \\ \begin{array}{cc} p & \quad g \end{array} & \end{array} \tag{1}$$

gdzie: $\Sigma_{12}' = \Sigma_{21}$ (x' - transpozycja):

$$\text{np.} \quad \Sigma_{11} = \begin{bmatrix} \text{war. } x_1x_1, \text{ kowar. } x_1x_2, \ldots, \text{kowar. } x_1x_p \\ \text{kowar. } x_2x_1, \text{ war } x_2x_2, \ldots, \text{kowar. } x_2x_p \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \text{kowar. } x_px_1, \text{ kowar. } x_px_2, \ldots, \text{war. } x_px_p \end{bmatrix} \quad \text{itd.}$$

Zakłada się, że macierz kowariancji [1] jest dodatnio określona. Niech rząd $\Sigma_{12}' = s$, gdzie $s \leqslant \min (p, g)$.

Rozpatrzmy funkcję liniową zmiennych x o postaci $u = l_1x_1 + l_2x_2 + \dots + l_px_p$ oraz zmiennych y o postaci $v = m_1y_1 + m_2y_2 + \dots + m_gy_g$. Funkcje powyższe można zapisać krócej w postaci wektorowej:

$$u = l'x \text{ oraz } v = m'y$$

Współczynniki l_i oraz m_i w powyższych funkcjach wyznacza się tak, aby korelacja między zmiennymi u i v była maksymalna. Obliczając współczynnik korelacji między tymi zmiennymi uzyskamy:

$$\text{war. } u = E\,[l'x - E\,(lx)\,]^2 = l'\Sigma_{11}l^{*} \tag{2}$$

$$\text{war. } v = E\,[m'y - E\,(m'y)]^2 = m'\Sigma_{22}m \tag{3}$$

$$\text{kowar. } (u, v) = E\,[l'x - E\,(l'x)\,]\,[\,m'y - E\,(m'y)\,] = l'\Sigma_{12}m \tag{4}$$

Stąd:

$$\text{kor. } (u, v) = \rho = \frac{l'\Sigma_{12}m}{\sqrt{(l'E_{11}l)\ (m'\Sigma_{22}m)}} \tag{5}$$

Wiadomo, że współczynnik korelacji między zmienną różniącą się od zmiennej u tylko stałym mnożnikiem i zmienną różniącą się od zmiennej v stałym mnożnikiem jest równy współczynnikowi korelacji między zmiennymi u i v. Korzystając z powyższego wektory l i m można więc normować dowolnie. Stąd też dla wygody obliczeń l i m dobiera się tak, aby:

$$\text{war. } (l'x) = l'\Sigma_{11}\,l = 1 \tag{6}$$

$$\text{war. } (m'y) = m'\Sigma_{22}\,m = 1 \tag{7}$$

Wtedy współczynnik korelacji między zmiennymi u i v jest równy:

$$\rho = \text{kor. } (u, v) = l'\Sigma_{12}m \tag{8}$$

---

*Można wykazać po przekształceniach algebraicznych.

Zadanie sprowadza się więc do zmaksymalizowania wyrażenia [8] przy dodatnich warunkach [6] i [7].

Funkcja, która ma być maksymalizowana ma postać:

$$F\ (l,\ m) = l' \sum_{12} m - \frac{\lambda}{2}\ (l' \sum_{11} l - 1) - \frac{\mu}{2} (m' \sum_{22} m - 1) \qquad (9)$$

gdzie: λ i μ są mnożnikami Lagrange'a

Po zróżniczkowaniu funkcji F (l, m) względem składowych wektorów l i m oraz przekształceniach algebraicznych, otrzymuje się więc układ równań o postaci:

$$\begin{cases} \left(\sum_{11}^{-1} \sum_{12} \sum_{22}^{-1} \sum_{21} - \rho^2 J\right) l = 0 \\ \left(\sum_{22}^{-1} \sum_{21} \sum_{11}^{-1} \sum_{12} - \rho^2 J\right) m = 0 \end{cases} \qquad (10)$$

