Franciszek A. S z c z o t k a /Warszawa/

PODSTAWY ANALIZY CZYNNIKOWEJ

1. Wstęp

Metoda ilościowego opisu zjawisk przyrodniczych okazała się metodą niezmiernie przydatną. Ma to swoje uzasadnienie przede wszystkim w tym, że opis ilościowy jest sprawdzalny i przekazywalny. Podstawą takiego opisu są pomiary określonych wielkości lub pomiary nasilenia występowania niektórych właściwości. Mierzone właściwości nazywamy cechami wielkościowymi. Zwykle zjawiska przyrodnicze są tak bardzo złożone,że ich opis wymaga zmierzenia wielu różnych cech. Tak np. wysokość ciała jest jedną z charakterystyk budowy, ale tylko w niewielu przypadkach można tę charakterystykę uznać za wystarczającą. Do subtelniejszej oceny budowy trzeba użyć dodatkowo pomiarów innych cech. Chociaż łatwo zgodzić się z poglądem, że wielocechowa charakterystyka badanego zjawiska jest wszechstronniejsza od jednocechowej, to stosunkowo rzadko zdarzało się, by zagadnienie było stawiane i rozwiązywane w formie wielocechowej. Ma to swoją przyczynę w pewnym ubóstwie metod matematycznych do opracowania zagadnień wielocechowych jak i w dużych na ogół trudnościach obliczeniowych, z jakimi wiąże się stosowanie wypracowanych dotychczas metod. Dopiero w ostatnich latach możliwość szybkiego wykonania skomplikowanych obliczeń - dzięki elektronicznej technice obliczeniowej - spowodowała intensywną pracę nad metodami wielocechowymi.

Jednym z ważnych i często stawianych w praktyce jest zagadnienie analizy zespołu cech. Mierzone cechy na ogół nie oddają adekwatnie badanego zjawiska biologicznego. Każda z nich tylko w pewnym stopniu mierzy zjawisko interesujące badacza, a ponadto wielkości leżące poza zainteresowaniami badawczymi oraz chwilowe przypadkowe zakłócenia. Oczywiście dla badającego interesujące jest przede wszystkim to, co jest mierzone wspólnie przez cechy rozważanego zespołu. Nasuwają się wtedy pytania:

1. Jaka część informacji zawarta w zespole cech dotyczy części mierzonej przez cechy wspólnie,
2. Jaki jest udział poszczególnych cech w określaniu części wspólnej,
3. Jaka jest struktura zjawiska mierzonego wspólnie przez zespół cech, czy jest to zjawisko jednorodne, czy niejednorodne w określonym sensie.

Zespołem bardzo ogólnych i wszechstronnych metod pozwalających odpowiedzieć na tego typu pytania jest analiza czynnikowa.

Ponieważ podstawą analizy są zależności między cechami /mierzone zwykle współczynnikami korelacji lub kowariancji/, analiza czynnikowa jako metoda statystyczna jest przede wszystkim analizą i swoistą interpretacją związków między cechami. Dla przeprowadzenia tej analizy wprowadza się pewne umowne zmienne losowe zwane czynnikami /lub faktorami/. Można więc najogólniej powiedzieć, że analiza czynnikowa jest zespołem metod do badania struktury zależności między cechami przy pomocy pojęcia czynników.

Często udaje się czynnikom nadawać sensowną interpretację przyrodniczą. Pamiętać jednak należy o tym, że są to umowne formalne wielkości pozwalające na opis matematyczny badanego materiału.

Jak każda pojęciowo nowa metoda badawcza, tak i analiza czynnikowa wymaga pewnego obycia i wyrobienia sobie nawyków myślowych ułatwiających posługiwanie się nią. Opracowanie niniejsze ma w tym pomóc przez omówienie podstawowych pojęć tej metody oraz pokazanie kilku przykładów zastosowań.

Obliczenia związane z przeprowadzeniem analizy czynnikowej są długie i tylko w nielicznych i bardzo prostych zagadnieniach możliwe do wykonania bez elektronowych maszyn liczących. Powszechność jednak takich maszyn w czasach obecnych umożliwia już szerokie stosowanie analizy czynnikowej, a to z kolei powoduje doskonalenie jej metod zarówno pod względem teoretycznym jak i samej techniki obliczeniowej.

Zdając sobie sprawę z tego, że obliczenia z zasady nie będą wykonywane przez samego badacza - niematematyka, nie podaję w niniejszym opracowaniu żadnych algorytmów obliczeń.

Ograniczyłem także cytowanie literatury, poprzestając na pracach o znaczeniu podstawowym dla rozwoju analizy czynnikowej lub opracowaniach współczesnych o charakterze monograficznym.

2. Definicje i symbole

Cechy będą oznaczone przez $X_1, X_2, \ldots, X_p$. Zakłada się, że są to cechy wielkościowe. Dysponujemy pomiarami tych cech uzyskanymi z badania próbki losowej. Ilość badanych osobników oznaczymy przez n. Wartość cechy X_i u osobnika o numerze j oznaczona będzie przez x_{ij}, tak że pierwszy wskaźnik /i/ jest numerem cechy, a drugi /j/ numerem badanego osobnika.

W opracowaniu użyte będą następujące symbole:

średnia arytmetyczna cechy X_i: $\bar{x}_i = \frac{1}{n} \sum_{j=1}^{n} x_{ij}$, średnia wartość cechy X_i w populacji /wartość oczekiwana cechy/: μ_i,

standardowe odchylenie cechy X_i: $s_i = \sqrt{\frac{\sum_{j=1}^{m} (x_{ij} - \bar{x}_i)^2}{n-1}}$

dyspersja /w populacji/: σ_i,

Kowariancja cech X_k i X_l w próbce: $c_{kl} = \frac{\sum_{j=1}^{m} (x_{kj} - \bar{x}_k)(x_{lj} - \bar{x}_l)}{n-1}$

Kowariancja cech X_k i X_l w populacji: C_{kl}

Z definicji wynika, że:

$$C_{kk} = \sigma_k^2$$

$$c_{kk} = s_k^2$$

Współczynnik korelacji w próbce: $\hat{r}_{kl} = \frac{c_{kl}}{\sqrt{c_{kk}\, c_{ll}}} = \frac{c_{kl}}{s_k\, s_l}$

Współczynnik korelacji w populacji $r_{kl} = \frac{C_{kl}}{\sqrt{C_{kk} C_{ll}}} = \frac{C_{kl}}{\sigma_k \cdot \sigma_l}$

Dla wygody używane będą także symbole E, War, Kow i Kor na oznaczenie odpowiednio wartości oczekiwanej, wariancji, kowariancji i korelacji /w populacji/.

W analizie czynnikowej ogranicza się rozważania do tych tylko parametrów statystycznych. Są one jedynymi charakterystykami statystycznymi w przypadku, gdy rozkłady rozważanych zmiennych losowych są normalne.

W opracowaniu stosowany będzie zapis macierzowy. Macierze z zasady oznaczone będą dużymi literami. Macierz M o w wierszach i k kolumnach zapisana będzie w postaci $M_{w \times k}$. Przez M' rozumie się transpozycję macierzy M, przez M^{-1} jej odwrotność, jeżeli taka istnieje. Symbolem diag M oznaczana będzie macierz przekątniowa, która ma na głównej przekątnej te same wyrazy co macierz M.

3. Podstawowe modele analizy czynnikowej

3.1. Spearmana model analizy czynnikowej z jednym czynnikiem wspólnym.

Twórcą analizy czynnikowej był C. Spearman. Sformułował on hipotezę, że we wszystkich czynnościach umysłowych ujawnia się jedna wspólna cecha, zwana przez Spearmana inteligencją ogólną, a poza nią właściwości specyficzne, różne dla każdego rodzaju czynności. Innymi słowy: rozwiązanie każdego zadania umysłowego wymaga pewnych ogólnych umiejętności umysłowych oraz umiejętności specyficznych dla tego zadania. W pracy "General Intelligence Objectivily Determined and Measured" C.Spearman /1904/ użył po raz pierwszy dla rozwiązania sformułowanych w tytule zadań metody, która później rozwinęła się i uzyskała nazwę analizy czynnikowej.

Oznaczmy za Spearmanem inteligencję ogólną przez G. Będziemy ją w dalszym ciągu nazywali czynnikiem ogólnym. Zdolności specyficznych będzie tyle, ile zadań testowych. Zdolności te będziemy oznaczali przez U_1, U_2, ..., U_p i nazywali czynnikami specyficznymi. Zarówno czynnik ogólny jak i czynniki specyficzne będziemy uważali za zmienne wielkościowe wyrażone w takiej skali, by ich średnie /w populacji/ były równe zeru, a dyspersje /w populacji/ równe jeden. Wyniki w p rozpatrywanych testach będziemy oznaczali przez X_1, X_2, ..., X_p. Hipotezę Spearmana można wypowiedzieć następująco: wynik w każdym teście uzyskany przez dowolną osobę jest sumą składnika proporcjonalnego do wielkości czynnika ogólnego u tej osoby i składnika proporcjonalnego do wielkości czynnika specyficznego dla tego testu. To sformułowanie słowne zapisuje się przy pomocy wzoru następująco:

/3.1.1/
$$
\begin{aligned}
X_1 &= l_1 G + v_1 U_1 \\
X_2 &= l_2 G + v_2 U_2 \\
&\dots\dots\dots\dots\dots \\
X_p &= l_p G + v_p U_p
\end{aligned}
$$

Jeżeli przez x_{ij} oznaczyć wartość i-tej cechy u j-ego osobnika, a przez u_{ij} i g_j wartość czynnika ogólnego i i-tego czynnika specyficznego, to dla tego osobnika powyższy układ równań przyjmuje postać

$$x_{1j} = l_1 g_j + v_1 u_{1j}$$
$$x_{2j} = l_2 g_j + v_2 u_{2j}$$
$$\dots\dots\dots\dots\dots\dots$$
$$x_{pj} = l_p g_j + v_p u_{pj}$$

Symbole g_j oraz u_{1j} ...., u_{pj} oznaczają odpowiednio wartości czynnika ogólnego G i wartości czynników specyficznych u j-tego osobnika. Liczby l_1, l_2 ..., l_p i v_1, v_2 ..., v_p są współczynnikami proporcjonalności. Analizując wzór stwierdzamy, że współczynniki proporcjonalności są takie same dla wszystkich osób.

Równania /3.1.1/ uzupełniamy założeniami o unormowaniu czynnika ogólnego i czynników specyficznych.

/3.1.2/ $EG = 0$, $\quad War\ G = 1$

/3.1.3/ $EU_i = 0$, $\quad War\ U_i = 1 \quad$ dla $i = 1, \dots, p$.

W tych i dalszych wzorach czynnik ogólny G i czynniki specyficzne U_i traktowane są jako zmienne losowe w badanej populacji osobników.

Założenie, że czynniki specyficzne są różne od czynnika ogólnego i różne między sobą, wyraża się warunkiem, by były to zmienne losowe wzajemnie nieskorelowane, czyli by

/3.1.4/ $Kor\ /G, U_i/ = 0 \quad i = 1, \dots p$

/3.1.5/ $Kor\ /U_k, U_l/ = 0 \quad$ dla $k, l = 1, \dots, p$ i $k \neq l$

Dodaje się zwykle założenie o unormowaniu wyników w testach

/3.1.6/ $EX_i = \mu_i = 0 \qquad War\ X_i = \sigma_i^2 = 1$

Podobną hipotezę i związany z nią model matematyczny można formułować w stosunku do dowolnego zespołu cech $X_1, \ldots, X_p$

Wzory /3.1.1/ - /3.1.6/ wiążą wielkości obserwowalne, mianowicie wartości cech $X_1, X_2, \ldots, X_p$, które mogą być znane dla każdej z badanych osób, oraz niewiadome jakimi są współczynniki $l_1, l_2, \ldots, l_p$, współczynniki $v_1, v_2, \ldots, v_p$, które są jednakowe dla wszystkich badanych osób, indywidualne wartości czynnika ogólnego g_j i indywidualne wartości czynników specyficznych $u_{1j}, u_{2j}, \ldots, u_{pj}$. Nie wiemy, czy model ten jest zgodny z faktami doświadczalnymi. Istnieją metody statystyczne, które pozwalają to sprawdzić. Metody te poznamy później przy omawianiu analizy czynnikowej w postaci ogólnej. Przyjmując, że model jest słuszny, powstaje problem liczbowego wyznaczenia nieznanych wielkości, a mianowicie współczynników $l_1, l_2, \ldots, l_p$ i $v_1, v_2, \ldots, v_p$ stanowiących parametry charakteryzujące całą populację oraz wartości czynnika g_j, charakteryzującego poszczególne osoby. Rozwiązanie tego zagadnienia także będzie omawiane w osobnych rozdziałach.

By zinterpretować nieznane współczynniki $l_1, \ldots, l_p$ oraz $v_1, \ldots, v_p$ załóżmy, że hipoteza sformalizowana wzorami /3.1.1/ - /3.1.6/ jest słuszna. Proste przekształcenia pokazują, że współczynniki $l_1, l_2, \ldots, l_p$ są równe współczynnikom korelacji czynnika ogólnego z poszczególnymi cechami, czyli

/3.1.7/ $l_i = Kor\ /X_i, G/.$

Współczynnik l_i nazywa się ładunkiem czynnika ogólnego G w cesze X_i /w języku angielskim: loading/ lub nasyceniem cechy X_i czynnikiem G / w języku francuskim: saturation/.

Przy pomocy formalnych przekształceń algebraicznych można wariancję cech przedstawić w postaci

/3.1.8/ $\sigma_i^2 = l_i^2 + v_i^2.$

Jeżeli założymy, że wariancja wyników testowych jest równa 1, to otrzymujemy w konsekwencji

/3.1.8a/ $l_i^2 + v_i^2 = 1$

Kwadrat ładunku l_i^2 to ta część wariancji zmiennej X_i, która jest wywołana zróżnicowaniem czynnika ogólnego G, natomiast v_i^2 część wariancji wywołana zróżnicowaniem czynnika specyficznego. Z tego powodu v_i^2 nazywa się zasobem zmienności specyficznej lub wariancją specyficzną.

Czynnik ogólny G ma z założenia średnią zero i dyspersję jeden. Jedyne zatem parametry, jakie go charakteryzują, to jego współczynniki korelacji z cechami, czyli ładunki /zakładając, że jego rozkład jest normalny/. Innymi słowy: czynnik ogólny określony jest przez ładunki. Znając ładunki możemy ze wzoru /3.1.8a/ obliczyć wariancje specyficzne.

Wykazaliśmy poprzednio, że $l_i^2 + v_i^2 = 1$. Można dalej łatwo wykazać, że:

$$r_{kl} = \mathrm{Kor}\ /X_k, X_l/ = E\, X_k X_l =$$

$$= E/l_k G + v_k U_k/\ /l_l G + v_l U_l/ = l_k l_l \quad \text{dla } k \neq l$$

/3.1.9/ $r_{kl} = l_k l_l$

Streszczając: jeżeli założenia modelu Spearmana są słuszne, to między wyrazami macierzy korelacji a ładunkami zachodzą następujące związki:

/3.1.10/ $$l_i^2 + v_i^2 = 1$$

$$l_k \cdot l_l = r_{kl} \qquad \text{dla } k \neq l$$

Równości te można napisać w postaci macierzowej następująco:

$$\begin{bmatrix} 1 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2p} \\ r_{p1} & r_{p2} & \dots & 1 \end{bmatrix} = \begin{bmatrix} l_{11}^2 & l_1 l_2 & \dots & l_1 l_p \\ l_2 l_1 & l_2^2 & \dots & l_2 l_p \\ l_p l_1 & l_p l_2 & \dots & l_p^2 \end{bmatrix} + \begin{bmatrix} v_1^2 & 0 & \dots & 0 \\ 0 & v_2^2 & \dots & 0 \\ 0 & 0 & \dots & v_p^2 \end{bmatrix}$$

lub też

$$\begin{bmatrix} 1 - v_1^2 & r_{12} & \dots & r_{1p} \\ r_{21} & 1 - v_2^2 & \dots & r_{2p} \\ r_{p1} & r_{p2} & \dots & 1-v_p^2 \end{bmatrix} = \begin{bmatrix} l_1^2 & l_1 l_2 & \dots & l_1 l_p \\ l_2 l_1 & l_2^2 & \dots & l_2 l_p \\ l_p l_1 & l_p l_2 & \dots & l_p^2 \end{bmatrix}$$

Oznaczmy przez $\underset{/px1/}{L} = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_p \end{bmatrix}$

/jednokolumnową/ macierz ładunków

a przez $\underset{/pxp/}{V}$ przekątniową macierz wariancji specyficznych. Wtedy powyższą równość można napisać krótko:

/3.1.11/ $$R = LL' + V$$

lub też

/3.1.12/ $$R - V = LL'$$

Macierz korelacji, w której od głównej przekątnej odjęto wariancje specyficzne, nazywa się w analizie czynnikowej zredukowaną macierzą korelacji.

