

Algorytm metody osi głównych bazuje na zredukowanej macierzy korelacji. W celu zbudowania macierzy zredukowanej wprowadza się na głównej przekątnej kwadraty korelacji wielorakiej kolejnych zmiennych z pozostałymi zmiennymi lub jedynki na przekątnej macierzy zredukowanej zastępuje się modułami najwyższych korelacji z kolejnych wierszy macierzy korelacji. Wartości wprowadzone na przekątną macierzy zredukowanej nazywane są zasobami zmienności wspólnej poszczególnych zmiennych.

Metoda największej wiarygodności dąży do wyznaczenia ładunków czynnikowych dla każdej zmiennej, przy założeniu, że oszacowane parametry modelu w najlepszy sposób będą odzwierciedlały rzeczywiste korelacje między zmiennymi. Działania prowadzone w tej metodzie powinny również spełniać warunek, że wyznacznik macierzy korelacji resztowych ma być największy oraz że konfiguracja czynników powinna być taka, że „korelacja kanoniczna między czynnikami wspólnymi i obserwowalnymi zmiennymi była maksymalna” (Sztemberg-Lewandowska w [Metody statystyczne analizy wielowymiarowej ... 2004, s. 204]). W metodzie głównych składowych rozwiązanie otrzymuje się na podstawie równania charakterystycznego (5.31), ale zamiast macierzy zredukowanej wprowadza się macierz:

$$\mathbf{R}_2 = \mathbf{U}^{-1}(\mathbf{R} - \mathbf{U}^2)\mathbf{U}^{-1} = \mathbf{U}^{-1}\tilde{\mathbf{R}}\mathbf{U}^{-1}. \quad (5.35)$$

Należy również określić liczbę czynników głównych, które są istotne dla przeprowadzonej analizy. Można wyróżnić następujące kryteria:

- kryterium „łokcia”, zaproponowane przez Cattela: z wartości własnych ułożonych w porządku nierosnącym tworzony jest wykres; numer wartości własnej, po przejściu której wykres ma przebieg łagodny, wskazując najlepszą liczbę czynników;
- kryterium Keisera: pod uwagę brane są tylko $\lambda_i > 1$;
- $\frac{\lambda_i}{\sum_{p=1}^p \lambda_p} > w_0$, gdzie $w_0 < 1$.

W analizie czynnikowej można napotkać problemy związane z interpretacją czynników. W takim przypadku stosuje się rotację układu odniesienia przy niezmienionej liczbie czynników i zasobach zmienności wspólnej. Wyróżnia się rotacje ortogonalne i ukośne. W pierwszym podejściu zachowuje się warunek ortogonalności czynników, a w drugim dodatkowo wprowadza się korelacje między czynnikami, które utrudniają, a nie jak się zakłada w ogólnym postępowaniu związanym z rotacją czynników – ułatwiają interpretację wyników.

Wśród rotacji ortogonalnych czynników często wyróżnia się technikę *quartimax*. Aby zastosować algorytm tej rotacji, maksymalizuje się wyrażenie:

$$Q = \sum_{j=1}^m \sum_{l=1}^p b_{jl}^4. \quad (5.36)$$

Zastosowanie tego kryterium jest korzystne z punktu widzenia interpretacji czynników, gdy „mała liczba zmiennych ma wysokie ładunki na tym czynniku, a pozostałe zmienne mają ładunki zerowe” (Sztemberg-Lewandowska w [Metody statystyczne analizy wielowymiarowej ... 2004, s. 213]).

Inne techniki rotacji ortogonalnej i ukośnej opisuje Sztemberg-Lewandowska w [Metody statystyczne analizy wielowymiarowej ... 2004, s. 211-217].

Przykład 5.2

W celu zaprezentowania postępowania w analizie czynnikowej badaniu poddano (hipotetyczne) dane opisujące ocenę cech jogurtu owocowego dokonaną na ośmiopunktowej skali Stapela (0-7). Cechami opisującymi produkt są:

- cena: 0 – bardzo tani, 7 – bardzo drogi (X_1);
- zawartość tłuszczu: 0 – dietetyczny, 7 – bardzo wysoka zawartość tłuszczu (X_2);
- konsystencja: 0 – bardzo gęsty, 7 – bardzo wodnisty (X_3);
- zawartość owoców: 0 – bardzo dużo owoców, 7 – bardzo mało owoców (X_4);
- dostępność na rynku: 0 – dostępny bez ograniczeń, 7 – nieosiągalny (X_5).

Oceny dokonane przez 20 respondentów zawiera tab. 5.4.

Tabela 5.4. Oceny jogurtu owocowego dokonane przez respondentów ze względu na analizowane cechy

		Numer zmiennej				
		X_1	X_2	X_3	X_4	X_5
Numer respondenta	1	7	5	7	2	2
	2	3	5	1	6	7
	3	0	0	4	0	0
	4	1	1	0	7	7
	5	5	7	7	3	3
	6	3	7	5	4	7
	7	3	5	6	3	6
	8	7	6	7	2	2
	9	3	3	0	7	7
	10	1	7	4	5	7
	11	0	1	6	0	0
	12	5	7	7	3	5
	13	5	7	1	7	7
	14	1	1	7	1	1
	15	7	7	7	2	1
	16	6	4	5	1	2
	17	3	2	1	7	6
	18	6	7	7	4	4
	19	4	4	2	6	7
	20	5	6	7	3	6

Źródło: opracowanie własne.