Aby istniały nietrywialne rozwiązania obu tych równań macierze w nich występujące powinny być osobliwe tzn.:

$$\begin{aligned} \left| \sum_{11}^{-1} \sum_{12} \sum_{22}^{-1} \sum_{21} - \rho^2 J \right| &= 0 \\ \left| \sum_{22}^{-1} \sum_{21} \sum_{11}^{-1} \sum_{12} - \rho^2 J \right| &= 0 \end{aligned} \qquad (11)$$

Pierwsze z tych równań ma co najwyżej p-pierwiastków, drugie zaś co najwyżej g-pierwiastków. Niezerowe pierwiastki tych równań są jednakowe, więc można je oznaczyć jednakowymi symbolami. Liczba niezerowych pierwiastków tych równań jest równa rzędowi macierzy $\sum_{12}$.

Wektory charakterystyczne odpowiadające pierwiastkom tych równań wyznacza się z układu, pamiętając o związkach zachodzących między nimi:

Niech $\rho_1^2 \geqslant \rho_2^2 \geqslant \rho_3^2 \geqslant \ldots \geqslant \rho_p^2$ będą pierwiastkami pierwszego równania, zaś l_1, l_2, ... l_p odpowiadającymi im wektorami charakterystycznymi. Analogicznie $\rho_1^2 \geqslant \rho_2^2 \geqslant \rho_3^2 \geqslant \ldots \geqslant \rho_g^2$ - pierwiastki drugiego równania, zaś $m_1, m_2, \ldots, m_g$ - odpowiadające im wektory charakterystyczne.

Zmienne $u_i = l_i'x$ nazywane są *zmiennymi kanonicznymi przestrzeni* x-ów zaś $v_i = m_i'y$ *zmiennymi kanonicznymi przestrzeni* y-ów.

Niezerowe wartości $\rho_1, \rho_2, \ldots, \rho_s$ będące dodatnimi pierwiastkami kwadratowymi z $\rho_1^2, \rho_2^2, \ldots, \rho_s^2$ nazywane są *korelacjami kanonicznymi*.

WŁASNOŚCI ZMIENNYCH KANONICZNYCH

1. Wariancje wszystkich zmiennych u_i oraz v_i są równe 1, stąd kowariancje są równe korelacjom.

2. Pierwsza zmienna kanoniczna przestrzeni x-ów jest skorelowana jedynie z pierwszą zmienną kanoniczną przestrzeni y-ów, druga z drugą itd. Korelacje te są odpowiednio równe:

$$\rho_1, \rho_2, \ldots, \rho_s$$

3. Korelacje kanoniczne zmiennych x i y oraz zmiennych unormowanych są identyczne.

4. Można wykazać, że:

$$\text{war. } x_i = \sum_{j=1}^{p} \text{kowar.}^2 (X_i, U_j)$$

$$\text{war. } y_k = \sum_{j=1}^{g} \text{kowar.}^2 (y_k, v_j)$$

Wzory te stanowią interpretację dla zmiennych kanonicznych. Widać, że zmienne kanoniczne w różnym stopniu wyjaśniają wariancję zmiennych pierwotnych. Miarą tego wyjaśnienia jest kwadrat kowariancji między zmienną kanoniczną i zmienną pierwotną.

5. Miarą zależności liniowej między funkcjami obu układów zmiennych jest tzw. złożony współczynnik determinacji. Jest on zdefiniowany jako stosunek sumy wariancji wyjaśnionych zmiennych $y_1 \dots y_g$ do sumy wariancji całkowitej tych zmiennych. Można go zapisać:

$$R^2_{y,x} = \sum_{i=1}^{s} \frac{\sum_{k=1}^{g} \rho_i^2 \text{ war. } y_k \cdot \text{kor.}^2(y_k, v_i)}{\sum_{k=1}^{g} \text{war. } y_k}$$

gdzie:

ρ_i jest i-tym pierwiastkiem charakterystycznym z próby.