Macierz korelacji, którą można tak zredukować, by dała się przedstawić jako iloczyn LL', Spearman nazywa macierzą o strukturze hierarchicznej. Mamy wobec tego wniosek, że jeżeli model Spearmana jest słuszny, to macierz korelacji jest macierzą o strukturze hierarchicznej. Fakt ten wykorzystuje się do sprawdzenia zgodności modelu z doświadczeniem. Jeżeli na podstawie obliczonej z badań macierzy korelacji odrzucamy hipotezę, że macierz w populacji jest hierarchiczna, to fakt ten obala słuszność modelu Spearmana, czyli słuszność występowania jednego tylko czynnika.

Przykładem macierzy korelacji o strukturze hierarchicznej jest macierz korelacji podana w tablicy 1.

Tablica 1 Macierz o strukturze hierarchicznej

1.00	0,26	0,18	0,22	0.28
0.26	1.00	0.20	0.24	0.31
0.18	0.20	1.00	0.16	0.21
0.22	0.24	0,16	1.00	0.25
0.28	0.31	0.21	0.25	1.00

Zasoby zmienności specyficznej wynoszą:

$v_1^2 = 0.24$, $v_2^2 = 0.29$, $v_3^2 = 0.14$, $v_4^2 = 0.19$, $v_5^2 = 0.32$

a ładunki:

$l_1 = 0.49$, $l_2 = 0.54$, $l_3 = 0.37$, $l_4 = 0.44$, $l_5 = 0.57$.

Istotną cechą macierzy o strukturze hierarchicznej jest następująca właściwość: Jeżeli weźmiemy cztery wyrazy zredukowanej macierzy, które stanowią wierzchołki prostokąta o bokach równoległych do brzegów tablicy, to różnica iloczynów przeciwległych wierzchołków równa się zeru. Łatwo to sprawdzić rachunkowo w podanym przykładzie, ale i nietrudno udowodnić. Jeżeli R jest macierzą o strukturze hierarchicznej, czyli jest sumą macierzy LL′ i macierzy przekątniowej V, to korzystając z tej właściwości łatwo obliczyć ładunki ze współczynników korelacji. Mianowicie biorąc cztery takie wyrazy, z czego jeden jest wyrazem głównej przekątnej, mamy:

/3.1.13/ $$l_k^2 = \frac{r_{kl}\, r_{km}}{r_{lm}}$$

Obliczywszy w ten sposób $l_1^2, l_2^2, \dots l_p^2$ możemy w końcu obliczyć:

/3.1.14/ $$v_i^2 = 1 - l_i^2$$

Rozważania dotychczasowe możemy streścić w następującym wniosku. Jeżeli słuszny jest model sformułowany przez Spearmana, czyli jeżeli słuszne są założenia sformułowane we wzorach /3.1.1/-/3.1.8a/, to zachodzi równość /3.1.11/ lub /3.1.12/, czyli macierz korelacji w populacji jest macierzą hierarchiczną. I na odwrót, jeżeli macierz korelacji w populacji jest macierzą hierarchiczną, to można według wzorów /3.1.13/ i /3.1.14/ obliczyć ładunki i wariancje specyficzne, a więc słuszny jest wtedy model.

Zespół cech, którego macierz korelacji ma strukturę hierarchiczną, można uważać za zespół o najprostszej strukturze zależności /po zespole cech nieskorelowanych/, bo występujące korelacje między wszystkimi parami cech wytłumaczyć można występowaniem w każdej cesze jednego wspólnego czynnika. Wielkość ładunków l_i charakteryzuje korelację cechy a czynnikiem wspólnym, l_i^2 zaś udział zmienności czynnika wspólnego w zmienności tej cechy.

Wszystko to odnosi się oczywiście do macierzy współczynników korelacji w populacji. Wiadomo, że wskutek błędu próbki równość /3.1.11/ nie będzie nigdy zachodziła dokładnie. Wartości $l_1, l_2, \ldots, l_p$ wyznacza się wtedy tak, by różnice między macierzą obliczonych współczynników korelacji a macierzą odtworzoną przy pomocy ładunków wg wzoru /3.1.10/ była możliwie mała. Zagadnieniem tym, które jest zagadnieniem estymacji, zajmiemy się szczegółowo w osobnym paragrafie.

3.2. Ogólny model analizy czynnikowej

Model Spearmana został uogólniony przez wielu innych autorów. W najogólniejszej postaci model czynnikowy podał L. L. Thurstone /1947/. W modelu przez niego sformułowanym dopuszcza się występowanie kilku czynników wspólnych. Model jest następujący:

1. O cechach zakłada się, że są unormowane na średnią zero, tzn., że

/3.2.1/ $$\mu_1 = \mu_2 = \mu_3 = \ldots = 0$$

2. Zakłada się dalej, że każda z cech da się przedstawić jako średnia ważona r zmiennych losowych wspólnych F_1, $F_2, \ldots, F_r$ i zmiennej losowej U specyficznej dla tej cechy:

$$X_1 = l_{11}F_1 + l_{12}F_2 + \ldots + l_{1r}F_r + v_1 U_1$$

$$X_2 = l_{21}F_1 + l_{22}F_2 + \ldots + l_{2r}F_r + v_2U_2$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

/3.2.2/

$$X_p = l_{p1}F_1 + l_{p2}F_2 + \ldots + l_{pr}F_r + v_pU_p$$

Zmienne $F_1, F_2, \dots, F_r$ nazywają się czynnikami wspólnymi, zmienne $U_1, U_2, \dots, U_p$ czynnikami specyficznymi.

Zmienne $F_1, \dots, F_r$ i $U_1, \dots, U_p$ są traktowane jako zmienne losowe. Przyjmują one więc inne wartości dla każdego badanego osobnika. Współczynniki l_{ik} oraz v_i są parametrami stałymi, jednakowymi dla wszystkich osobników.

3. Ponieważ jest obojętnie, w jakiej skali wyrażają się zmienne $F_1, F_2, \dots, F_r$ oraz zmienne $U_1, U_2, \dots, U_p$ zakłada się, że wartości oczekiwane /średnie w populacji/ tych zmiennych są równe zeru, a ich wariancje równe jeden, czyli

/3.2.3/ $EF_k = 0, \quad \text{War } F_k = 1, \quad k = 1, \dots, r$

/3.2.4/ $EU_i = 0, \quad \text{War } U_i = 1, \quad i = 1, \dots, p$

4. Zakłada się, że czynniki wspólne są nieskorelowane, czyli

/3.2.5/ $\text{Kor } /F_k, F_l/ = 0 \quad \text{dla } k \neq l$

5. Podobnie jak w modelu o jednym czynniku wspólnym zakłada się, że czynniki specyficzne są parami nieskorelowane oraz nieskorelowane z czynnikami wspólnymi:

/3.2.6/ $\text{Kor } /F_k, U_l/ = 0, \quad \text{dla } k = 1, \dots r \quad \text{oraz } l = 1, \dots, p$

/3.2.7/ $\text{Kor } /U_k, U_l/ = 0, \quad \text{jeżeli tylko } k \neq l.$

Warunki /3.2.1/ - /3.2.7/ określają model, który będziemy nazywali ogólnym modelem analizy czynnikowej z nieskorelowanymi czynnikami.

Często do warunku /3.2.1/ dodaje się jeszcze warunek /3.2.1a/, by cechy były unormowane na dyspersję 1, czyli

/3.2.1a/ $\text{War } X_i = {}_i^2 = 1 \quad \text{dla } i = 1, \dots, p$

Założenie to ma dla analizy czynnikowej znaczenie istotne. Okazuje się mianowicie /por. p. 4.4/, że wartości współczynników l_{ik} oraz v_i zależą od jednostek, w jakich wyrażone są cechy. Dlatego też w dalszym ciągu rozgraniczyć trzeba przypadek ogólny i przypadek cech unormowanych na wariancję jeden.

Współczynniki przy czynnikach wspólnych i specyficznych mają podobną interpretację, co w modelu z jednym czynnikiem wspólnym. Mianowicie współczynniki $l_{11}, l_{21}, \ldots, l_{p1}$ są równe kowariancjom kolejnych cech z czynnikiem pierwszym. Dzieląc je przez dyspersje cech otrzymamy współczynniki korelacji kolejnych cech z czynnikiem pierwszym.

Jeżeli rozważamy cechy unormowane, to ich dyspersje są równe jedności i wtedy współczynniki $l_{11}, l_{21}, \ldots, l_{p1}$ są wprost równe współczynnikom korelacji. Współczynniki $l_{12}, l_{22}, \ldots, l_{p2}$ są równe kowariancjom lub, po podobnym unormowaniu, współczynnikom korelacji cech i drugiego czynnika wspólnego. Podobną interpretację mają pozostałe współczynniki. Zachodzi ogólnie następujący związek:

$$/3.2.8/ \quad l_{ik} = \mathrm{Kow}\ /X_i, F_k/ \quad \text{oraz} \quad \mathrm{Kor}\ /X_i, F_k/ = l_{ik}/\sigma_i.$$

Także w tym przypadku współczynniki l_{ik} noszą nazwę ładunków. Zwykle ładunki zestawia się w tablicy postaci:

$$L = \begin{bmatrix} l_{11}, & l_{12}, & \ldots & l_{1r} \\ l_{21}, & l_{22} & \ldots & l_{2r} \\ \vdots & & & \\ \vdots & & & \\ l_{p1} & l_{p2} & \ldots & l_{pr} \end{bmatrix}$$

Tablicę taką nazywa się macierzą ładunków.

Można wyrachować przy pomocy przekształceń algebraicznych, że zachodzi następująca równość:

/3.2.9/ $$\sigma_i^2 = l_{i1}^2 + l_{i2}^2 + \ldots l_{ir}^2 + v_i^2$$

Wariancja cechy X_i jest więc równa sumie kwadratów ładunków w tej cesze czynników wspólnych oraz ładunku czynnika specyficznego. Wprowadźmy oznaczenie

/3.2.10/ $$h_i^2 = l_{i1}^2 + l_{i2}^2 + \ldots + l_{ir}^2$$

wtedy

/3.2.11/ $$\sigma_i^2 = h_i^2 + v_i^2 \quad \text{oraz}$$

/3.2.11a/ $$\frac{h_i^2}{\sigma_i^2} + \frac{v_i^2}{\sigma_i^2} = 1$$

Wielkość h_i^2 / σ_i^2 wskazuje, jaka część wariancji cechy X_i wynika ze zróżnicowania czynników wspólnych, wielkość v_i^2 / σ_i^2 zaś wskazuje jaka część wariancji wynika ze zróżnicowania czynnika specyficznego. Wielkość v_i^2 będę nazywał wariancją specyficzną, $\frac{v_i^2}{\sigma_i^2}$ względną wariancją specyficzną, a wielkość h_i^2 zasobem zmienności wspólnej cechy X_i.

Wzór /3.2.9/ wiąże wariancję cechy X_i i jej ładunki. Można wyprowadzić podobne zależności dla kowariancji. Mają one postać:

/3.2.12/ $$c_{kl} = l_{k1}l_{l1} + l_{k2}l_{l2} + \ldots + l_{kr}l_{lr}.$$

Oba te wzory można słownie wyrazić następująco: Wariancja cechy jest równa sumie kwadratów ładunków czynników wspólnych w tej cesze i jej wariancji specyficznej. Kowariancja dwóch cech natomiast wyraża się jako suma iloczynów ładunków w obu tych cechach. Równania /3.2.9/ i /3.2.12/ wiążące wariancje cech z ładunkami czynników wspólnych i wariancjami specyficznymi są równaniami podstawowymi dla analizy czynnikowej. Równania powyższe można napisać przy pomocy macierzy następująco

$$\begin{bmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \dots & \dots & \dots & \dots \\ c_{p1} & c_{p2} & \dots & c_{pp} \end{bmatrix} = \begin{bmatrix} l_{11} & \dots & l_{1r} \\ l_{21} & \dots & l_{2r} \\ \dots & \dots & \dots \\ l_{p1} & \dots & l_{pr} \end{bmatrix} \begin{bmatrix} l_{11} & \dots & l_{p1} \\ l_{12} & \dots & l_{p2} \\ \dots & \dots & \dots \\ l_{1r} & \dots & l_{pr} \end{bmatrix} + \begin{bmatrix} v_1^2 & 0 & \dots & 0 \\ 0 & v_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & v_p^2 \end{bmatrix}$$

lub krótko:

/3.2.13/ $$C = LL' + V$$

Jeżeli rozpatrujemy unormowany zespół cech, miejsce macierzy kowariancji C zajmuje macierz korelacji R i równanie podstawowe będzie wtedy miało postać.

/3.2.14/ $$R = LL' + V$$

Równania podstawowe /3.2.13/ można napisać w postaci nieco rozszerzonej. Oznaczmy przez L_k macierz kolumnową ładunków k-tego czynnika, czyli

$$L_k = \begin{bmatrix} l_{1k} \\ l_{2k} \\ \vdots \\ l_{pk} \end{bmatrix}, \quad k = 1, 2, \dots, r$$

Przy takim oznaczeniu równanie podstawowe /3.2.13/ da się napisać:

/3.2.15/ $$C = L_1L_1' + L_2L_2' + \ldots + L_rL_r' + V$$

Jest to rozbicie macierzy kowariancji na sumę r macierzy hierarchicznych i macierzy przekątniowej. Jest to równocześnie przedstawienie struktury związków między badanymi cechami.

W tablicy 2 podano macierz /fikcyjnych/ współczynników korelacji wyników w 7 konkurencjach lekkoatletycznych: 1. bieg na 100 m, 2. bieg przez płotki, 3. skok wzwyż, 4. skok o tyczce, 5. pchnięcie kulą, 6. rzut dyskiem, 7. rzut oszczepem. /W konkurencjach biegowych zmieniono znaki/.

Tablica 2. Macierz hipotetycznych współczynników korelacji wyników w 7 konkurencjach lekkoatletycznych

	1	2	3	4	5	6	7
1	1.000	0.553	0.352	0.344	0.221	0.248	0.217
2	0.553	1.000	0.470	0.472	0.250	0.290	0.264
3	0.352	0.470	1.000	0.306	0.354	0.372	0.293
4	0.344	0.472	0.306	1.000	0.206	0.227	0.196
5	0.221	0.250	0.354	0.206	1.000	0.700	0.496
6	0.248	0.290	0.372	0.227	0.700	1.000	0.487
7	0.217	0.264	0.293	0.196	0.496	0.487	1.000

Korelacje te można wytłumaczyć występowaniem dwóch czynników o następujących ładunkach:

	l_{11}	l_{12}	v_1^2
Bieg na 100 m	0.201	0.601	0.599
Bieg przez płotki	0.210	0.850	0.233
Skok wzwyż	0.378	0.459	0.647

	l_{i1}	l_{i2}	v_i^2
Skok o tyczce	0.192	0.508	0.705
Pchnięcie kulą	0.846	0.085	0.277
Rzut dyskiem	0.814	0.140	0.318
Rzut oszczepem	0.569	0.170	0.647

Pierwszy z czynników koreluje wysoko z wynikami w pchnięciu kulą /l_{51} = 0.846/ i w rzucie dyskiem /l_{61} = 0.814/, a trochę mniej z wynikami w rzucie oszczepem /l_{71} = 0.569/, natomiast bardzo nisko koreluje z wynikami w biegach /l_{11} = 0.201 i l_{21} = 0.210/. Czynnik pierwszy można więc interpretować jako siłę.

