

120

**UTJECAJ KRITERIJA ZA ZAUSTAVLJANJE
EKSTRAKCIJE FAKTORA NA STRUKTURU
DOBIVENU OBLIMIN I VARIMAX
TRANSFORMACIJOM ZNAČAJNIH GLAVNIH
OSOVINA**

Konstantin Momirović, Vojislav Kovačević, Ignjat
Ignjatović, Smiljka Rađenović-Horga, Dobrivoje
Radovanović, Milko Mešovšek, Janez Štalec, Nata-
ša Viskiđ-Štalec.

Odjel za kineziološkijsku psihologiju i sociologiju

Influence of the Criterion for Number of Factors Extracted, on the Structure Achieved by Oblimin and Varimax Transformations of Significant Principal Axes.

If only those principal axes of image matrixes of covariances extracting at least $1/2 \cdot 10^{-2}$ of the relative variance of those matrixes are considered as being significant, there can be achieved relatively small, but consistent number of factors, which are, in most cases, sufficient for the expliation of image covariances.

It is probable that criteria based on the components of image covariances matrix's entropy are worth while further investigation.

Влияние критерия прекращения экстракции факторов на структуру, полученную при помощи облимин и варимакс трансформации значимых главных осей

Если значимыми считать только те главные оси изображения матриц ковариации, выделяющие, по крайней мере, $1/2 \cdot 10^{-2}$ относительной дисперсии этих матриц, получается относительно небольшое, но consistente число факторов, достаточных чаще всего для объяснения изображения ковариации.

Вероятно, что критерии, основанные на компонентах энтропии матрицы изображения ковариации, достойны дальнейших исследований.

0. UVOD

Određivanje broja značajnih glavnih osovina image matrice kovarijanci vjerojatno je najveći problem ove metode faktorske analize. Kako nema egzaktnog statističkog kriterija značajnosti karakterističnih korjenova image matrice kovarijanci, u praksi se primjenjuju empirijski kriteriji, od kojih su najčešći:

- (1) broj značajnih glavnih osovina, k , ravan je $n/2$, gdje je n red image matrice (Guttman)
- (2) broj značajnih glavnih osovina treba biti određen tako, da zbroj k karakterističnih korjenova odgovara nekoj frakciji p traga image matrice. Horst preporučuje, da p bude između 0.80 i 0.85.

Oba se predložena kriterija loše ponašaju u praksi. Prvi dovodi, očito, do produkcije suvišnih faktora, koji se tek gdjekada mogu ukloniti u toku transformacijskih procesa. Drugi, u nekim slučajevima, uistinu daje k koji je aproksimativno ravan statističkom rangju image matrice kovarijanci, ali u mnogim slučajevima proizvodi suviše malo faktora, a, što je još gore, na image matricama kovarijanci dobivenim na uzorcima iz iste populacije entiteta, može dati vrlo različit k .

Otuda su komparativne analize faktorskih struktura dobivenih s pomoću transformacija značajnih glavnih osovina image matrice kovarijanci vrlo otežane, budući je invarijantnost broja faktora osnovni preduvjet invarijantnosti faktorskih solucija. Vjerojatno se i uslijed toga image analiza relativno rijetko primjenjuje, unatoč svojih očitih prednosti nad drugim ekstrakcionim tehnikama. Naime u image analizi:

- (1) komunaliteti su točno određeni
- (2) uticaj unikne varijance je uklonjen, te je nemoguće dobiti artifičijelne single i dual faktore, što je jedan od bitnih nedostataka metode glavnih komponenata
- (3) varijable su normalizirane, jer su image varijable linearna kombinacija $n-1$ ostalih varijabli, a sukladno Laplaceovoj teoremi, linearne kombinacije ma kako distribuiranih varijabli, su normalne varijable ako je n dovoljno velik.

1. CILJ ISTRAŽIVANJA

Neka je

$$Z = (z_{ij})$$

$$\begin{aligned} i &= 1, \dots, n \\ i &= 1, \dots, N \end{aligned}$$

matrica standardiziranih vrijednosti N entiteta u n varijabli.

(iz istraživanja »Efikasnost krivičnih sankcija prema maloljetnim počiniocima krivičnih djela s posebnim osvrtom na povratništvo kod maloljetnika«).

Neka je

$$z'_i = (z_{i1} \dots z_{iN})$$

$$i = 1, \dots, n$$

vektor rezultata u bilo kojoj varijabli, i neka je

$$Z_{)i(} = (z_{kj})$$

$$k = 1, \dots, n-1$$

$$k \neq i$$

matrica rezultata u svim varijablama izuzev varijable i .

Operacijama

$$Z_{)i(} Z'_{)i(} N^{-1} = R_{)i(}$$

$$i = 1, \dots, n$$

dobit će se n korelacijskih matrica reda $n-1$, iz kojih je, sukcesivno, isključena varijabla i , $i = 1, \dots, n$.