Poszczególne składniki tej sumy można interpretować następująco: i-ty składnik sumy jest frakcją sumy wiariancji wyjaśnionych przy pomocy i-tej zmiennej kanonicznej przestrzeni y-ów.

6. Ocena zmienności zmiennych zależnych (objaśnianych) $y_1 \dots y_g$ za pomocą wszystkich zmiennych kanonicznych przestrzeni x-ów sprowadza się do oceny za pomocą tylko s-zmiennych kanonicznych odpowiadających niezerowym pierwiastkom równania. Tym samym o zmiennych kanonicznych odpowiadających zerowym pierwiastkom można powiedzieć, że nie wnoszą żadnej informacji, stąd więc można się ograniczyć tylko do s pierwszych zmiennych kanonicznych.

PRZYKŁAD ZASTOSOWANIA ANALIZY KANONICZNEJ DO OCENY RODÓW ZIEMNIAKA

Zastosowano analizę korelacji zmiennych kanonicznych do określenia zależności między dwiema grupami cech ziemniaka. Eksperymentalne dane dotyczyły plonów uzyskanych w trzech terminach wczesności zbioru oraz oceny morfologicznej regularności kształtu i oceny głębokości oczek rodów bardzo wczesnych. Wyniki doświadczeń wstępnych z 1979 r. uzyskano z Instytutu Ziemniaka w Boninie. W doświadczeniu badano 12 rodów bardzo wczesnych oraz 2 odmiany wzorcowe Irys i Frezja pochodzące z dwóch źródeł rozmnożenia (Bonin i Zamarte). Doświadczenia przeprowadzono metodą układów kratowych w 5 miejscowościach. Do analizy użyto średnich plonów z doświadczeń. Łączna liczba wyników użytych w analizie kanonicznej wynosiła n = 80. Tych 80 wyników dla 5 obserwowanych zmiennych traktowano jako wielowymiarową próbę losową dla populacji ziemniaków bardzo wczesnych.

Zmiennymi objaśnionymi y_1, y_2, y_3 są tutaj odpowiednio plony w I terminie, II terminie i III terminie w t z ha, zaś zmiennymi objaśniającymi x_1, x_2 są oceny regularności kształtu i oceny głębokości oczek wyrażone w stopniach bonitacyjnych.

W tabeli 1 podane są średnie z rodów i miejscowości dla poszczególnych cech.

Tabela 1

Średnie z rodów i miejscowości dla analizowanych cech

$\bar{x}_1$	$\bar{x}_2$	$\bar{y}_1$	$\bar{y}_2$	$\bar{y}_3$
5,97	6,33	8,774	16,929	28,065

Macierz kowariancji ma postać:

$$
\left[\begin{array}{cc|ccc}
4,77 & 3,43 & -357,66 & -344,51 & -709,21 \\
3,43 & 16,18 & 19,05 & 590,67 & 493,79 \\
\hline
-357,66 & 19,05 & 226921,39 & 244721,75 & 388757,11 \\
-344,51 & 590,67 & 244721,75 & 341650,99 & 475527,69 \\
-709,21 & 493,79 & 388757,11 & 475527,69 & 817519,91
\end{array}\right]
$$

Do obliczenia kanonicznych zmiennych i korelacji można użyć również macierzy korelacji. Macierz współczynników korelacji jest macierzą blokową:

$$
\left[\begin{array}{c|c}
R_{11} & R_{12} \\
\hline
R_{21} & R_{22}
\end{array}\right]
$$

W przedstawionym doświadczeniu ma ona postać:

$$
\left[\begin{array}{cc|ccc}
1 & 0,30 & -0,34 & -0,26 & -0,35 \\
0,39 & 1 & 0,00 & 0,25 & 0,13 \\
\hline
-0,34 & 0,00 & 1 & 0,87 & 0,90 \\
-0,26 & 0,25 & 0,87 & 1 & 0,89 \\
-0,35 & 0,13 & 0,90 & 0,89 & 1
\end{array}\right]
$$