Drugi czynnik jest wysoko skorelowany z wynikami w biegach /l_{12} = 0.601 i l_{22} = 0.850/, mało natomiast z wynikami w rzutach /l_{52} = 0.085, l_{62} = 0.140 i l_{72} = 0.170/. Czynnik ten można interpretować jako szybkość. Wyniki skoku wzwyż korelują umiarkowanie z czynnikiem jednym i drugim: korelacja z siłą wynosi l_{31} = 0.378, z szybkością l_{32} = 0.459.

Na podstawie wariancji specyficznych można powiedzieć, że najbardziej specyficzną konkurencją, tzn. konkurencją najmniej skorelowaną z wyodrębnionymi czynnikami, jest skok o tyczce /v_4^2 = 0.705/. Następnie idą kolejno: rzut oszczepem /v_7^2 = 0.647/, skok wzwyż /v_2^2 = 0.657/ i bieg na 100 m /v_1^2 = 0.599/. Wariancje specyficzne są równocześnie miarą stopnia nieskorelowania z pozostałymi cechami.

Podobnie jak w przypadku jednego czynnika równania /3.2.1/ - /3.2.8/ wiążą wielkości dostępne obserwacji, mianowicie badane cechy, i wielkości nieznane, jakimi są wartości czynników wspólnych i specyficznych oraz ładunki tych czynników. Także w ogólnym modelu ładunki czynników wspólnych i wariancje specyficzne są parametrami liczbowymi charakteryzującymi całą badaną populację. Dlatego też najbardziej interesujące dla badacza są te właśnie wielkości.

Równocześnie ładunki są jedynymi charakterystykami statystycznymi czynników /jeżeli założyć, że rozkład czynników jest normalny/. Innymi słowy czynniki zdefiniowane są przez ładunki. Dwa czynniki będziemy uważali za identyczne, jeżeli mają w poszczególnych cechach te same ładunki, natomiast za różne, jeżeli chociaż w jednym ładunku się różnią. Dla badacza jest więc przede wszystkim interesująca ilość występujących czynników wspólnych oraz korelacje czynników z cechami, czyli ładunki.

Czynnik, który jest silnie skorelowany z cechami, jest czynnikiem ujawniającym się w dużym stopniu w badanych cechach. Czynnik taki będziemy uważali za czynnik ważny /w badanym zespole cech/. Czynnik słabo skorelowany z cechami jest czynnikiem ujawniającym się w stopniu niewielkim w badanym zespole cech. Jako miary ważności czynnika można użyć sumy kwadratów unormowanych ładunków tegoż czynnika ze wszystkimi cechami, czyli wielkość $\sum_{i=1}^{p} l_{ik}^2$. Wariancja specyficzna v_i^2 określa, w jakim stopniu cecha nasycona jest czynnikiem specyficznym, lub inaczej na ile cecha X_i odbiega od pozostałych.

Równaniami wiążącymi ładunki z parametrami rozkładu badanych cech są równania /3.2.13/ lub /3.2.14/. Gdybyśmy znali prawdziwe wariancje i kowariancje /lub prawdziwe współczynniki korelacji/ równania powyższe przedstawiałyby układ równań kwadratowych o niewiadomych l_{ik} oraz v_i. Powstaje wtedy pytanie o jednoznaczności rozwiązania. Okazuje się że równania /3.2.13/ ani /3.2.14/ nie określają jednoznacznie ładunków. Mianowicie, jeżeli L jest macierzą ładunków spełniającą równanie /4.2.13/ a Θ jest dowolną macierzą ortogonalną, to macierz $\tilde{L} = L\Theta$ także spełnia powyższe równanie. Wynika z tego, że znając jedno rozwiązanie, można z niego w sposób formalny uzyskać nieskończenie wiele rozwiązań, a więc nieskończenie wiele macierzy ładunków równoważnych z punktu widzenia matematycznego. Ponieważ różnią się one ładunkami, każde z nich określa inny /w interpretacji/ układ czynników. Ustalenie zasad wyboru jednego spośród wielu możliwych rozwiązań nazywa się zagadnieniem rotacji. Zagadnienie to omówione będzie w dalszej części.

Drugie zagadnienie, to pytanie jak postąpić, gdy nie znamy wariancji i kowariancji w populacji, a jedynie ich oszacowania z próbki, czyli wielkości obarczone błędem losowym. Zagadnienie to nazywa się zagadnieniem estymacji i omówione będzie w następnym rozdziale.

3.3. Inne modele

Rozpatruje się czasem modele odbiegające od sformułowanego. Modele takie tworzy się na podstawie informacji posiadanych z wcześniejszych badań o badanym zjawisku, lub też formułuje się je celem sprawdzenia określonych hipotez naukowych. Czasem w sposób zamierzony posługujemy się uproszczonym modelem wymagającym mniejszej pracy obliczeniowej w badaniach wstępnych, mających na celu uzyskanie ogólnego rozeznania w strukturze czynnikowej.

Zmiany w modelu mogą dotyczyć trzech rzeczy: założeń co do wariancji specyficznych, założeń co do struktury macierzy ładunków oraz założenia niezależności czynników wspólnych.

3.3.1. Model z równymi wariancjami specyficznymi

Model ten różni się od modelu ogólnego tym, że wprowadza się dodatkowo założenie o równości wariancji specyficznych. Możliwe tu są dwa warianty:

a/ zakłada się równość wariancji specyficznych wprost

/3.3.1/ $$v_1^2 = v_2^2 = \dots v_p^2 = v^2$$

lub też

b/ zakłada się równość względnych wariancji specyficznych, czyli

/3.3.2/
$$\frac{v_1^2}{\sigma_1^2} = \frac{v_2^2}{\sigma_2^2} = \ldots = \frac{v_p^2}{\sigma_p^2} = \tilde{v}^2$$

To drugie założenie sprowadza się do założenia o równości wariancji specyficznych w równaniu dla cech unormowanych. W pierwszym przypadku równania podstawowe mają postać

/3.3.3/
$$C = LL' + v^2 I \quad ,$$

w drugim

/3.3.4/
$$R = LL' + \tilde{v}^2 I \quad .$$

gdzie I jest macierzą jednostkową.

Trzeba zwrócić uwagę na to, że rozwiązania uzyskane z macierzy kowariancji są nieporównywalne z rozwiązaniami uzyskanymi dla macierzy korelacji. Innymi słowy, z rozwiązania równań /3.3.3/ nie można obliczyć bezpośrednio rozwiązania równań /3.3.4/ i na odwrót.

Równania podstawowe są przy założeniu równości wariancji specyficznych łatwiejsze do rozwiązania. Jednak konsekwencją tego założenia jest to, że dla przedstawienia macierzy kowariancji w postaci czynnikowej wymagana jest zwykle większa ilość czynników wspólnych.

3.3.2. Modele z określoną strukturą macierzy ładunków

W modelach takich z góry żąda się, by niektóre ładunki były zerami, tzn. by określone czynniki wspólne nie występowały w niektórych cechach. Takie postępowanie może być uzasadnione zagadnieniem, wcześniejszą wiedzą lub też hipotezą roboczą, którą należy skonfrontować z doświadczeniem.

Wróćmy do przykładu z p. 3.2. Można postawić pytanie, czy da się dla wyników w 7 konkurencjach lekkoatletycznych znaleźć dwa takie czynniki, z których pierwszy nie korelowałby z biegami, a drugi nie korelowałby z rzutami. Równania byłyby wtedy następujące:

$$
\begin{aligned}
X_1 &= & & l_{12}F_2 + v_1U_1 \\
X_2 &= & & l_{22}F_2 + v_2U_2 \\
X_3 &= l_{31}F_1 &+& l_{32}F_2 + v_3U_3 \\
X_4 &= l_{41}F_1 &+& l_{42}F_2 + v_4U_4 \\
X_5 &= l_{51}F_1 & & \quad\quad\;\; + v_5U_5 \\
X_6 &= l_{61}F_1 & & \quad\quad\;\; + v_6U_6 \\
X_7 &= l_{71}F_1 & & \quad\quad\;\; + v_7U_7
\end{aligned}
$$

a macierz ładunków miałaby postać:

$$
L = \begin{bmatrix} 0 & l_{12} \\ 0 & l_{22} \\ l_{31} & l_{32} \\ l_{41} & l_{42} \\ l_{51} & 0 \\ l_{61} & 0 \\ l_{71} & 0 \end{bmatrix}
$$

3.3.3. Modele z czynnikami skorelowanymi

Już Thurstone, twórca współczesnej analizy czynnikowej, dyskutował założenie o nieskorelowaniu czynników wspólnych wyrażając pogląd, że w niektórych sytuacjach należałoby także rozpatrywać skorelowane czynniki wspólne /Thurstone,1947/. Oznaczmy współczynnik korelacji między czynnikami F_k i F_l przez m_{kl}, czyli Kor $/F_k, F_l/ = m_{kl}$. Macierz korelacji dla wszystkich czynników wspólnych oznaczmy przez M, czyli

$$M = \begin{bmatrix} 1 & m_{12} & \dots & m_{1r} \\ m_{21} & 1 & \dots & m_{2r} \\ \cdot & \cdot & \cdot & \cdot \\ m_{r1} & m_{r2} & \dots & 1 \end{bmatrix}$$

Wtedy równania podstawowe przyjmą postać.

$$/4.3.6/ \qquad C = LML' + V$$

Jeżeli macierz M jest nieznana, to klasa możliwych rozwiązań będzie bardzo liczna i dla wyboru jednego z nich trzeba sformułować dodatkowe założenia. Moim zdaniem wprowadzenie czynników skorelowanych wymaga gruntownie przemyślanych i ściśle sformułowanych hipotez badawczych, opartych o wcześniejsze badania empiryczne.

4. Metody estymacji

4.1. Zagadnienie estymacji.

Podstawowe założenia analizy czynnikowej /3.2.13/ wiążą ładunki czynników wspólnych i specyficznych z wariancjami i kowariancjami zespołu cech. Te ostatnie jako parametry populacji są nieznane, możemy jedynie obliczyć ich oszacowanie z losowo wybranej próbki. Jak wiadomo, oszacowania takie nie są dokładne, równe parametrom w populacji, a różnią się od nich o wielkość zwaną błędem próbki. Dlatego też, jeżeli w równaniach podstawowych /3.2.13/ napisać zamiast prawdziwych wariancji i kowariancji ich oszacowanie, trzeba równość zastąpić znakiem równości przybliżonej. To samo dotyczy sytuacji, gdy podstawą analizy są współczynniki korelacji.

Dla wyjaśnienia na czym polega estymacja posłużymy się następującym rozumowaniem.

Niech będzie ustalona liczba czynników wspólnych $r < p$. Oznaczmy przez

$$\hat{L} = \begin{bmatrix} \hat{l}_{11} & \hat{l}_{12} & \dots & \hat{l}_{1r} \\ \hat{l}_{21} & \hat{l}_{22} & \dots & \hat{l}_{2r} \\ \dots & \dots & \dots & \dots \\ \hat{l}_{p1} & \hat{l}_{p2} & \dots & \hat{l}_{pr} \end{bmatrix} \quad \text{i} \quad \hat{V} = \begin{bmatrix} \hat{v}_1^2 & 0 & \dots & 0 \\ 0 & \hat{v}_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \hat{v}_p^2 \end{bmatrix}$$

macierze nieznanych jeszcze oszacowań ładunków i wariancji specyficznych. Sumę kwadratów wyrazów i-tego wiersza macierzy $\hat{L}$ powiększoną o $\hat{v}_i^2$ oznaczymy przez q_{ii}

czyli $$q_{ii} = \hat{l}_{i1}^2 + \hat{l}_{i2}^2 + \dots + \hat{l}_{ip}^2 + \hat{v}_i^2$$

Sumę iloczynów wyrazów dwóch różnych wierszy macierzy $\hat{L}$ oznaczymy przez q_{kl}, czyli

$$q_{kl} = \hat{l}_{kl}\hat{l}_{1l} + \hat{l}_{k2}\hat{l}_{12} + \ldots + \hat{l}_{kr}\hat{l}_{lr}$$

Obliczone w ten sposób wartości q_{ii} oraz q_{kl} będziemy odpowiednio nazywali odtworzonymi wariancjami lub kowariancjami.

Estymacja ładunków i wariancji specyficznych polega na obliczeniu z empirycznych wariancji i kowariancji takiej macierzy ładunków $\hat{L}$ i macierzy wariancji specyficznych $\hat{V}$, by odtworzone z nich wariancje i kowariancje różniły się od obliczonych mało w określonym sensie. W zależności od kryterium określającego zgodność otrzymuje się różne sposoby estymacji. Obliczone liczby $\hat{l}_{ik}$ oraz $\hat{v}_i^2$ nazywamy estymatorami ładunków i wariancji specyficznych, lub empirycznymi ładunkami i wariancjami specyficznymi.

4.2. Metody estymacji. I.

Ponieważ obliczenia związane z estymacją wykonywane są z zasady przez maszyny, nie będę w niniejszym artykule mówił o sprawach obliczeniowych ani algorytmach. Bieżący paragraf zawiera przegląd najczęściej stosowanych metod estymacji oraz omówienie założeń i najważniejszych konsekwencji. W paragrafie następnym natomiast będzie przeprowadzona szczegółowa dyskusja, już przy większym użyciu aparatu matematycznego.

Jedną z pierwszych metod estymacji ładunków jest metoda centroidalna wymyślona przez Thurstone`a /1947/. Jest to metoda rachunkowo dość prosta ale żmudna. Ponieważ istnieją już metody estymacji o lepszych właściwościach, metody centroidalnej na ogół nie stosuje się.

Zastosowanie do estymacji ładunków znanej w statystyce metody minimum kwadratów pochodzi od Thomsona /1950/. Metoda polega na takim wyznaczeniu ładunków, by suma kwadratów różnic między obliczonymi i odtworzonymi kowariancjami /lub korelacjami/

była możliwie mała, czyli by wyrażenie

$$\sum_{k}\sum_{\neq l} /c_{kl} - q_{kl}/^2$$

osiągnęło minimum. Postulat ten prowadzi do układu równań, którego rozwiązaniami są ładunki i wariancje specyficzne. Metodę tę będziemy krótko nazywali metodą Thomsona. Metoda ta zależy od jednostek, w jakich mierzono cechy. W szczególności daje ona inne wyniki zastosowana do macierzy korelacji a inne dla macierzy kowariancji. Ważną właściwością tej metody jest to, że sumy iloczynów ładunków każdej pary czynników są równe zeru. Wektory ładunków są więc ortogonalne.

Lawley /1963/ rozwiązał zagadnienie estymacji przy pomocy znanej w statystyce metody maximum wiarogodności. Metodę tę będziemy nazywali metodą ML I /od maximum likelihood/. Estymatory uzyskane tą metodą są niezależne od jednostek w jakich wyrażone są cechy. W szczególności daje ona te same wyniki zastosowana do macierzy kowariancji co zastosowana do macierzy korelacji.

Zastosowanie tej samej zasady estymacji maksimum wiarogodności do modelu o równych wariancjach specyficznych prowadzi do wyników związanych ze składowymi głównymi Hotellinga /Hotelling, 1933, Anderson, 1958/. Dlatego też metodę tę często nazywa się metodą Hotellinga. Ponieważ autorem tej metody jest także Lawley /1963/ będziemy ją nazywali metodą ML II. Metoda ta jest bardzo dogodna do obliczeń maszynowych. Jej wadą jest to, że wyniki zależą od jednostek, w jakich wyrażone są cechy.

Metodą, która łączy obie zalety - prostotę obliczeniową i niezależność wyników od jednostek - jest metoda stworzona przez K.G. Jöreskoga /1963/.

4.3. Metody estymacji. II.

Dla czytelnika bardziej zainteresowanego w metodach estymacji przedstawię w tym paragrafie równania służące do estymacji ładunków oraz interpretację matematyczną poszczególnych metod.

4.3.1. Metoda ML II /Lawley, 1963/

Metodą ML II nazywaliśmy estymowanie ładunków i wariancji specyficznej metodą maksimum wiarogodności w modelu z równymi wariancjami specyficznymi.