Celem badania jest zredukowanie liczby analizowanych zmiennych. W postępowaniu wykorzystano metodę głównych składowych.

Sprawdzono zasadność stosowania analizy czynnikowej dla zebranych danych. W tym celu wykorzystano test sferyczności Bartletta. Po wyznaczeniu $\log(\det \mathbf{R}) = -1,962$ obliczono statystykę $\chi^2 = 32,373$. Dla $\frac{m(m-1)}{2} = 10$ oraz poziomu istotności $\alpha = 0,05$ z tablic odczytano $\chi^2_{\alpha} = 18,307$. Tak więc zbiór współczynników korelacji jest istotny i istnieją przesłanki do zastosowania analizy czynnikowej.

W wyniku przeprowadzenia analizy otrzymano pięć wartości własnych, które zapisano w tab. 5.5.

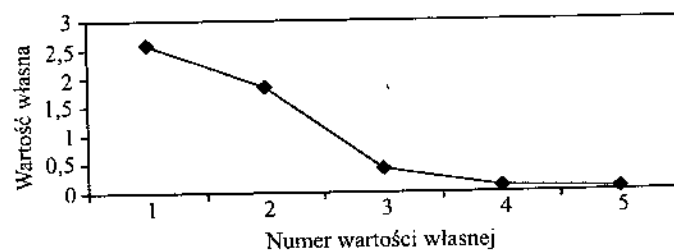
Tabela 5.5. Wartości własne

	Wartość własna	Zasób zmienności wspólnej (%)	Skumulowane wartości własne	Procent wyjaśniania zmienności wspólnej
1	2,58	51,65	2,58	51,65
2	1,85	36,90	4,43	88,55
3	0,42	8,33	4,85	96,88
4	0,10	2,04	4,95	98,92
5	0,05	1,08	5	100

Źródło: opracowanie własne.

Na podstawie wartości zapisanych w tab. 5.5 można stwierdzić, że dwa pierwsze czynniki główne zawierają łącznie 89% zmienności wspólnej.

W celu wskazania, ile czynników jest istotnych, zastosowano kryteria „łokcia” (rys. 5.2) oraz Keisera.



Rys. 5.2. Wykres wartości własnych – kryterium „łokcia”

Źródło: opracowanie własne.

Na podstawie rys. 5.2 można wybrać trzy czynniki główne odpowiadające trzem największym wartościom własnym. Stosując drugie kryterium, można zmniejszyć liczbę czynników do dwóch (dwie wartości własne są większe od 1).

W tabeli 5.6 zaprezentowano ładunki czynnikowe otrzymane metodą głównych składowych.

Tabela 5.6. Ładunki czynnikowe

	Czynnik 1	Czynnik 2
X_1	-0,31573	0,82855
X_2	-0,08271	0,943625
X_3	-0,90674	0,254345
X_4	0,941389	0,267993
X_5	0,876136	0,362941

Źródło: opracowanie własne.

Na podstawie wartości zawartych w tab. 5.6 należy stwierdzić, że pierwszy czynnik główny jest wynikiem redukcji zmiennych X_2 , X_3 , X_4 (najwyższe co do modułu wartości), natomiast drugi czynnik główny tworzą zmienne X_1 oraz X_2 . Pierwszy czynnik główny jest związany z walorami smakowymi, natomiast drugi z cechami rynkowymi jogurtu.

Po zastosowaniu innych opisanych technik wyodrębniania czynników wspólnych i rotacji *quartimax* otrzymano wyniki zgodne z zaprezentowanymi powyżej (zob. tab. 5.7).

Tabela 5.7. Wyniki zastosowania innych technik analizy czynnikowej

		Metoda osi głównych		Metoda największej wiarygodności		Metoda osi głównych z rotacją <i>quartimax</i>	
Numer czynnika		1	2	1	2	1	2
Wartość własna		2,52	1,66	2,48	1,73	2,52	1,66
Zasób zmienności wspólnej (%)		50,35	33,29	49,69	34,57	50,35	33,29
Skumulowane wartości własne		2,52	4,18	2,48	4,21	2,52	4,18
Procent wyjaśniania zmienności wspólnej		50,35	83,64	49,69	84,26	50,35	83,64
Ładunki czynnikowe	X_1	-0,25	0,74	-0,11	0,78	0,12	0,78
	X_2	-0,04	0,93	0,13	0,92	-0,10	0,92
	X_3	-0,87	0,31	-0,81	0,48	0,82	0,44
	X_4	0,95	0,22	0,98	0,03	-0,98	0,07
	X_5	0,88	0,33	0,92	0,18	-0,92	0,17

Źródło: opracowanie własne.

Zastosowania analizy czynnikowej w badaniach marketingowych

Analiza czynnikowa pojawia się najczęściej w badaniach marketingowych prowadzonych w trakcie segmentacji rynku. Innymi obszarami zastosowań są badanie produktu i określanie preferencji konsumentów, analiza cen, badania efektywności działań reklamowych.