Z warunku [11] otrzymujemy pierwiastki charakterystyczne:

$$\rho_1^2 = 0,3122; \quad \rho_2^2 = 0,1401$$

Stąd współczynniki korelacji kanonicznej są następujące:

$$\rho_1 = 0,5588; \quad \rho_2 = 0,3743$$

Zmienne kanoniczne odpowiadające współczynnikowi korelacji kanonicznej ρ_1 mają postać:

$$u_1 = 0{,}4892\, x_1 - 0{,}8722\, x_2$$

$$v_1 = 0{,}6900\, y_1 - 0{,}6528\, y_2 - 0{,}3127\, y_3$$

oraz dla ρ_2 postać:

$$u_2 = 0{,}9769\, x_1 + 0{,}2136\, x_2$$

$$v_2 = -0{,}5705\, y_1 + 0{,}6683\, y_2 - 0{,}4773\, y_3$$

Aby zbadać istotność statystyczną zależności obu zespołów zmiennych w przypadku gdy $s = p = 2$ lub $s = g = 2$ można podać dokładne testy [1]. Hipotezę o braku zależności między zbiorami zmiennych odrzucamy na obranym poziomie istotności wtedy i tylko wtedy, gdy wartość funkcji testowej $F_{emp.}$ przekroczy odpowiednią wartość graniczną z rozkładu F Fishera-Snedecora. Zmienne $F_{emp.}$ ma postać:

$$F_{emp_1} = \frac{(1 - \sqrt{\lambda})\ (n-g-1)}{\sqrt{\lambda}\ g} \quad \text{gdy } p = 2$$

$$F_{emp_2} = \frac{(1 - \sqrt{\lambda})\ (n-p-1)}{\sqrt{\lambda}\ p} \quad \text{gdy } g = 2$$

gdzie: λ jest statystyką Wilksa określoną następująco:

$$\lambda = \prod_{i=1}^{s} (1 - \rho_i^2)$$

Zmienna F_{emp_1} ma liczbę stopni swobody $n_1 = 2g$ i $n_2 = 2(n - g - 1)$ zaś zmienna F_{emp_2} odpowiednio $n_1 = 2p$ i $n_2 = 2\ (n - p - 1)$.

W przedstawianym przykładzie z dokładnego testu dla którego $F_{emp_1} = 7,51$ przy wartości granicznej F ($\alpha = 0,05$, $n_1 = 6$, $n_2 = 150$) $= 2,16$ wynika, że korelacja kononiczna między obu zbiorami cech jest istotna, czyli bądź oba współczynniki korelacji są istotne bądź większy z nich.

Dla szczegółowego sprawdzenia istotności korelacji między zmiennymi kanonicznymi u_2 i v_2, o mniejszym współczynniku korelacji kanonicznej $\rho_2 = 0,3743$, zastosujemy funkcję testową o rozkładzie χ^2.

Funkcja ta ma postać:

$$\chi^2_{emp.} = -\left(n - \frac{p + g + 3}{2}\right) \ln (1 - \rho_k^2) \ldots (1 - \rho_p^2)$$

gdzie:

k - określa ostatnie korelacje nieistotne.

Tutaj k jest równe 2. Zmienna $\chi^2_{emp.}$ ma liczbę stopni swobody $n_3 = (p - k + 1)\ (g - k + 1)$.

W przypadku gdy $\chi^2_{emp.} \geq \chi^2_{1-\alpha}$ hipotezę odrzucamy [3].

W rozpatrywanym doświadczeniu otrzymano $\chi^2_{emp.} = 11,47$ przy $\chi^2_{(\alpha = 0,05; n_3 = 2)} = 9,49$ więc należy wnioskować, że korelacja między zmiennymi u_2 i v_2 jest istotna.