Oznaczmy przez $\hat{\lambda}_1 > \hat{\lambda}_2 > \ldots > \hat{\lambda}_r > \hat{\lambda}_{r+1} > \ldots > \hat{\lambda}_p > 0$

wartości własne empirycznej macierzy kowariancji $\hat{C}$ uporządkowane według wielkości. Oszacowaniem wariancji specyficznej /z założenia takiej samej dla wszystkich cech/ jest średnia arytmetyczna /p-r/ najmniejszych wartości własnych, czyli

/4.3.1/ $$\hat{v}^2 = \hat{\lambda}_0 = \frac{1}{p-r} / \hat{\lambda}_{r+1} + \hat{\lambda}_{r+2} + \ldots + \hat{\lambda}_p /$$

Niech

/4.3.2/ $$\underline{a}_k = \begin{bmatrix} a_{1k} \\ a_{2k} \\ \cdot \\ \cdot \\ \cdot \\ a_{pk} \end{bmatrix}$$

oznacza wektor własny macierzy $\hat{C}$ odpowiadający wartości własnej λ_k. Wektory te są wzajemnie prostopadłe i tak wyznaczone, by ich długość była równa jeden.

Ładunki k-tego czynnika wspólnego określone są następująco:

/4.3.3./ $$\hat{l}_{1k} = \sqrt{\hat{\lambda}_k - \hat{\lambda}_0}\; a_{1k}$$

$$\hat{l}_{2k} = \sqrt{\hat{\lambda}_k - \hat{\lambda}_0}\, \hat{a}_{2k}$$

$$\cdots$$

$$\hat{l}_{pk} = \sqrt{\hat{\lambda}_k - \hat{\lambda}_0}\, a_{pk}$$

W metodzie ML II ładunki k-tego czynnika są więc proporcjonalne do współrzędnych k-tego wektora własnego. Współczynnikiem proporcjonalności jest $\sqrt{\hat{\lambda}_k - \hat{\lambda}_0}$

Dla obliczonych tą metodą ładunków okazuje się, że suma iloczynów ładunków dwóch różnych czynników jest równa zeru, czyli $\sum_{i=1}^{p} \hat{l}_{ik} \hat{l}_{il} = 0$ jeżeli $k \neq l$. Wektory ładunków są więc prostopadłe.

4.3.2. Metoda Thomsona

Równania służące do obliczenia ładunków tą metodą są w zapisie macierzowym następujące /Anderson, Rubin, 1956/:

/4.3.4/
1. $\hat{L} / \hat{L}' \hat{L} / = / \hat{C} - \hat{V} / \hat{L}$
2. $\operatorname{diag} \hat{V} = \operatorname{diag} / \hat{C} - \hat{L} \hat{L}' /$
3. macierz $\hat{L}' \hat{L}$ jest macierzą przekątniową.

Macierz $\hat{L}$ jest macierzą szukanych ładunków, a $\hat{V}$ macierzą szukanych wariancji specyficznych. Macierz $\hat{L}' \hat{L}$ jest macierzą sum kwadratów i sum iloczynów ładunków. Warunek, by macierz ta była macierzą przekątniową, oznacza, że wektory ładunków mają być prostopadłe. Układ powyższych równań rozwiązuje się metodą kolejnych przybliżeń. Interpretacja wyników tej metody jest następująca. Oznaczamy przez $\hat{v}_1^2, \hat{v}_2^2, \ldots, \hat{v}_p^2$, wariancje specyficzne obliczone metodą Thomsona. Odejmując je od wariancji otrzymujemy zredukowaną macierz kowariancji. Okazuje się, że ładunki czynników wspólnych są proporcjonalne do wektorów własnych zredukowanej macierzy kowariancji.

4.3.3. Metoda ML I

Równania do wyznaczenia estymatorów są następujące:

$$1.\quad \hat{L}\,/I_r + \hat{L}'\,\hat{V}^{-1}\hat{L}/ = \hat{C}\,\hat{V}^{-1}\hat{L}$$

/4.3.5/
$$2.\quad \mathrm{diag}\,\hat{V} = \mathrm{diag}\,/\hat{C} - \hat{L}\,\hat{L}'/$$

3. macierz $\hat{L}'\hat{V}^{-1}\hat{L}$ jest macierzą przekątniową /Lawley i Maxwell, 1963, Kieloch i Oktaba, 1971/.

Interpretacja wyników tej metody estymacji jest następująca.

Rozpatrzmy macierz $\hat{\hat{C}} = \hat{V}^{-1/2}/\hat{C} - \hat{V}/\,\hat{V}^{-1/2}$.

Jest to zredukowana macierz kowariancji unormowana na dyspersje specyficzne. Niech $\gamma_1 > \gamma_2 > \ldots\ \gamma_r > 0$ będą wartościami własnymi tej macierzy, a $\underline{\hat{\hat{a}}}_1, \underline{\hat{\hat{a}}}_2 \ldots \underline{\hat{\hat{a}}}_r$ jej wektorami własnymi. Wtedy kolumny $\underline{\hat{l}}_k$ macierzy ładunków $\hat{L}$ wyrażają się wzorami

$$\underline{\hat{l}}_k = \begin{bmatrix} \hat{\hat{a}}_{1k}\sqrt{\lambda_k} \\ \hat{\hat{a}}_{2k}\sqrt{\lambda_k} \\ \cdot \\ \cdot \\ \cdot \\ \hat{\hat{a}}_{pk}\sqrt{\lambda_k} \end{bmatrix}$$

Macierz $\hat{L}'\hat{V}^{-1}\hat{L} = G$, gdzie G jest macierzą przekątniową. Wyrazami przekątnej są wartości własne $\gamma_1, \gamma_2, \ldots, \gamma_r$. Widzimy więc, że metoda maksimum wiarogodności daje ładunki proporcjonalne do wektorów własnych macierzy $\hat{\hat{C}}$.

Rozwiązanie równań /4.3.5./ odbywa się metodą kolejnych przybliżeń.

4.3.4. Metoda Jöreskoga.

Obliczenia wykonuje się na macierzy C^* określonej wzorem:

/4.3.7/
$$C^* = /\mathrm{diag}\,\hat{C}^{-1}/\,\tfrac{1}{2}\;\hat{C}\;/\mathrm{diag}\,\hat{C}^{-1}/\,\tfrac{1}{2}\,,$$

gdzie $\hat{C}$ może być macierzą kowariancji badanego zespołu cech wyrażonych w dowolnych jednostkach /Jöreskog, 1963/. Zapis /diag $\hat{C}^{-1}$/ $\frac{1}{2}$ oznacza macierz przekątniową, w której na głównej przekątnej znajdują się pierwiastki kwadratowe wyrazów głównej przekątnej macierzy $\hat{C}^{-1}$. Niech wartościami własnymi macierzy C^* będą liczby

$$\lambda_1^* > \lambda_2^* > \ldots > \lambda_r^* > \lambda_{r+1}^* > \ldots > \lambda_p^* > 0$$

którym odpowiadają wektory własne $\underline{a}_1^*$, $\underline{a}_2^*$, ..., $\underline{a}_p^*$.

Oznaczmy przez λ_0^* średnią arytmetyczną p-r najmniejszych wartości własnych. Przez $W_{/p \times r/}$ oznaczmy macierz, której kolumnami są wektory

$$\underline{w}_1 = \sqrt{\lambda_1^* - \lambda_0^*}\ \underline{a}_1^*$$

$$\underline{w}_2 = \sqrt{\lambda_2^* - \lambda_0^*}\ \underline{a}_2^*$$

. . . .

$$\underline{w}_r = \sqrt{\lambda_r^* - \lambda_0^*}\ \underline{a}_r^*$$

Macierz ładunków $\hat{L}$ otrzymuje się ze wzoru:

$$\hat{L} = /\mathrm{diag}\ \hat{C}^{-1}/^{-\frac{1}{2}}\ W.$$

Oszacowania wariancji specyficznych oblicza się według wzoru

$$\hat{v}_i^2 = c_{ii} - \sum_{k=1}^{r} \hat{l}_{ik}^2 = s_i^2 - \sum_{k=1}^{r} \hat{l}_{ik}^2$$

Śledząc powyższe wzory łatwo stwierdzić, że metoda ta jest w istocie metodą składowych głównych /ML II/ zastosowaną do macierzy C^*. Pomnożenie macierzy $\hat{C}$ przez /diag $\hat{C}^{-1}$/$^{1/2}$ gwarantuje niezależność od jednostek.

4.4. Wpływ jednostek pomiarowych

Omawiając metody estymacji zwrócono uwagę na to, że niektóre dają wyniki niezależne od tego w jakich jednostkach są wyrażone cechy i kowariancje. W tym paragrafie przedyskutujemy to zagadnienie szczegółowo. Weźmy pod uwagę jakiś zespół cech i obliczmy jego macierz kowariancji $\hat{C}$. Następnie wyraźmy wszystkie lub niektóre cechy tego zespołu w innych jednostkach, co odpowiada pomnożeniu wartości tych cech przez stałe. Otrzymamy wtedy inną macierz kowariancji $\hat{C}_1$. W szczególności, jeżeli cechy wyrazimy w dyspersjach, czyli gdy je unormujemy, nową macierzą kowariancji będzie macierz korelacji. Znając stałe,przez które dzielono lub mnożono cechy, łatwo z jednej macierzy kowariancji obliczyć drugą /bez sięgania do wartości cech/.Natomiast między wartościami własnymi oraz między wektorami własnymi obu macierzy nie ma naogół bezpośrednich związków.W szczególności więc różnią się wartości i wektory własne macierzy kowariancji i macierzy korelacji. Zarówno w metodzie Thomsona jak i w metodzie ML II /dla modelu o różnych wariancjach specyficznych/ ładunki są proporcjonalne do wektorów własnych macierzy kowariancji na której wykonano obliczenia. Stąd wniosek, że estymowane ładunki będą zależały od jednostek, w jakich wyraża się cechy.

Tej niedogodności nie ma metoda ML I. Metoda ta daje te same rozwiązania niezależnie od tego, czy stosuje się ją do macierzy korelacji, czy do macierzy kowariancji cech wyrażonych w dowolnych jednostkach. Właściwość ta wynika z tego,że ładunki obliczone metodą ML I są proporcjonalne do wektorów własnych macierzy kowariancji cech unormowanych na wariancje specyficzne, a więc na jednostki "naturalne" z punktu widzenia analizy czynnikowej. W metodzie Jöreskoga natomiast niezależność od jednostek uzyskuje się przez odpowiednie unormowanie macierzy,na której wykonuje się obliczenia /por.wzór 4.3.7/.

4.5. Przykład

W tablicy 3 podano empiryczną macierz korelacji wyników w 7 konkurencjach lekkoatletycznych: biegu na 100 m, biegu przez płotki, skoku wzwyż, skoku o tyczce, pchnięciu kulą, rzucie dyskiem, rzucie oszczepem /w konkurencjach biegowych zmieniono znaki/. Ilość badanych wynosiła 300.

Tablica 3. Macierz współczynników korelacji 7 konkurencji lekkoatletycznych.

	1	2	3	4	5	6	7
1	1.000	0,557	0,357	0,331	0,244	0,225	0,187
2	0,557	1.000	0,455	0,461	0,238	0,298	0,266
3	0,357	0,455	1.000	0,380	0,379	0,348	0,261
4	0,331	0,461	0,380	1.000	0,197	0,208	0,234
5	0,244	0,238	0,379	0,197	1.000	0.700	0,485
6	0,225	0,298	0,348	0,208	0.700	1.000	0,499
7	0,187	0,266	0,261	0,234	0,485	0,499	1.000

Metodą maksimum wiarogodności znaleziono ładunki dwóch czynników, które podano w tablicy 4.

Tablica 4. Ładunki 2 czynników i wariancje specyficzne obliczone metodą maksimum wiarogodności z macierzy korelacji 7 konkurencji lekkoatletycznych.

	$\hat{l}_{i1}$	$\hat{l}_{i2}$	$\hat{v}_i^2$
Bieg na 100 m	0,520	0,362	0,599
Bieg przez płotki	0,677	0,556	0,233
Skok wzwyż	0,577	0,142	0,647
Skok o tyczce	0,458	0,292	0,705
Pchnięcie kulą	0,729	-0,438	0,277
Rzut dyskiem	0,736	-0,374	0,318
Rzut oszczepem	0,558	-0,204	0,647

5. Testowanie hipotez dotyczących ilości czynników wspólnych

5.1. Zagadnienie ilości czynników

Przy omawianiu podstawowego modelu analizy czynnikowej zwróciliśmy uwagę na to, że jednym z podstawowych parametrów jest ilość czynników wspólnych. Zagadnieniem ilości czynników wspólnych nie zajmowaliśmy się przy omawianiu metod estymacji. Zakładaliśmy, że estymację ładunków i wariancji specyficznych wykonuje się przy z góry ustalonej ilość czynników. Niewątpliwym jest, że im większą przyjmujemy ilość czynników, tym większa będzie zgodność między obliczoną i odtworzoną macierzą kowariancji. Wynika stąd, że omówione metody estymacji nie pozwalają na oszacowanie ilości czynników wspólnych. Zagadnienie wyznaczenia ilości czynników, które ma dla interpretacji merytorycznej wyników zasadnicze znaczenie, poświęcono od powstania analizy czynnikowej dużo uwagi. Współcześnie podchodzi się do tego zagadnienia następująco: Ilość czynników traktujemy jako hipotezę którą należy sprowadzić odpowiednim testem statystycznym. Innymi słowy, traktujemy ilość czynników r jako parametr liczbowy i sprawdzamy hipotezę $r = r_0$, gdzie r_0 jest ilością czynników wspólnych, przy której przeprowadzono estymację. Jeżeli test odrzuca tę hipotezę, to na korzyść hipotezy, że ilość czynników jest większa niż r_0, czyli że $r > r_0$. Natomiast przyjęcie hipotezy oznacza, że już przy pomocy r_0 czynników można wytłumaczyć korelację między cechami z dokładnością do błędu próbki. Oczywiście czynników tych może być więcej niż r_0, ale uzasadnienie tego wymagałoby bądź argumentacji spoza analizy czynnikowej, bądź przeprowadzenia nowych badań i nowej analizy.

Jako ilość czynników wspólnych przyjmuje się najmniejszą wartość r_0, przy której test nie odrzuca hipotezy o równości. Postępowanie takie jest postępowaniem wielodecyzyjnym. Łączny poziom istotności takiego wielokrotnego postępowania jest trudny do ocenienia, napewno jednak jest większy niż przyjęty poziom w pojedyńczym teście.

Jak zobaczymy dalej, test zależy od ilości badanych osób. Okazuje się, że wartość statystyki testującej wzrasta z ilością badanych osób, a co zatem idzie, ilość czynników wykrywanych jako istotne, wzrasta z licznością próby. Inaczej mówiąc, jeżeli z dwóch badań obliczymy takie same macierze kowariancji i zastosujemy do nich tę samą metodę estymacji, to może się okazać, że test uzna w materiale z liczniejszych badań więcej czynników wspólnych za istotne, niż w materiale z badań obejmujących mniej osób.

Jeżeli test odrzuci hipotezę, że ilość czynników jest równa liczbie r_o, przy której przeprowadzono estymację, wtedy trzeba przeprowadzić estymację ponownie przy większej ilości czynników. Wymaga to w niektórych metodach estymacji /ML I, metoda Thomsona/ powtórzenia wszystkich obliczeń, w innych metodach powtórzenia niewielkich partii obliczeń /w metodzie ML II i metodzie Jöreskoga/.

