

WYBRANE PROBLEMY STATYSTYKI WIELOWYMIAROWEJ

18. Stu uczniów zdawało 5 egzaminów: z fizyki (Z), biologii (Z), geometrii (O), algebry (O) i statystyki (O) których wyniki były podane w skali 1-100. Litera przy nazwie egzaminu oznacza, czy egzamin był przeprowadzony przy otwartym (O), czy zamkniętym podręczniku (Z). Składowe główne i ich wariancje podane są w poniższej tabeli:

	Współczynniki składowych głównych				
	1	2	3	4	5
fizyka (Z)	0,51	0,75	-0,30	0,30	0,08
biologia (Z)	0,37	0,21	0,42	-0,78	0,19
geometria (O)	0,35	-0,08	0,15	-0,00	-0,92
algabra (O)	0,45	-0,30	0,60	0,52	0,29
statystyka (O)	0,53	-0,55	-0,60	-0,18	0,15
Wariancja	679,2	199,8	102,6	83,7	31,8

- (a) Jaka część całkowitej zmienności jest wyjaśniona przez pierwsze dwie składowe główne?
 (b) Narysuj zmienne oryginalne w kole korelacyjnym dla dwóch pierwszych składowych głównych, zakładając, że wszystkie zmienne oryginalne mają tą samą wariancję.
 (c) Spróbuj nazwać i opisać znaczenie pierwszych dwóch składowych głównych posługując się kołem korelacyjnym

19. Wyznacz dwie pierwsze składowe główne dla macierzy S z zadania 17 (lista 4). Zinterpretuj te składowe, analizując współczynniki korelacji między oryginalnymi zmiennymi i składowymi głównymi. Narysuj zmienne oryginalne w kole korelacyjnym dla dwóch pierwszych składowych głównych.

20. Wyznacz macierz korelacji K odpowiadającą macierzy S z zadania 17, a następnie dla tej macierzy powtórz zadanie 18. Czy są różnice między składowymi głównymi i ich interpretacją dla macierzy S i K?

21. Korzystając ze wzoru $S = \sum_{j=1}^4 l_j a^{(j)} a^{(j)'}$, gdzie l_j jest j-tą wartością własną, a $a^{(j)}$ j-tym wektorem własnym macierzy S, wyznacz kolejne przybliżenia macierzy kowariancji S z zadania 17 i ich macierze resztowe: $S_k = \sum_{j=1}^k l_j a^{(j)} a^{(j)'}$ i $R_k = S - S_k$ dla $k=1,2$

22. Dane są fragmentem badań longitudinalnych przeprowadzonych na grupie dzieci. Podane są wartości ilorazu inteligencji IQ dla 6 dzieci w wieku 2,4,8 i 13 lat:

dziecko	1	2	3	4	5	6
2_lata	120	131	126	120	126	125
4_lata	115	109	115	102	125	109
8_lat	109	113	113	111	114	96
13_lat	106	95	90	121	96	87

- (a) Oblicz macierz kowariancji dla tych danych. Wyznacz dwie pierwsze składowe główne dla tej macierzy. Zinterpretuj te składowe, analizując współczynniki korelacji między oryginalnymi zmiennymi i składowymi głównymi. Narysuj zmienne oryginalne w kole korelacyjnym dla dwóch pierwszych składowych głównych.
 (b) Wyznacz współrzędne dla dzieci 1,5 i 6 w układzie pierwszych dwóch składowych głównych.

Wskazówka.

Wartości własne i wektory własne można obliczyć w programie Matlab (funkcja eig), Mathematica (funkcja Eigensystem) lub wykorzystując procedurę w programie Statistica