Obie pary zmiennych kanonicznych wnoszą więc istotne informacje o zależności między obu grupami zmiennych.

Funkcje u_1, u_2, v_1 i v_2 można traktować jako funkcje dyskryminacyjne [9] i wykorzystać je do uporządkowania obiektów rozpatrywanych w przestrzeni x-ów u_1 i u_2 i przestrzeni y-ów v_1 i v_2.

Średnie rzeczywiste rodów bardzo wczesnych ziemniaków dla podanych cech oraz średnie unormowane podano w tabelach 2 i 3.

Tabela 2

Średnie rzeczywiste

Badane rody	Ocena		Plony w terminie		
	regularności kształtu x_1	głębokości oczek x_2	I y_1	II y_2	III y_3
B-34501	5,9	5,9	76,9	141,3	247,4
B-34503	6,0	6,4	72,9	144,9	268,8
B-34547	6,0	6,4	97,9	175,3	274,9
B-34896	5,8	6,0	56,9	134,2	259,6
P-72.610-5	5,9	6,1	117,9	182,5	274,9
B-Irys	6,0	6,7	114,9	191,4	314,6
Z-Frezja	6,1	6,6	83,7	167,5	296,2
Z-51695	6,0	6,8	120,2	183,7	290,1
Z-66144	5,8	6,2	110,3	169,3	287,1
Z-66188	6,0	6,4	43,3	109,9	216,1
Z-67092	5,9	6,5	102,4	181,9	319,7
Z-67231	5,9	5,9	130,0	203,5	287,1
Z-67224	5,9	6,1	124,1	189,1	290,1
Z-67228	6,1	6,1	112,3	183,7	275,3
Z-Irys	6,1	6,6	112,3	198,1	316,7
Z-Frezja	6,2	6,6	84,7	162,1	275,3

T a b e l a 3

Średnie unormowane

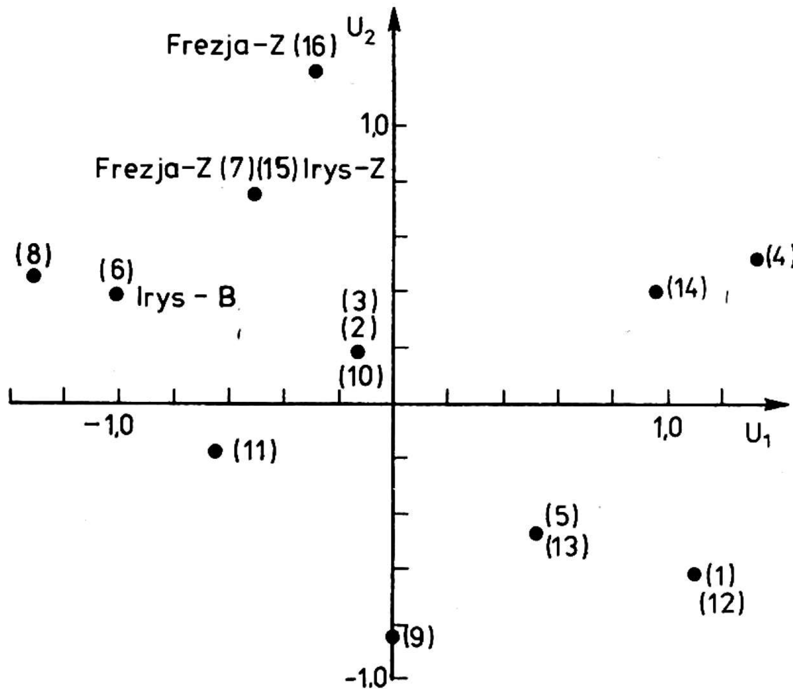
Badane rody	x_1	x_2	y_1	y_2	y_3
B-34501	-0,31	-1,43	-1,05	-1,34	-1,16
B-34503	0,13	0,23	-1,25	-1,16	-0,42
B-34547	0,13	0,23	0,01	0,29	-0,20
B-34896	0,76	-1,10	-2,06	-1,68	-0,74
P.72.610-5	-0,31	-0,77	1,02	0,63	-0,20
B-Irys	0,13	1,23	0,86	1,06	1,19
Z-Frezja	0,58	0,90	-0,71	-0,09	0,54
Z-51695	0,13	1,57	1,13	0,69	0,33
Z-66144	-0,76	-0,43	0,63	0,00	0,23
Z-66188	0,13	0,23	-2,74	-2,84	-2,26
Z-67092	-0,31	0,57	0,23	0,60	1,37
Z-67231	-0,31	-1,43	1,62	1,63	0,23
Z-67224	-0,31	-0,77	1,33	0,95	0,33
Z-67228	0,58	-0,77	0,73	0,69	-0,19
Z-Irys	0,58	0,90	0,73	1,38	1,26
Z-Frezja	1,03	0,90	-0,66	-0,34	-0,19