5.2. Testy.

Oznaczmy jak poprzednio przez $\hat{C}$ obliczoną a przez Q odtworzoną macierz kowariancji. Macierz $\hat{C}$ można przedstawić jako sumę dwóch macierzy: macierzy $\hat{L}\hat{L}'$ i macierzy, która na głównej przekątnej ma wariancje specyficzne a poza nią różnice między obliczonymi a odtworzonymi kowariancjami, oznaczone symbolami d_{kl}:

$$/5.2.1/ \quad \begin{bmatrix} c_{11} & c_{12} & \dots c_{1p} \\ c_{21} & c_{22} & \dots c_{2p} \\ \cdot & \cdot & \cdot \\ c_{p1} & c_{p2} & \dots c_{pp} \end{bmatrix} = \begin{bmatrix} \hat{l}_{1k}^2 & \hat{l}_{1k}\hat{l}_{2k} \dots & \hat{l}_{1k}\hat{l}_{pk} \\ \hat{l}_{1k}\hat{l}_{2k} & \hat{l}_{2k}^2 \quad \dots & \hat{l}_{2k}\hat{l}_{pk} \\ \cdot & \cdot & \cdot \\ \hat{l}_{1k}\hat{l}_{pk} & \hat{l}_{2k}\hat{l}_{pk} \dots & \hat{l}_{pk}^2 \end{bmatrix} +$$

$$+\begin{bmatrix} \hat{v}_1^2 & d_{12} & \dots & d_{1p} \\ d_{21} & \hat{v}_2^2 & \dots & d_{2p} \\ \dots & \dots & \dots & \dots \\ d_{p1} & d_{p2} & \dots & \hat{v}_p^2 \end{bmatrix}$$

Macierz

$$\begin{bmatrix} \hat{v}_1^2 & d_{12} & \dots & d_{1p} \\ d_{21} & \hat{v}_2^2 & \dots & d_{2p} \\ \dots & \dots & \dots & \dots \\ d_{p1} & d_{p2} & \dots & \hat{v}_p^2 \end{bmatrix} \qquad /5.2.2/$$

jest empiryczną macierzą kowariancji czynników specyficznych $U_1, U_2, \dots, U_p$. Można z niej łatwo obliczyć empiryczną macierz korelacji czynników specyficznych. Będzie to macierz

$$\begin{bmatrix} 1 & \frac{d_{12}}{\hat{v}_1 \hat{v}_2} & \dots & \frac{d_{1p}}{\hat{v}_1 \hat{v}_p} \\ \frac{d_{12}}{\hat{v}_1 \hat{v}_2} & 1 & \dots & \frac{d_{2p}}{\hat{v}_2 \hat{v}_p} \\ \dots & \dots & \dots & \dots \\ \frac{d_{1p}}{\hat{v}_1 \hat{v}_p} & \frac{d_{2p}}{\hat{v}_2 \hat{v}_p} & \dots & 1 \end{bmatrix} \qquad /5.2.3/$$

Jeżeli model jest słuszny, to czynniki specyficzne są nieskorelowane w populacji. Empiryczne współczynniki korelacji tych czynników, czyli liczby $\frac{d_{kl}}{\hat{v}_k \hat{v}_l}$ muszą być bliskie zeru. Innymi słowy, jeżeli hipoteza jest słuszna, wyrazy poza główną przekątną mogą tylko na tyle odbiegać od zera, na ile pozwala błąd próbki.

Testowanie hipotezy, że ilość czynników wspólnych równa się liczbie r_o sprowadza się właśnie do testowania, czy empiryczna macierz korelacji czynników specyficznych jest estymatorem macierzy korelacji zmiennych nieskorelowanych, czyli macierzy przekątniowej o jedynkach na głównej przekątnej.

Niech łączny rozkład rozpatrywanego zespołu p cech będzie p - wymiarowym rozkładem normalnym. Przypuśćmy, że macierz kowariancji $\hat{C}$ obliczona została z próbki losowej liczącej /n+1/ osób. Macierz ładunków $\hat{L}$ i wariancje specyficzne obliczone zostały metodą ML I przy założeniu, że ilość czynników wspólnych równa się r_o. Rozpatrzmy macierz

$$\hat{V}^{-\frac{1}{2}} \hat{C} \hat{V}^{-\frac{1}{2}}$$

Niech jej wartościami własnymi będą liczby:

$$\tilde{\lambda}_1 \geqslant \tilde{\lambda}_2 \geqslant \ldots \geqslant \tilde{\lambda}_{r_o} \geqslant \tilde{\lambda}_{r_o+1} \geqslant \ldots \geqslant \tilde{\lambda}_p > 0$$

Tworzy się wielkość

$$\gamma = -n \sum_{i=r_o+1}^{p} \log \tilde{\lambda}_i$$

Wiadomo, że jeżeli spełnione są założenia /3.2.1/ - /3.2.7/ określające model, to wielkość γ ma asymptotycznie /przy nieograniczenie wielkim n/ rozkład chi - kwadrat o $d_f = \frac{1}{2}\left\{ /p-r_o/^2 - /p+r_o/ \right\}$ stopniach swobody /Lawley, Maxwell, 1963/. Wielkość γ może więc służyć do testowania hipotezy, że ilość czynników wspólnych $r = r_o$. Wartość krytyczną odczytuje się z tablic rozkładu chi-kwadrat na przyjętym poziomie istotności przy podanej powyżej ilości stopni swobody df. Można także użyć testu przybliżonego: obliczamy empiryczną macierz korelacji czynników specyficznych według wzoru /5.2.3/.

Wielkość $\gamma_1 = n \sum_{k<l} \sum \left(\frac{d_{kl}}{\hat{v}_k \; \hat{v}l} \right)^2$ ma w przybliżeniu także rozkład chi-kwadrat o $df = \frac{1}{2} \{ /p-r_o/^2 - /p+r_o/ \}$ stopniach swobody /Lawley, Maxwell, 1963/. Wielkość γ jest n-krotną sumą kwadratów wyrazów nad /lub pod/ główną przekątną macierzy korelacji czynników specyficznych.

Jeden i drugi test można stosować, jeżeli użyto do estymacji ładunków i wariancji specyficznych metody maksimum wiarogodności. Jeżeli wielkości te estymowano inną metodą, stosowanie tych testów może prowadzić do błędnych wniosków.

Omówimy teraz testowanie hipotezy dotyczącej liczby czynników w modelu z równymi wariancjami specyficznymi. Model ten zakłada, że wartości własne macierzy kowariancji w populacji od /r+1/ - tego poczynając są równe, czyli

$$\lambda_{r+1} = \lambda_{r+2} = \ldots = \lambda_p.$$

A zatem, jeżeli model jest słuszny, to p-r najmniejszych wartości własnych empirycznej macierzy kowariancji może różnić się tylko wskutek błędu próbki. Testy do sprawdzania takiej hipotezy pochodzą od Bartletta /1950, 1953/. Statystyką do sprawdzenia hipotezy, że ilość czynników wspólnych równa się r_o, jest wielkość

$$\chi^2 = n' \left[(p-r_o) \ln \frac{1}{p-r_o} \sum_{i=r_o+1}^{p} \lambda_i - \sum_{i=r_o+1}^{p} \ln \lambda_i \right]$$

gdzie $n' = \left(n - r_o - \frac{1}{6} \right) \left[2/p-r_o/ + 1 + \frac{2}{p-r_o} \right]$

Statystyka ta, przy takich samych założeniach co poprzednio, ma w przybliżeniu rozkład chi-kwadrat o $df = \frac{1}{2}\left\{ /p-r_o + 2/ \times /p-r_o - 1/ \right\}$ stopniach swobody.

Podobne testy stosuje się, jeżeli obliczenia wykonano na macierzy korelacji.

W tablicy 5 podano macierz, w której na głównej przekątnej znajdują się wariancje specyficzne, a poza nią różnice między obliczonymi i odtworzonymi korelacjami. W tablicy 6 natomiast podano obliczoną z niej macierz empirycznych współczynników korelacji czynników specyficznych.

Tablica 5. Różnica między faktyczną i odtworzoną macierzą korelacji powiększona o wariancję specyficzną na głównej przekątnej.

1	0,599	0,004	0,006	-0,013	0,023	-0,022	-0,029
2	0,004	0,233	-0,015	-0,011	-0,012	0,008	0,002
3	0,006	-0,015	0,647	0,074	0,021	-0,024	-0,032
4	-0,013	-0,011	0,074	0,705	-0,009	-0,020	0,038
5	0,023	-0,012	0,021	-0,009	0,277	0,000	-0,011
6	-0,022	0,008	-0,024	-0,020	0,000	0,318	0,012
7	-0,029	0,002	-0,032	0,038	-0,011	0,012	0,647

Tablica 6. Macierz empirycznych współczynników korelacji czynników specyficznych /korelacji reszt/.

1	1,000	0,011	0,010	-0,020	0,056	-0,050	-0,047
2	0,011	1.000	-0,039	-0,027	-0,047	0,029	0,005
3	0,010	-0,039	1.000	0,110	0,050	-0,053	-0,049
4	-0,020	-0,027	0,110	1.000	-0,020	-0,042	0,056
5	0,056	-0,047	0,050	-0,020	1.000	0,000	-0,026
6	-0,050	0,029	-0,053	-0,042	0,000	1.000	0,026
7	-0,047	0,005	-0,049	0,056	-0,026	0,026	1,000

Suma kwadratów wyrazów nad główną przekątną wynosi:

$$\sum_{k} < \sum_{l} \left(\frac{d_{kl}}{\hat{v}_k \quad \hat{v}_l} \right)^2 = 0{,}04025$$

Ponieważ ilość badanych liczyła 300 osób

$$\gamma_1 = 12.08$$

Ilość stopni swobody df = 8

Wartość krytyczna odczytana z tablicy rozkładu chi-kwadrat

$$\chi^2_{0.05;8} = 15.507$$

Ponieważ obliczona wartość chi-kwadrat nie przekracza wartości krytycznej, godzimy się na to, że ilość czynników wspólnych wynosi 2. Przeprowadzając ten sam rachunek przy wyestymowanych ładunkach jednego tylko czynnika otrzymano $\gamma_1 = 300 \cdot 0{,}822 = 246{,}6$ co przekraczało wartość krytyczną wynoszącą

$$\chi^2_{0{,}05;\ 14} = 23{,}685$$

6. Rotacja czynników

6.1. Interpretacja geometryczna czynników.

Przyjęła się następująca pożyteczna interpretacja geometryczna ładunków. Weźmy pod uwagę ładunki dwóch czynników wspólnych. Narysujmy prostokątny układ współrzędnych. Oś poziomą nazwijmy osią pierwszego czynnika F_1, oś pionową osią drugiego czynnika F_2. Każdej cesze odpowiadają dwa ładunki: ładunek czynnika pierwszego l_{i1} i ładunek czynnika drugiego l_{i2}. Traktując te ładunki jako współrzędne na płaszczyźnie, każdej cesze będzie odpowiadał punkt $P_i = /l_{i1}, l_{i2}/$ na płaszczyźnie F_1, F_2 lub - jeżeli punkt ten połączymy z początkiem układu - wektor X_i o składowych $/l_{i1}, l_{i2}/$. Rysunek wszystkich takich wektorów ilustruje tzw. strukturę czynnikową. Podobny rysunek można sporządzić dla ładunków unormowanych, tzn. dla ładunków podzielonych przez pierwiastek z ich sumy kwadratów $\frac{l_{ik}}{h_i}$. W tym przypadku wszystkie punkty będą leżały na kole o promieniu 1.

Na rysunku 1 przedstawiono strukturę obu czynników dla 7 konkurencji lekkoatletycznych /tablica 4/.

Jeżeli w zespole cech wyodrębniono 3 czynniki, dla zilustrowania struktury czynnikowej trzeba by sporządzić model trójwymiarowy, przedstawiający punkty lub wektory odpowiadające cechom w układzie trzech osi prostopadłych odpowiadających czynnikom F_1, F_2 i F_3. Dla takich zespołów cech, dla których wyodrębniono więcej czynników narysowanie lub zbudowanie modelu ilustrującego strukturę jest niemożliwe, ale zachowamy koncepcję takiej interpretacji geometrycznej. Będziemy mianowicie identyfikowali cechy z wektorami w przestrzeni o tylu wymiarach, ile jest czynników wspólnych. Wektory te mają składowe równe ładunkom /lub ładunkom unormowanym/ kolejnych czynników. Pomocą w poglądowym przedstawieniu są rysunki wektorów cech robione dla poszczególnych par czynników.

6.2. Przekształcenia dopuszczalne macierzy ładunków.

W paragrafie 3.2 wspomnieliśmy, że rozwiązania podstawowego równania są niejednoznaczne. To znaczy, że jeżeli przy ustalonej ilości czynników r istnieje macierz ładunków spełniająca podstawowe równanie, to istnieje nieskończenie wiele innych macierzy spełniających to równanie. Znając jedno rozwiązanie można w sposób prosty otrzymać wszystkie pozostałe. Wszystkie te rozwiązania są jednakowo dobre z matematycznego punktu widzenia.

To samo zachodzi także dla macierzy ładunków obliczonej jedną z metod estymacji.

Ponieważ jednak w poszczególnych rozwiązaniach czynniki różnią się ładunkami, każde rozwiązanie określa układ czynników o innej interpretacji przyrodniczej. Mamy więc taką sytuację,że zagadnienie ma nieskończenie wiele rozwiązań jednakowych pod względem matematycznym ale o różnej interpretacji przyrodniczej. Wybór ostatecznego rozwiązania pozostawia się przyrodnikowi. Ta sytuacja ściągnęła na analizę czynnikową opinię metody subiektywnej.

Dla uzyskania jednoznaczności rozwiązania istnieją dwie drogi: Albo sformułować zagadnienie przyrodnicze tak, by wynikało z tego rozwiązanie jednoznaczne, co można osiągnąć przez nałożenie szczegółowych warunków dotyczących struktury macierzy ładunków. Albo przyjąć pewne leżące poza zagadnieniem przyrodniczym kryteria formalne, jakie ma spełniać macierz ładunków. W dalszym ciągu zajmiemy się rozwiązaniem drugim. Przedtem jednak omówimy działania, jakim można poddać macierz ładunków, by jednak nie przestała ona być rozwiązaniem równania zasadniczego. Przekształcenia te będziemy nazywali przekształceniami dopuszczalnymi.

Weźmy pod uwagę ładunki ustalonego czynnika, tzn. jedną kolumnę macierzy ładunków. W tej kolumnie zmieńmy wszystkie znaki na przeciwne. Ponieważ ładunki są współczynnikami korelacji czynnika z cechami, taka zmiana znaków odpowiada zmianie znaku czynnika i jego interpretacji. Do wzorów równania podstawowego ładunki te wchodzą zawsze jako iloczyny dwóch z nich. Jeżeli zaś w iloczynie zmienić znaki obu czynników, to wynik nie zmieni się. Stąd wniosek, że po zmianie znaku całej kolumny macierzy ładunków otrzymamy nową macierz, która także spełnia równania podstawowe. Zmiana znaku jednej lub kilku kolumn macierzy ładunków jest więc przekształceniem dopuszczalnym macierzy ładunków. Geometrycznie takiej zmianie znaków odpowiada zmiana zwrotu osi odpowiadającej czynnikowi.

Drugie przekształcenie dopuszczalne omówimy najpierw geometrycznie. Weźmy pod uwagę ładunki dwóch czynników F_{k_1} i F_{k_2}. Narysujmy dwie osie prostopadłe oznaczone F_1 i F_2. W takim układzie współrzędnych narysujmy wektory odpowiadające cechom o składowych równych ładunkom tych czynników /rys.1/. Na tym samym rysunku narysujmy dwie inne osie F_1^* i F_2^* obrócone o kąt α w stosunku do osi F_1 i F_2. Rzutując końce wektorów na oś F_1^* i F_2^* możemy odczytać nowe współrzędne l_{i1}^* oraz l_{i2}^*. Jeżeli teraz w wyjściowej macierzy ładunków L zastąpimy kolumny l_{i1} oraz l_{i2} odczytanymi wartościami l_{i1}^* i l_{i2}^* otrzymamy nową macierz ładunków L^*. Macierz L^* także spełnia równania podstawowe. Można się o tym przekonać następująco: wyrażenia $\sum_{s=1}^{r} l_{ks} l_{ls}$ są równe iloczynom skalarnym. Iloczyny skalarne zależą tylko od długości i wzajemnego położenia obu wektorów, a nie od umieszczenia układu współrzędnych. Ponieważ przez obrót układu współrzędnych nie zmieniła się długość ani położenie wektorów, więc taka zmiana współrzędnych nie zmienia iloczynu skalarnego. Dlatego L^* także spełnia równania podstawowe.