Odpowiednie wartości funkcji dyskryminacyjnych dla unormowanych średnich podano w tabeli 4.

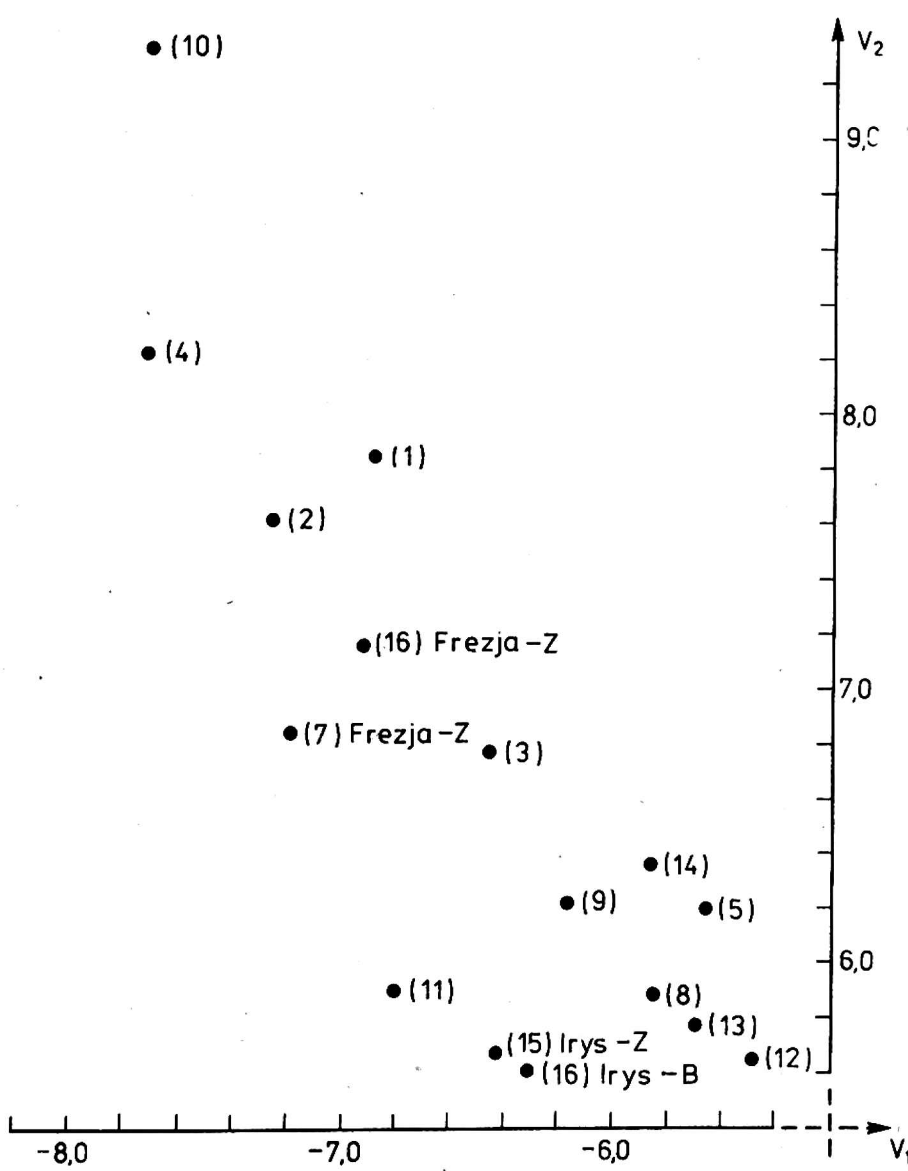
Tabela 4

Wartości funkcji u_i i v_i dla średnich unormowanych

Badane rody	u_1	u_2	v_1	v_2
B-34501	1,1	-0,61	-6,89	7,84
B-34503	-0,14	0,18	-7,26	7,60
B-34547	-0,14	0,18	-6,46	6,77
B-34896	1,33	0,51	-7,72	8,21
P.72.610-5	0,52	-0,47	-5,76	6,20
B-Irys	-1,01	0,39	-6,31	5,62
Z-Frezja	-0,50	0,76	-7,19	6,83
Z-51695	-1,31	0,46	-5,85	5,88
Z-66144	0,00	-0,83	-6,17	6,21
Z-66188	-0,14	0,18	-7,71	9,32
Z-67092	-0,65	-0,18	-6,80	5,90
Z-67231	1,10	-0,61	-5,48	5,65
Z-67224	0,52	-0,47	-5,71	5,77
Z-67228	0,96	0,40	-5,96	6,36
Z-Irys	-0,50	0,76	-6,42	5,67
Z-Frezja	-0,28	1,20	-6,92	7,15



Rys. 1. Zmienne kanoniczne przestrzeni x-ów



Rys. 2. Zmienne kanoniczne przestrzeni y-ów

Na rysunkach 1 i 2 podano rozmieszczenie badanych rodów w układzie osi współrzędnych u_1 i u_2 (cechy morfologiczne) oraz v_1 i v_2 (plony).

PODSUMOWANIE WYNIKÓW

1. Na podstawie przeprowadzonej analizy kanonicznej można wnioskować, że oceny regularności kształtu i głębokości oczek są skorelowane z plonami ziemniaków populacji rodów bardzo wczesnych.