Całe to postępowanie można wykonać rachunkowo. Oznaczmy kąt obrotu przez α. Wtedy nowe ładunki oblicza się według wzorów

$$l_{i1} = l_{i1} \cos\alpha + l_{i2} \sin\alpha$$

$$l_{i2} = l_{i1} \cdot /-\sin\alpha/ + l_{i2} \cos\alpha \qquad /6.2.1/$$

Widzimy więc, że dopuszczalnym przekształceniem macierzy L są obroty w płaszczyźnie wyznaczonej przez dowolne dwa czynniki. W ogólnym przypadku r czynników wspólnych wyobraźmy sobie obrót całego układu współrzędnych $F_1, F_2, \ldots, F_r$ dokoła początku układu do nowego położenia $F_1^*, F_2^*, \ldots, F_r^*$. Współrzędne wektorów odpowiadające cechom w tym nowym układzie tworzą macierz ładunków L^*, która jest także rozwiązaniem podstawowego równania. Ponieważ każdy obrót układu współrzędnych można otrzymać przez kolejne obroty dwóch osi w płaszczyźnie przez nie wyznaczonej, można nowe ładunki obliczyć stosując wzory /6.2.1/ do różnych par czynników. Najważniejszym dla nas wnioskiem jest to, że drugim dopuszczalnym przekształceniem macierzy ładunków są obroty.

Przydatny będzie wzór ogólny na obliczenie macierzy ładunków po obrocie. Oznaczmy przez Θ macierz, której wyrazami są kosinusy kątów, jakie tworzą z sobą parami osie pierwotne i osie po obrocie. Macierz taka jest macierzą ortogonalną. Macierz ładunków po obrocie L oblicza się z macierzy wyjściowej według wzoru

/6.2.2./ $$L^{*} = L\,\Theta$$

Mnożąc więc macierz ładunków przez macierz ortogonalną otrzymujemy nową macierz ładunków, która także spełnia równania podstawowe.

6.3. Zagadnienie rotacji czynników.

Działanie matematyczne określone wzorem /6.2.2./ czyli pomnożenie macierzy ładunków przez macierz ortogonalną nazywa się rotacją czynników. Wybór określonej macierzy Θ_o, której odpowiada macierz

$$L_o = L\,\Theta_o$$

sensowna z punktu widzenia zastosowań, nazywa się zagadnieniem rotacji.

Thurston /1947/ sformułował jako kryterium wyboru ostatecznej macierzy ładunków L_o tzw. zasadę struktury prostej. Według niego, czynniki tworzą strukturę prostą, jeżeli poszczególne czynniki są skorelowane z niewielką ilością cech, ale z nimi wysoko, oraz jeżeli poszczególne cechy korelują z małą ilością czynników. Takiej strukturze odpowiada macierz ładunków, która w poszczególnych kolumnach ma dużo ładunków bliskich zeru i która ma dużo wierszy, w których są ładunki bliskie zera. Tego rodzaju macierze Thurston uważał za macierze ładunków o istotnym sensie przyrodniczym.

Intuicje Thurston'a zostały sprecyzowane przez O.Reiersola /Anderson, Rubin, 1956/. Kryteria te są jednak kryteriami jakościowymi, nie dającymi żadnych algorytmów dla znalezienia macierzy o strukturze prostej drogą rachunkową.

6.4. Zasada varimax

Spośród wielu zasad analitycznych na uwagę zasługuje zasada varimax sformułowana przez Kaisera /Harman, 1960/. Zwrot varimax jest skrótem słów maximum of the variance. Zasada ta brzmi następująco: Jako ostateczne rozwiązanie bierze się spośród wszystkich dopuszczalnych macierzy ładunków tę, dla której suma wariancji kwadratów ładunków dla wszystkich czynników osiąga maksimum. Suma wariancji kwadratów ładunków wyraża się wzorem:

/6.4.1./ $$K = \sum_{k=1}^{r} \frac{1}{p} \sum_{i=1}^{p} \left(l_{ik}^2 - \overline{l_k^2} \right)^2$$

gdzie $$\overline{l_k^2} = \frac{1}{p} \sum_{i=1}^{p} l_{ik}^2$$

Według zasady varimax należy więc wybrać jako rozwiązanie ostateczne tę macierz ładunków, dla której powyższa wartość K jest największa.

Wiemy, że dowolną dopuszczalną macierz ładunków można przedstawić w postaci $L^{*} = L\Theta$, gdzie Θ jest jakąś macierzą ortogonalną. Ładunki dowolnej dopuszczalnej macierzy są więc funkcją wyrazów macierzy Θ. Wyrażenie K jest więc także funkcją wyrazów Θ. Zagadnienie znalezienia maksimum tej wartości sprowadza się do znalezienia maksimum funkcji wielu zmiennych, które rozwiązuje się standardowymi metodami rachunku różniczkowego. Rozwiązanie wyznacza jednoznacznie macierz Θ_0 i w konsekwencji macierz $L_0 = L\,\Theta_0$.

Postulat, by suma wariancji kwadratów ładunków była możliwie duża, powoduje to, że w poszczególnych kolumnach ładunki stają się bądź bliskie zeru, bądź możliwie duże. W konsekwencji jako rozwiązanie otrzymuje się macierz ładunków o strukturze zbliżonej do prostej i łatwej do interpretacji.

Znalezienie macierzy obrotu dla dwóch czynników, która doprowadziłaby je do struktury spełniającej postulat varimax, nie wymaga dużej ilości obliczeń. Mogą one być wykonane na arytmometrze.

Oznaczmy ładunki obu czynników przez l_{i1} i l_{i2}. Ładunki te normuje się dzieląc je przez pierwiastek kwadratowy zasobu zmienności wspólnej. Otrzymane wielkości oznaczmy dla pierwszego czynnika przez x_i a dla drugiego przez y_i, czyli:

$$x_i = l_{i1} / h_i$$

$$y_i = l_{i2} / h_i$$

Następnie oblicza się:

$$u_i = x_i^2 - y_i^2$$

$$v_i = 2\,x_i y_i$$

$$A = \sum u_i$$

$$B = \sum v_i$$

$$C = \sum /u_i^2 - v_i^2/$$

$$D = 2 \sum u_i v_i$$

$$L = D - 2 AB /p$$

$$M = C - /A^2 - B^2/ p$$

Kąt obrotu wyznacza się z warunków

$$tg\, 4\alpha = L/M$$

Wartość kąta $\varphi = 4\alpha$ odczytuje się z tablic. Dzieląc odczytany kąt przez 4 otrzymuje się kąt α. Ze względu na niejednoznaczność rozwiązania takiego równania przy obliczaniu samego kąta α trzeba kierować się regułami podanymi w niżej podanej tablicy.

Znak L	Znak M	Znak tg 4α	przedziały dla kąta 4α	przedziały dla kąta α
+	+	+	$0 < 4\alpha < 90^o$	$0 < \alpha < 22,5^o$
+	-	-	$90^o < 4\alpha < 180^o$	$22,5 < \alpha < 45^o$
-	-	+	$-180^o < 4\alpha < -90^o$	$-45^o < \alpha < -22,5^o$
-	+	-	$-90^o < 4\alpha < 0^o$	$-22,5^o < \alpha < 0^o$

Macierz obrotu jest postaci:

$$\Theta = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

Nową macierz ładunku oblicza się według wzoru

$$L^{*} = \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \\ \vdots & \vdots \\ l_{p1} & l_{p2} \end{bmatrix} \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

lub według wzorów rozpisanych szczegółowo:

$$l^{*}_{i1} = l_{i1} \cos\alpha + l_{i2} \sin\alpha$$

$$l^{*}_{i2} = -l_{i1} \sin\alpha + l_{i2} \cos\alpha$$

Jako przykład niech posłuży wykonanie rotacji metodą varimax ładunków obu czynników obliczonych dla wyników w 7 konkurencjach lekkoatletycznych /Tablica 4/. Obliczenia ujęte są w schemacie:

l_{i1}	l_{i2}	h_i^2	h_i	x_i	y_i	u_i	v_i
0,520	0,362	0,401	0,634	0,821	0,571	0,348	0,938
0,677	0,556	0,767	0,876	0,773	0,635	0,194	0,981
0,577	0,142	0,353	0,594	0,971	0,239	0,886	0,464
0,458	0,292	0,295	0,543	0,843	0,538	0,422	0,907
0,729	-0,438	0,723	0,851	0,857	-0,515	0,469	-0,883
0,736	-0,374	0,682	0,826	0,891	-0,453	0,589	-0,807
0,558	-0,204	0,353	0,594	0,939	-0,343	0,764	-0,644

A = 3,672

B = 0,956 $\qquad$ L = -1,145

C = -2,453

D = -0,1417 $\qquad$ $\dfrac{A^2 - B^2}{p} = 1,796$

$\frac{2AB}{p} = 1,003$ $\qquad$ $M = -4,249$

$$tg\ 4\alpha = 0,2695$$

Ponieważ $L < 0$ i $M < 0$ kąt 4α ma zawierać się między -180^o a -90^o, sam zaś kąt α między -45^o a $-22,5^o$.

Z tablic wartości funkcji trygonometrycznych odczytuje się

$$4\alpha = 164,92^o$$
$$\alpha = -41,23^o$$
$$\cos\alpha = 0,7521$$
$$\sin\alpha = -0,6591$$

Macierzą obrotu jest więc macierz

$$\Theta = \begin{bmatrix} 0,7521 & 0,6591 \\ -0,6591 & 0,7521 \end{bmatrix}$$

Ładunek pierwszego czynnika po obrocie w pierwszej cesze obliczamy następująco:

$$l_{11}^{*} = l_{11} \cos\alpha + l_{12} \sin\alpha =$$
$$= 0,520 \cdot 0,7521 - 0,362 \cdot 0,6591 = 0,152$$

Ładunek drugiego czynnika po obrocie w pierwszej cesze obliczamy wg wzoru:

$$l_{12}^{*} = -l_{11} \sin\alpha + l_{12} \cos\alpha =$$
$$= + 0,520 \cdot 0,6591 + 0,362 \cdot 0,7521 = 0,615$$

Podobnie oblicza się ładunki obu czynników po obrocie w pozostałych cechach. Ładunki te są następujące:

l_{i1}^{*}	l_{i2}^{*}
0,152	0,615
0,143	0,864
0,340	0,487

0,152	0,521
0,837	0,151
0,800	0,204
0,554	0,214

Osie odpowiadające nowym czynnikom oznaczono na rysunku 1 przez F_1^* i F_2^*.

Okazuje się, że pierwszy czynnik koreluje wysoko z wynikami w pchnięciu kulą i rzucie oszczepem, niskie natomiast są jego korelacje z konkurencjami biegowymi. Drugi czynnik naodwrót, koreluje wysoko z wynikami w biegach, nisko natomiast z wynikami w rzutach. Upoważnia to do interpretowania czynnika pierwszego jako czynnika siły, a drugiego jako czynnika szybkości.

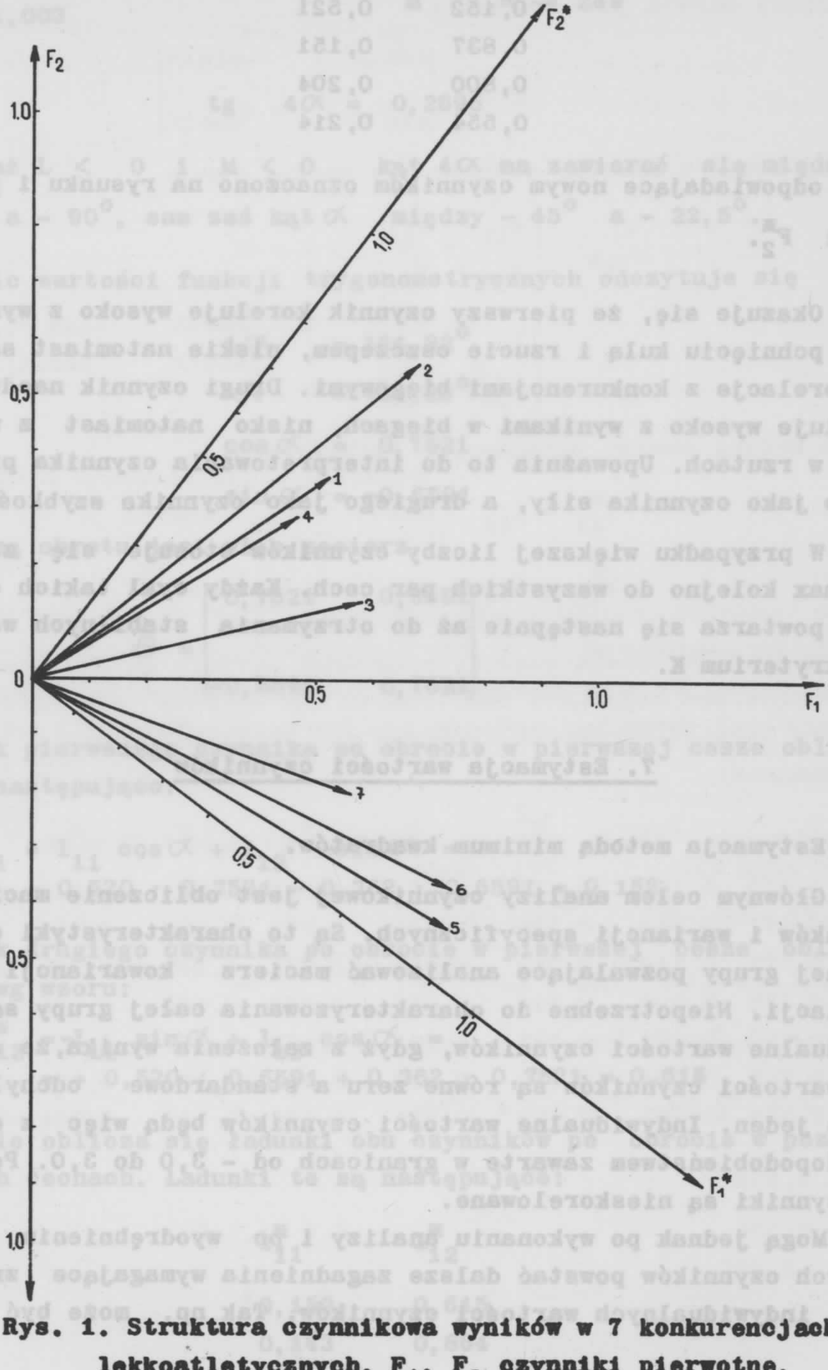
W przypadku większej liczby czynników stosuje się zasadę varimax kolejno do wszystkich par cech. Każdy cykl takich obliczeń powtarza się następnie aż do otrzymania stabilnych wartości kryterium K.

7. Estymacja wartości czynników

7.1. Estymacja metodą minimum kwadratów.

Głównym celem analizy czynnikowej jest obliczenie macierzy ładunków i wariancji specyficznych. Są to charakterystyki całej badanej grupy pozwalające analizować macierz kowariancji lub korelacji. Niepotrzebne do charakteryzowania całej grupy są indywidualne wartości czynników, gdyż z założenia wynika, że średnie wartości czynników są równe zeru a standardowe odchylenie równe jeden. Indywidualne wartości czynników będą więc z dużym prawdopodobieństwem zawarte w granicach od - 3,0 do 3,0. Ponadto czynniki są nieskorelowane.

Mogą jednak po wykonaniu analizy i po wyodrębnieniu sensownych czynników powstać dalsze zagadnienia wymagające znajomości indywidualnych wartości czynników. Tak np. może być wte-



Rys. 1. Struktura czynnikowa wyników w 7 konkurencjach lekkoatletycznych. F_1, F_2 czynniki pierwotne, F_1^*, F_2^* czynniki po rotacji

dy, gdy chcemy korelować wartości czynników z cechami, które nie weszły do analizy czynnikowej, badać średnie wartości czynników w określonych podgrupach lub też podejmować decyzje na podstawie wartości czynników.