2. Przyjmując arbitralne kryteria np. według wprowadzonych wzorców, można uzyskać podział obiektów badanych na pewne podgrupy pod względem łącznej oceny cech objaśniających i cech objaśnianych.

LITERATURA

1. Anderson T. W.: An introduction to multivariate statistical analysis. New York, Wiley, 1958.
2. Bertier P., Bouroche J. M.: Analyse des données multidimensionnelles, Presses Universitaires de France, Paris, 1977.
3. Dagnelle P.: Analyse statistique a plusieurs variables. Les Presses Agronomiques de Gembloux, Beligique, 1977.
4. Dmochowski K.: Uwagi o metodach i kryteriach oceny wartości gospodarczej odmian ziemniaka. Biul. Oceny Odmian, 4: 97-111, 1973.
5. Gnanadesikan R.: Methods for statistical date analysis of multivariate observations. John Wiley and Sons, New York, 1977.
6. Hotelling H.: Relations between two sets of variates. Biometrika, 28: 321-377, 1936.
7. Kadłubiec W.: Wartość informacyjna cech użytkowych mieszańców F_2 pszenicy ozimej. Praca doktorska z Instytutu Hodowli Roślin i Nasiennictwa, AR Wrocław 1979.
8. Kettenring J. R.: Canonical analysis of several sets of variables. Biometrika, 58: 433-452, 1971.