Punktem wyjściowym do wyznaczenia indywidualnych wartości czynników są równania, które otrzymuje się wypisując wzory /3.2.2/ dla określonego obiektu. Równania te są następujące:

$$x_{1j} = l_{11} f_{1j} + l_{12} f_{2j} + \ldots + l_{1r} f_{rj} + v_1 u_{1j}$$

$$x_{2j} = l_{21} f_{1j} + l_{22} f_{2j} + \ldots + l_{2r} f_{rj} + v_2 u_{2j}$$

/7.1.1/...

$$x_{pj} = l_{p1} f_{1j} + l_{p2} f_{2j} + \ldots + l_{pr} f_{rj} + v_p u_{pj}$$

W równaniach tych j oznacza numer badanej osoby. Załóżmy że występujące w tych wzorach ładunki czynników wspólnych i specyficznych są znane. W praktyce zastępuje się je wartościami obliczonymi jedną z metod estymacji. Wtedy wzory te przedstawiają dla każdego osobnika układ równań. Niewiadomymi w tym układzie są wartości r czynników wspólnych f_{kj} oraz wartości p czynników specyficznych. Łatwo sprawdzić, że w układzie tym jest więcej niewiadomych niż równań, układ jest więc nieoznaczony. Jeżeli zrezygnujemy z obliczenia wartości czynników specyficznych, to układ staje się układem nadoznaczonym. Wynika stąd, że nie potrafimy dla badanego osobnika obliczyć jednoznacznie wartości czynników wspólnych. Jednym ze sposobów rozwiązania takiego zadania jest zastosowanie zasady minimum sumy kwadratów. Zgodnie z tą zasadą wyznaczmy wartości czynników tak, by na czynniki specyficzne przypadło możliwie mało, tzn. by wartość $\sum_{i=1}^{p} u^2_{ij}$ była możliwie mała dla każdej osoby.

Powyższy postulat można przepisać w postaci:

$$\sum_{i=1}^{p} \frac{\left(x_{ij} - \sum_{k=1}^{r} l_{ik} f_{kj}\right)^2}{v_i^2} = \text{minimum} \tag{7.1.2}$$

Idea takiego sposobu estymacji wartości czynników wspólnych pochodzi od Bartletta /1937/. Przy pomocy standardowych metod matematycznych oblicza się, że wartości czynników, które minimizują powyższe wyrażenie wyznaczone są wzorem:

$$\hat{\underline{f}}_j = /L' V^{-1} L/^{-1} L' V^{-1} /x_j - \bar{x}/, \tag{7.1.3}$$

gdzie $\hat{\underline{f}}_j$ jest wektorem kolumnowym oszacowanych wartości kolejnych czynników a $/\underline{x}_j - \bar{x}/$ jest wektorem kolumnowym odchyleń indywidualnych wartości kolejnych cech od wektora średnich.

W praktyce we wszystkich wzorach występujących w tym paragrafie macierze L i V zastępuje się ich estymatorami. Dla uproszczenia zapisu zrezygnowano z "daszków" oznaczających estymatory.

Jeżeli analizę czynnikową przeprowadzono na macierzy korelacji, wtedy zamiast wektora $/\underline{x}_j - \bar{x}/$ należy we wzorze /7.1.3/ podstawić wektor wartości unormowanych

$$\underline{x}_j^* \quad V = \begin{bmatrix} \dfrac{x_{1j} - \bar{x}_1}{s_1} \\ \dfrac{x_{2j} - \bar{x}_2}{s_2} \\ \cdots\cdots \\ \dfrac{x_{pj} - \bar{x}_p}{s_p} \end{bmatrix}$$

Z powyższego wzoru wynika bezpośrednio, że średnia wartość estymowanych wartości czynników równa się zeru. Macierz wariancji i kowariancji estymowanych wartości czynników oznaczmy przez M. Jest ona określona wzorem

$$M = I_r + /L' V^{-1} L/^{-1} \qquad /7.1.4/$$

Może się okazać, że estymowane wartości czynników są skorelowane, chociaż z założenia czynniki nimi nie są. Korelacja ta jest wynikiem estymacji czynników.

Jak wynika z /7.1.4/ właściwości statystyczne wyznaczonych wartości czynników zależą od właściwości macierzy ładunków występującej we wzorze. Podane wzory są wzorami ogólnymi. Rozpatrzmy przypadki szczegółowe.

1. Macierze L i V wyznaczone metodą ML I.

Wzory przyjmują wtedy postać:

$$\hat{f}_{1j} = \frac{1}{\gamma_1}\left[\frac{l_{11}}{v_1} \cdot \frac{x_{1j}-\bar{x}_1}{v_1} + \frac{l_{21}}{v_2} \cdot \frac{x_{2j}-\bar{x}_2}{v_2} + \ldots + \frac{l_{p1}}{v_p} \cdot \frac{x_{pj}-\bar{x}_p}{v_p}\right] \qquad /7.1.5/$$

$$\hat{f}_{2j} = \frac{1}{\gamma_2}\left[\frac{l_{12}}{v_1} \cdot \frac{x_{1j}-\bar{x}_1}{v_1} + \frac{l_{22}}{v_2} \cdot \frac{x_{2j}-\bar{x}_2}{v_2} + \ldots + \frac{l_{p2}}{v_p} \cdot \frac{x_{pj}-\bar{x}_p}{v_p}\right]$$

$$\cdots$$

$$\hat{f}_{rj} = \frac{1}{\gamma_r}\left[\frac{l_{1r}}{v_1} \cdot \frac{x_{1j}-\bar{x}_1}{v_1} + \frac{l_{2r}}{v_2} \cdot \frac{x_{2j}-\bar{x}_2}{v_2} + \ldots + \frac{l_{pr}}{v_p} \cdot \frac{x_{pj}-\bar{x}_p}{v_p}\right]$$

W tym przypadku wartości czynników są średnimi ważonymi wartości cech unormowanych na średnią zero i wyrażonych w dyspersjach specyficznych przy czym wagi są wprost proporcjonalne do ładunków a odwrotnie proporcjonalne do dyspersji specyficznych. Łatwo sprawdzić, że średnie wartości czynników są równe zeru. Wariancje wartości czynników wynoszą:

$$\mathrm{War}\ \hat{f}_k = \gamma_k^{-1}, \qquad k = 1, \ldots , r.$$

Wartości różnych czynników są nieskorelowane.

Jeżeli do estymacji wartości czynników użyto macierzy ładunków L^{*} po obrocie, to wartości czynników mogą stać się skorelowane.

2. Macierze L i V wyznaczone metodą ML II

Dla metody tej $v_k = \lambda_o$, $k = 1, \ldots, r$, oraz

$$L^{\prime} L = \begin{bmatrix} \lambda_1 - \lambda_o & 0 & \ldots & 0 \\ 0 & \lambda_2 - \lambda_o & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \lambda_r - \lambda_o \end{bmatrix}$$

Po przekształceniu wzoru /7.1.3/ otrzymujemy następujące wzory na wartości czynników wspólnych:

/7.1.6/

$$\hat{f}_{1j} = \frac{1}{\lambda_1 - \lambda_o}\Big[l_{11}\,/x_{1j} - \bar{x}_1/ + l_{21}\,/x_{2j} - \bar{x}_2/ + \ldots + l_{p1}\,/x_{pj} - \bar{x}_p/\Big]$$

/7.1.6/
$$\hat{f}_{2j} = \frac{1}{\lambda_2 - \lambda_0} \left[l_{12} /x_{1j} - \bar{x}_1/ + l_{22} /x_{2j} - \bar{x}_2/ + \ldots + l_{p2} /x_{pj} - \bar{x}_p/ \right]$$

$$\ldots$$

$$\hat{f}_{rj} = \frac{1}{\lambda_r - \lambda_0} \left[l_{1r} /x_{1j} - \bar{x}_1/ + l_{2r} /x_{2j} - \bar{x}_2/ + \ldots + l_{pr} /x_{pj} - \bar{x}_p/ \right]$$

Ponieważ, jak wiemy, ładunki są proporcjonalne do wektorów własnych macierzy kowariancji /lub korelacji/, więc wartości czynników są proporcjonalne do tzw. składowych głównych.

W tym przypadku estymowane wartości czynników są nieskorelowane. Estymacja wartości czynników z macierzy ładunków po obrocie wprowadza znowu korelacje między czynnikami.

7.2. Estymacja wartości czynników z warunkiem, by były one unormowane i nieskorelowane.

Można estymować wartości indywidualne czynników metodą minimum kwadratów żądając dodatkowo, by średnie kwadratowe odchylenia były równe 1, a korelacje równe zeru /Anderson, Rubin, 1956/. Wartości czynników wyznaczone są wzorami:

/7.2.1/
$$\hat{\underline{f}}_j = H\, L'\, V^{-1} /\underline{x}_j - \bar{\underline{x}}/$$

Macierz H zdefinowana jest następującymi warunkami:

a/ H jest macierzą symetryczną

b/ $H\, H = L'\, V^{-1}\, C\, V^{-1}\, L$

7.3. Estymacja przy pomocy regresji

Ideà takiego otrzymania wartości czynników pochodzi od Thomsona /1950/. Rozpatrzmy łączny rozkład cech $X_1, X_2, \ldots, X_p$ wraz z czynnikiami $F_1, F_2 \ldots, F_r$. Załóżmy, że rozkład ten jest /p+r/ wymiarowym rozkładem normalnym.

Macierz kowariancji tych zmiennych można zapisać jako macierz

$$\begin{bmatrix} C & L \\ L' & I \end{bmatrix},$$

gdzie C jest populacyjną macierzą kowariancji cech $X_1, \ldots X_p$, L macierzą ładunków, czyli macierzą kowariancji cech z czynnikami. Macierz kowariancji czynników jest macierzą jednostkową, bo z założenia czynniki są nieskorelowane i unormowane.

Równanie regresji dla k-tego czynnika wspólnego u j-ego osobnika jest postaci

$$f_{kj} = \beta'_k \ /\underline{x}_j - \bar{x}/$$

gdzie $\underline{x}_j$ jest wektorem wartości cech u tego osobnika, a $\bar{x}$ wektorem średnich arytmetycznych. Wektor współczynników regresji $\underline{\beta}_k$ otrzymuje się jako rozwiązanie równań kanonicznych, które są postaci

$$\beta_k = C^{-1} \ l_k$$

Oznaczmy przez $\underline{f}_j$ r-wymiarowy wektor wartości czynników wspólnych u j-tego osobnika, a przez B macierz, której kolumnami są wektory współczynników regresji $\underline{\beta}_k$.

Wtedy równania regresji dla wszystkich czynników można napisać w postaci zwartej następująco

$$\underline{f}_j = B' /\underline{x}_j - \bar{x}/$$

Macierz B określona jest wzorem

$$B = C^{-1} L$$

stąd

$$f_j = L' C^{-1} /x_j - \bar{x}/$$

Podstawienie na miejsce C macierzy $LL' + V$, odpowiednie przekształcenie i zastąpienie prawdziwych wartości ich estymatorami daje w końcu wynik

$$/7.3.1/ \qquad \hat{f}_j = /I_r + \hat{L}'\hat{V}^{-1}\hat{L}/^{-1} \hat{L}'\hat{V}^{-1} /x_j - \bar{x}/$$

Podobnie jak poprzednio, średnie wartości czynników są równe zeru. Macierz kowariancji tak obliczonych wartości czynników jest równa

$$/7.3.2/ \qquad M = /I + L'V^{-1} L^{-1}/ L' V^{-1} L .$$

7.4. Przykład

Dla zilustrowania estymacji wartości czynników wspólnych wykorzystano dane A.Skibińskiej /por. Skibińska, Szczotka,1969/.

W tablicy 7 podano wartości średnie i standardowe odchylenia 20 cech charakteryzujących budowę ciała. Pomiary przeprowadzono w grupie 117 mężczyzn. Analiza czynnikowa wyodrębniła 5 czynników wspólnych. Estymację przeprowadzono metodą maksimum wiarogodności na macierzy kowariancji cech. Otrzymane wartości ładunków poddano rotacji według zasady varimax. W tablicy 8 podano wartości unormowane ładunków $\hat{l}^{*}_{ik} = \hat{l}_{ik}/s_i$ po rotacji, względne zasoby zmienności wspólnej $\hat{h}^2_i$ i względne wariancje specyficzne $\hat{v}^2_i$. Wartości $\hat{l}^{*}_{ik}$ są więc empirycznymi współczynnikami korelacji czynników wspólnych z cechami. W ostatnim wierszu tablicy podano sumy kwadratów ładunków $\sum_{i=1}^{20} \hat{l}^{*2}_{ik}$ dla k=1,..., 5.

Sumy te stanowią miarę ważności poszczególnych czynników wspólnych.

Pierwszy czynnik koreluje wysoko z długością kończyny dolnej i długością przedramienia, dlatego nazwano go "długością". Drugi czynnik koreluje wysoko z grubością fałdy tłuszczowej na brzuchu i fałdy tłuszczowej pod łopatką, czynnik ten nazwano "otłuszczeniem tułowia". Trzeci czynnik koreluje wysoko z grubością tłuszczu na kończynach /cechy numer 15, 16, 18 i 20/, czynnik ten reprezentuje "otłuszczenie kończyn". Ostatnie dwa czynniki wspólne charakteryzują tężyznę budowy ciała. "Tężyzna I" koreluje wysoko z obwodami przedramienia i podudzia, grubością mięśni ramienia i uda, szerokością kości ramienia, szerokością łokcia oraz szerokością klatki piersiowej. Czynnik ten okazał się czynnikiem o największym znaczeniu. Czynnik "tężyzna II" koreluje wysoko tylko z szerokością kostki.

Wartości czynników wspólnych estymowano metodą minimum kwadratów z dodatkowym założeniem, by estymowane wartości czynników były w badanej próbce nieskorelowane i unormowane. Wartości te otrzymuje się ze wzoru /7.2.1/

$$\hat{\underline{f}}_j = H L' V^{-1} / \underline{x}_j - \bar{\underline{x}} / = M / \underline{x}_j - \bar{\underline{x}} /$$

gdzie L jest macierzą ładunków /nieunormowanych/, V macierzą przekątniową wariancji specyficznych /także nieunormowanych/ macierz H zaś macierzą symetryczną, spełniającą warunek

$$HH' = L' V^{-1} C V^{-1} L$$

Transponowana macierz $M=HL' V^{-1}$ podana jest w tablicy 9. Dla uproszczenia zapisu wszystkie wyrazy pomnożono przez 100. Kolumny tej macierzy stanowią współczynniki przy $x_{ij} - \bar{x}_i$ we wzorze na estymowaną wartość czynnika. Tak więc wartość pierwszego czynnika dla j-ej osoby oblicza się ze wzoru

$$\hat{f}_{1j} = 0,01 \left\{ 0,25/x_{1j}-\bar{x}_1/+ 1,28/x_{2j}-\bar{x}_2/ + 3,99/x_{3j}-\bar{x}_3/ + \dots + +0,23/x_{19,j}-\bar{x}_{19}/ + 0,02\, /x_{20,j}-\bar{x}_{20}/ \right\}.$$

W tablicy 10 wypisano estymowane wartości wszystkich 5 czynników wspólnych dla dziesięciu wybranych osób. Surowe wartości cech tych osób podane są w tablicy 11.

Pierwsze dwie osoby charakteryzują się wysoką wartością czynnika "długość", ale różnią się w pozostałych czynnikach o mniej więcej jedno standardowe odchylenie. Dwie dalsze osoby mają bardzo niskie wartości czynnika "długość". Osoba czwarta ma wyjątkowo małą wartość czynnika "otłuszczenie kończyn" /f_{34}=-3,19/.

Osoby 5 i 6 są przeciętne pod względem czynnika "długość", różnią się one znacznie w czynniku "tęgość I" /f_{45}= -1,79, f_{46} = = 1,21/, Osoba nr 6 ma bardzo niską wartość czynnika "otłuszczenie tułowia" /f_{26}=-2,10/. Przeciwieństwem do niej pod względem tego czynnika jest osoba nr 9 /f_{29}=2,75/. Osobami o skrajnych wartościach czynnika "tężyzna II" są osoby nr 7 i 8 /f_{47}=-2,45, f_{48}=2,28/. Osoba nr 10 zaś mą wysoką wartość czynnika "otłuszczenie kończyn" /$f_{3,10}$=2,05/.