9. Krzyśko M., Ratajczak W.: Analiza kanoniczna. Listy Biometryczne, 65, 66, 67, Wrocław, 1979.
10. Mejza S.: Korelacje kanoniczne i ich zastosowanie w badaniach rolniczych. Materiały V Colloqium Metodologicznego z Agrobiometrii, 1975.
11. Nawrocki Z.: Studia biometryczno-hodowlane. Cz. I, O metodzie dyskryminacji wielocechowych populacji hodowlanych buraków cukrowych. Ann. UMCS, Lublin, 6: 309-318, 1951.
12. Nowosadzki M.: Analiza kanoniczna i analiza redundacji. Materiały V Colloqium Metodologicznego z Agrobiometrii, 1975.
13. Pillai K. C. S., Gupta A. K.: One the exact distribuation of Wilks's criterion. Biometrika, 56 109-118, 1969.
14. Simmonds N. W.: Prospects of potato improvement. Ann. Rep.: 18-38, 1969.
15. Ubysz-Borucka L.: Fenotypowa zmienność ziemniaka. Zesz. Probl. Post. Nauk. Rol., 191: 283-295, 1977.

Мария Парлиньска, Леокадия Убыш-Боруцка

ПРЕДЛОЖЕНИЕ ПРИСПОСОБЛЕНИЯ КАНОНИЧЕСКОГО АНАЛИЗА К ОЦЕНКЕ ЛИНИЙ КАРТОФЕЛЯ

Резюме

В ст тье приводится общий обзор теоретических основ оценки одного из методов анализа многомерных изменчивых, или т.наз. канонического анализа. Основной целью было использование канонического анализа в оценке линий картофеля. Экспериментальные данные касались урожаев полученных в 3 срока ранности уборки, а также морфологической оценки формы и глубины глазков у очень ранних линий. Результаты предварительных испытаний проведенных в 1979 г. были получены из Института картофелеводства в Бонине. В опытах испытывали 12 очень ранних линий и 2 образцовых сорта: Ирис и Фрезия, полученных из двух источников размножения (Бонин и Замарте). Многомерный случайный образец для популяции очень раннего картофеля составляли 5 охваченных наблюдениями признаков (изменчивых) в 80 повторениях. С помощью метода канонического анализа доказано существование зависимости между обеими группами признаков, т.е. морфологическими признаками клубней и урожаями картофеля. Максимальный коэффициент канонической корреляции составляет $\rho_1 = 0,5588$. В оценке зависимости между обеими группами признаков использовали две пары канонических изменчивых. Коэффициент детерминации урожаев при учете морфологических признаков составляет 45%. Использование канонических изменчивых как дискриминационных изменчивых позволило разделить испытуемые объекты на некоторые подгруппы.

Maria Parlińska, Leokadia Ubysz-Borucka

PROPOSAL OF ADAPTATION OF THE CANONICAL ANALYSIS TO ESTIMATION OF POTATO STRAINS

S u m m a r y

Theoretical fundamentals of one of the methods of analysis of multidimensional variables, i.e. so-called canonical analysis, are outlined in the paper. A main purpose was to apply the canonical analysis in the estimation of potato strains. Experimental data concerned the yields obtained at 3 har vest earliness dates and the morphological regularity of the shape and of eye depth of very early strains. Results of the preliminary experiments carried out in 1979 were obtained from the Institute of Potato Cultivation at Bonin. In the experiments 12 very early strains and 2 pattern varieties: Irys and Frezja, obtained from two multiplication sources (Bonin and Zamarte) were tested. A multidimensional random sample for the population of very early potatoes constituted 5 features (variable) under observations in 80 replications. The method of canonical analysis proved a relationship between both groups of the features, i.e. morphological estimates of tubers and potato yields. The maximum coefficient of canonical correlation was $\rho_1 = 0.5588$. In estimation of the relationship between both groups of the features two couples of canonical variables were used. The yield determination coefficient under consideration of morphological features amounts to 45%. While using the canonical variables as discrimination variables, the objects tested could be divided into some subgroups.