Wartości czynników wspólnych odzwierciedlają wielkość wartości unormowanych tych cech, które najsilniej z danym czynnikiem korelują. Tak więc np. osoba nr 4 o niskich wartościach czynnika "długość" /f_{14}= -2,19/ i "otłuszczenie kończyn" /f_{34}=-3,19/ ma relatywnie małą długość kończyny dolnej /731 mm/ długość przedramienia /237 mm/ i wysokość siedzeniową /890 mm/ oraz małą grubość tłuszczu na ramieniu /148 mm/, tłuszczu na udzie /201 mm/, cienką fałdę tłuszczową na ramieniu /136 mm/ i małą grubość tłuszczową na ramieniu /136 mm/ i małą grubość tłuszczu na krętarzu /134 mm/. W podobny sposób można porównać wartości czynników z względnymi wartościami tych cech, które z nimi korelują.

Tablica 7. Wartości średnie $/\bar{x}_i/$ i standardowe odchylenia $/s_i/$ 20 cech budowy ciała /w mm/

	$\bar{x}_i$	s_i
1. Wysokość siedzeniowa	924,8	32,42
2. Długość kończyny dolnej	823,0	43,33
3. Długość przedramienia	253,3	11,67
4. Szerokość pleców	402,6	16,26
5. Szerokość klatki piersiowej	291,8	14,40
6. Szerokość miednicy	281,4	15,18
7. Szerokość łokcia	69,6	3,19
8. Szerokość kostki	76,5	3,46
9. Szerokość kości ramiennej /rtg/	22,2	1,41
10. Szerokość kości uda /rtg/	30,4	2,08
11. Obwód przedramienia	268,8	14,00
12. Obwód podudzia	361,9	18.54
13. Grubość mięśni ramienia /rtg/	607,8	52,22
14. Grubość mięśni uda /rtg/	1334,8	82,46
15. Tłuszcz na ramieniu /rtg/	243,1	42,34
16. Tłuszcz na udzie /rtg/	303,9	31,71
17. Grubość fałdy tłuszczowej na brzuchu	169,1	12,24
18. Grubość fałdy tłuszczowej na ramieniu	161,2	12,95
19. Grubość fałdy tłuszczowej pod łopatką	177,2	8,98
20. Tłuszcz na krętarzu /rtg/	178,1	15,69

Tablica 8. Ładunki pięciu czynników budowy ciała.

	$\hat{l}^{*}_{i1}$	$\hat{l}^{*}_{i2}$	$\hat{l}^{*}_{i3}$	$\hat{l}^{*}_{i4}$	$\hat{l}^{*}_{i5}$	$\frac{\hat{h}_i^2}{s_i^2}$	$\frac{\hat{v}_i^2}{s_i^2}$
1. Wysokość siedzeniowa	0.490	0.092	-0.144	0.334	0.411	0.549	0.451
2. Długość kończyny dolnej	0.884	-0.001	-0.006	0.049	0.158	0.808	0.192
3. Długość przedramienia	0.856	-0.039	-0.005	0.149	0.106	0.768	0.232
4. Szerokość pleców	0.492	0.111	-0.040	0.442	0.148	0.473	0.527
5. Szerokość klatki piersiowej	0.133	-0.091	0.075	0.593	0.122	0.398	0.602
6. Szerokość miednicy	0.330	-0.026	0.049	0.241	0.363	0.302	0.698
7. Szerokość łokcia	0.274	-0.110	0.010	0.510	0.375	0.488	0.512
8. Szerokość kostki	0.237	-0.028	0.044	0.413	0.775	0.830	0.170
9. Szerokość kości ramiennej /rtg/	0.230	0.013	-0.155	0.505	0.205	0.375	0.625
10. Szerokość kości uda /rtg/	0.310	0.123	-0.197	0.322	0.250	0.317	0.683
11. Obwód przedramienia	0.028	0.023	0.012	0.954	0.056	0.915	0.085
12. Obwód podudzia	0.079	0.019	0.115	0.740	0.217	0.614	0.386
13. Grubość mięśni ramienia /rtg/	-0.108	0.028	-0.131	0.615	-0.144	0.428	0.572
14. Grubość mięśni uda /rtg/	-0.034	0.073	0.067	0.571	0.025	0.338	0.550
15. Tłuszcz na ramieniu /rtg/	-0.019	0.339	0.590	0.124	-0.007	0.479	0.521
16. Tłuszcz na udzie /rtg/	-0.019	0.121	0.799	0.181	0.084	0.694	0.306
17. Grubość fałdy tłuszcz. na brzuchu	-0.040	-0.680	0.321	0.162	0.100	0.603	0.397
18. Grubość fałdy tłuszcz. na ramieniu	0.021	0.298	0.524	0.103	-0.014	0.375	0.625
19. Grubość fałdy tłuszcz. pod łopatką	-0.043	0.778	0.150	0.276	-0.109	0.718	0.282
20. Tłuszcz na krętarzu /rtg/	0.012	0.275	0.721	-0.002	0.127	0.612	0.388
$\sum_{i=1}^{20} \hat{l}^{*2}_{ik}$	2.428	2.261	2.037	3.854	1.336		

Tablica 9. Transponowana macierz M współczynników we wzorze na wartości czynników wspólnych. /Wszystkie współczynniki pomnożone zostały przez 100/

Nr cechy					
1.	0,25	0,12	0,07	-0,02	0,37
2.	1,28	0,07	0,04	-0,12	-0,37
3.	3,99	-0,43	0,41	0,05	-0,85
4.	0,66	0,38	-0,27	0,21	-0,17
5.	0,15	-0,56	0,32	0,44	-0,18
6.	0,16	-0,05	0,03	-0,02	0,54
7.	0,59	-1,95	0,24	1,15	3,04
8.	-4,81	1,70	-3,22	-1,85	2,92
9.	0,17	0,22	-0,51	0,27	0,23
10.	0,14	0,46	-0,50	0,03	0,30
11.	0,43	-1,56	0,02	5,47	-3,00
12.	-0,01	-0,28	0,25	0,58	0,15
13.	0,00	0,00	-0,10	0,17	-0,18
14.	0,00	0,00	0,01	0,07	-0,04
15.	0,03	0,10	0,45	0,01	-0,09
16.	0,03	-0,71	1,79	0,08	-0,18
17.	-0,18	3,38	-0,34	-0,23	0,79
18.	0,14	0,22	1,10	0,01	-0,28
19.	0,23	8,03	-2,38	0,18	-0,48
20.	0,02	-0,09	2,18	-0,20	0,20

Tablica 10. Wartości czynników wspólnych dla 10 wybranych osób

Osoby	Długość $\hat{f}_{1j}$	Otłuszcz. tułowia $\hat{f}_{2j}$	Otłuszcz. kończyn $\hat{f}_{3j}$	Tęgość I $\hat{f}_{4j}$	Tęgość II $\hat{f}_{5j}$
1	2.70	0,47	-0,14	0,73	0,17
2	2.15	-0.83	-2.01	-0.25	-1.06
3	-2.60	0.10	-0,72	1.01	-1.61
4	-2.19	-0.98	-3.19	-0.15	-0,50
5	0.01	0.76	-0.20	-1.79	0.90
6	-0.11	-2.10	0.63	1.21	0.57
7	-0.90	0.00	-1.11	-2.45	-0.17
8	0.25	-0.24	1.74	2.28	0.10
9	0.81	2.75	-0.65	1.72	0.42
10	0.48	-0.63	2.05	-0.66	-0.35

Tablica 11. Wartości cech budowy ciała dla 10 wybranych osób

Nr cechy	Osoby: 1	2	3	4	5	6	7	8	9	10
1	1001	929	904	890	929	936	896	958	989	942
2	941	908	693	731	828	807	816	835	856	835
3	275	276	223	237	254	250	240	248	254	260
4	437	411	397	365	404	422	382	423	432	407
5	303	292	285	285	280	328	259	313	297	282
6	293	294	278	262	284	303	285	281	281	294
7	71	72	66	69	67	72	64	71	71	68
8	80	75	70	72	76	79	72	80	80	75
9	226	219	225	207	233	245	198	226	237	202
10	310	310	300	295	314	303	309	324	328	285
11	280	262	278	262	242	282	233	302	295	260
12	373	355	378	373	348	388	315	406	379	357
13	676	634	655	663	542	579	578	668	677	567
14	1351	1388	1441	1381	1286	1469	1366	1397	1441	1399
15	193	201	242	148	209	240	185	294	273	279
16	273	245	280	201	291	324	277	359	305	353
17	165	130	153	160	172	153	156	181	199	170
18	145	140	178	136	157	167	151	186	159	167
19	183	178	185	164	177	164	171	179	200	172
20	178	143	162	134	194	172	157	200	186	203

8. Porównanie analiz czynnikowych

8.1. Analiza czynników w podpopulacji.

Weźmy pod uwagę populację, w której każdy osobnik scharakteryzowany jest p cechami, np. populację chłopców 16-letnich, których poddano pewnej ilości prób sprawności fizycznej. Oznaczmy średnie wartości cech przez $\mu_1, \mu_2, \ldots \mu_p$, a macierz współczynników korelacji przez $R = \left[r_{ij}\right]$. Przypuśćmy, że ograniczamy nasze badania do części populacji. Jeżeli taką częścią populacji jest próbka wybrana losowo, to jej średnie, kowariancje i korelacje z dużym prawdopodobieństwem będą się różniły niewiele od tychże wielkości w populacji. Inaczej może być wtedy, gdy podpopulacja będzie wybrana według jakiejś właściwości związanej z badanym zespołem cech. Taką podpopulację będziemy nazywać podpopulacją wyselekcjonowaną. Wyselekcjonowaną podpopulacją będą np. chłopcy 16-letni uprawiający lekką atletykę. Parametry statystyczne w podpopulacji mogą zasadniczo różnić się od parametrów całej populacji. Mogą wystąpić różnice nie tylko w wartościach średnich badanych cech, ale także w dyspersjach, kowariancjach i korelacjach. Interpretacja tych różnic, a więc szczegółowe wyjaśnienie mechanizmu selekcji nastręcza w takich sytuacjach wiele trudności. Często pomocne tu być może badanie właściwości czynników wyodrębnionych w całej populacji.

Opiszmy populację zgodnie z ogólnym modelem przy pomocy r nieskorelowanych czynników wspólnych. Przypuśćmy, że z populacji wybrano podpopulację. Zespół rozpatrywanych cech w podpopulacji oznaczmy przez $Y_1, Y_2, \ldots, Y_p$, ich średnie przez $\eta_1, \eta_2, \ldots, \eta_p$ a macierz kowariancji przez C^*.

Ponieważ każdy osobnik podpopulacji jest równocześnie osobnikiem populacji ogólnej i scharakteryzowany tymi samymi cechami, możemy napisać, że

$$Y_1 = l_{11} F_1 + l_{12} F_2 + \ldots + l_{1r} F_r + v_1 U_1$$

$$Y_2 = l_{21} F_1 + l_{22} F_2 + \ldots + l_{2r} F_r + v_2 U_2$$

....

$$Y_p = l_{p1}F_1 + l_{p2}F_2 + \ldots + l_{pr}F_r + v_pU_p$$

lub w postaci macierzowej $Y = LF + VU$

Dalsze założenia o unormowaniu i nieskorelowaniu czynników odnoszą się do całej populacji i mogą być niespełnione w podpopulacji. Załóżmy jednak, że ewentualna selekcja nie dotyczyła czynników specyficznych i że są one w podpopulacji nadal unormowane i nieskorelowane z sobą i z czynnikami wspólnymi.

Oznaczmy przez $\varphi_1, \varphi_2, \ldots, \varphi_r$ średnie wartości czynników wspólnych w podpopulacji, a przez $/\varphi_{rx1}/$ wektor kolumnowy tych średnich. Zachodzi wtedy związek.

$$E^{*}\,\underline{Y} = L\varphi,$$

gdzie symbol E^{*} oznacza wartość oczekiwaną w podpopulacji. Jeżeli przez η oznaczyć **wektor średnich wartości cech** $Y_1, \ldots, Y_p$, to otrzymujemy równość

$$\eta = L\,\varphi,$$

a stąd można obliczyć średnie wartości czynników wspólnych w podpopulacji:

/8.1.1/

$$\varphi = /L'L/^{-1}\,L'\,\eta$$

Oznaczmy przez M macierz kowariancji czynników w podpopulacji. Wtedy między macierzą kowariancji C^{*} cech w podpopulacji a macierzą ładunków, macierzą wariancji specyficznych i macierzą kowariancji czynników w podpopulacji zachodzi następujący związek:

$$C^{*} = L\,M\,L' + V$$

Stąd przez przekształcenia algebraiczne otrzymujemy macierz kowariancji czynników w podpopulacji:

/8.1.2/ $$M = /L'L/^{-1}\,L'/C^{*} - V/\,L\,/L'L/^{-1}$$

Z powyższych wzorów wynika, że jeżeli znamy macierz ładunków i macierz czynników specyficznych w całej populacji, to można obliczyć średnie wartości, wariancje i kowariancje czynników w podpopulacji.

Przytoczone wzory dają możność obliczenia i zinterpretowania efektów selekcji w kategoriach czynników.

8.2. Porównanie analiz czynnikowych dwóch lub więcej populacji.

Przypuśćmy, że w próbkach pochodzących z dwu różnych populacji pomierzono te same cechy. Obie próbki opracowano tymi samymi metodami analizy czynnikowej. Powstaje pytanie w jakim stopniu wyniki obu analiz są porównywalne.

Przystępując do analizy czynnikowej robi się z powodu braku innej informacji założenia, które są jednak konieczne dla rozwiązania zadania: zakładamy więc, że w obu populacjach średnie czynników są równe zeru, ich standardowe odchylenia równe jeden oraz, że czynniki są nieskorelowane. Estymowaną macierz ładunków poddaje się rotacji według zasady odnoszącej się do samej macierzy ładunków. Widzimy więc, że wykonując niezależnie analizy czynnikowe, a priori zrównujemy w pewien sposób wyniki w jednej i drugiej populacji. Być może, że eliminujemy w ten sposób właśnie te różnice, które mogłyby być ciekawe. Porównywanie wyników analiz czynnikowych dwu lub więcej grup nastręcza więc zasadnicze trudności.

Wybrana literatura

Anderson T.W., 1958, Introduction to Multivariate Statistycal Analysis, Wiley and Sons, London, New-York.

Anderson T.W., Rubin H., 1956, Statistical Inference in Factor Analysis, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 5, p. 11-50.

Bartlett M.S., 1937, The statistical conception of mental factors, British Journal of Psychology, 28, p. 97-104.

Bartlett M.S., 1950, Tests of significance in factor analysis, British Journal of Psychology, Stat. Sec.,3, p. 77-85.

Bartlett M.S.,1953,Factor analysis in psychology as a statistician sees it, Uppsala Symposium on Psychological Factor Analysis, Almquist and Wiksell, Uppsala.

Harman H.H., 1960, Modern Factor Analysis, The University of Chicago Press, Chicago.

Hotelling H., 1933, Analysis of a complex of statistical variables into principal components, Journal of Educational Psychology, 24, p. 417-441 i p. 498-520.

Jahn W., Vahle H., 1970, Die Factorenalyse, Verlag Die Wirtschaft, Berlin.

Jöreskog K.G., 1963, Statistical Estimation in Factor Analysis, Almqvist-Wiksell, Stockholm.

Kieloch A., Oktaba W., 1971, Factor Analysis by the Method of Maximum Likelihood, Zastosowania Matematyki /Applicationes Mathematicae/, XII, 1, str. 63-78.

Lawley D.N., Maxwell A.E., 1963, Factor Analysis as a Statistical Method, Butterworths, London.

Morrison, D.F., 1967, Multivariate Statistical Methods, McGraw-Hill Book Company, New York.

Skibińska A., Szczotka F.A., 1969, Budowa ciała w świetle analizy czynnikowej, Wychowanie Fizyczne i Sport, XII, 4, p.17-24.

Spearman C., 1904 General Intelligence Objectivily Determined and Measured, American Journal of Psychology, 15, p. 201-293.

Thomson G.H., 1950, L Analyse factorielle des aptitudes humaines, Presses Universitaires de France, Paris.

Thurstone L.L., 1947, Multiple Factor Analysis, University of Chicago Press, Chicago.

Szczegółowy spis treści