Operacijom

$$\hat{Z}_i = r'_i R_{)i(}^{-1} Z_{)i(}$$

$$i = 1, \dots, n$$

gdje je

$$r_i = Z_{)i(} z_i N^{-1}$$

vektor korelacija varijable i sa ostalih $n-1$ varijabli k , dobiti će se, klasičnom regresionom analizom, n linalnih kombinacija $n-1$ varijabli. Varijable z_i nazivaju se image varijablama.

Varijance image varijabli su, očito,

$$D_i^2 = r'_i R_{)i(}^{-1} r_i$$

tj. ravni su koeficijentima determinacije svake image varijable.

Kovarijance image varijabli, dobivene operacijom

$$\hat{Z} \hat{Z} \cdot N^{-1} = C$$

gdje je

$$\hat{Z} = (z_{ij})$$

matrica image rezultata N entiteta u n image varijabli, očišćene su od bilo kojih utjecaja, osim utjecaja zajedničkih faktora; naime, image uklanja činioce, koji mogu dovesti do korelacijskih artefakta.

kata, kao što su abnormalnost distribucija, nedostatak homoscedaciteta ili utjecaj grešaka uzroka. Rješavanjem sistema karakterističnih jednažbi

$$(C - \lambda_i I) x_i = 0$$

$$i = 1, \dots, n$$

dobit će se dijagonalna matrica karakterističnih korjenova

$$\lambda = (\lambda_i)$$

i matrica ortonormalnih karakterističnih vektora

$$X = (x_i)$$

U praksi, matrica image kovarijanci računa se operacijom

$$C = R + S R^{-1} S - 2 S$$

gdje je

$$S = I - D^2$$

a

$$D^2 = (D_i^2)$$

dijagonalna matrica varijanci image varijabli.

Dijagonalizacija matrice C obično se vrši Hausholderovom metodom.

Cilj ovog istraživanja bio je da se odredi broj karakterističnih korjenova i karakterističnih vektora k, koji je dovoljan za reprodukciju vandijagonalnih elemenata matrice C, tj. da se matrice X i λ reduciraju na matrice \hat{X} i $\hat{\lambda}$ ranga k tako, da vrijedi

$$C \simeq H H' + E$$

gdje je E dijagonalna matrica neobjašnjeg dijela varijanci image varijabli, a

$$H = \hat{X} \hat{\lambda}^{1/2}$$

matrica značajnih glavnih osovina matrice C.

U tu svrhu ispitano je ponašanje slijedećih kriterija:

$$K 1: \sum_{i=1}^k \lambda_i n^{-1} \geq 0.80$$

$$K 2: \sum_{i=1}^k \lambda_i n^{-1} \geq 0.90$$

$$K 3: \sum_{i=1}^{k+1} \lambda_i n^{-1} - \sum_{i=1}^k \lambda_i n^{-1} \leq e 10^{-2}$$

$$K 4: \sum_{i=1}^{k+1} \lambda_i n^{-1} - \sum_{i=1}^k \lambda_i n^{-1} \leq$$

$$\leq 1/2 \pi e 10^{-2}$$

$$K 5: k = K, \quad K = \text{apriorna konstanta}$$

$$K 6: \lambda_k \geq 1.00$$

Kriterij 1 predložio je Horst; kriterij 2 dobio je nakon jedne serije pokusnih faktorskih analiza; kriterij 3 ispitano je radi toga, što je e konstanta funkcije kumulativne distribucije doprinosa karakterističnih korjenova objašnjenju ukupnog varijabiliteta image varijabli; kriterij 4 izveden je na temelju entropije matrice kovarijanci. Naime, entropija matrice C je

$$H_C = \sum_{i=1}^k \lambda_i + \frac{1}{2} k \pi e$$

a $1/2 \pi e$ je mehanički doprinos svake dimenzije; kriterij je formiran tako, da doprinos glavne osovine bude veći od mehaničkog doprinosa te osovine. Kriterij 5 uveden je zbog usporedbe transformiranih solucija. Kriterij 6, analogan kriteriju za značajne glavne komponente, ima iste racionalne kao i kriterij 4.

Dalji cilj ovog istraživanja bio je da se usporede varimax i obliimin transformacije značajnih glavnih osovina dobivenih prema nekim od predloženih kriterija, posebno K1, K2 i K5. Varimax transformacije učinjene su tako, da se maksimizira

$$V = n \sum_{p=1}^k \sum_{i=1}^n (a_{ip} h_i^{-1})^4 -$$

$$- \sum_{p=1}^k \left(\sum_{i=1}^n a_{ip}^2 h_i^{-2} \right)^2$$

gdje su a_{ip} elementi matrice A dobivene transformacijom

$$H T = A$$

pod uvjetima

$$T' T = T T' = I$$

Oblimin transformacija učinjena je direktnom oblimin metodom, tako da se minimizira

$$S = \sum_{q < p}^k \sum_{i=1}^n b_{ip}^2 b_{iq}^2$$

gdje su b_{ip} elementi matrice B dobivene transformacijom

$$H \tau^{-1} = B$$

pod uvjetima

$$\tau \tau' = M$$

$$\text{dg } M = I$$

2. METODE RADA

Na uzorku iz populacije jugoslavenskih maloljetnih delinkvenata, muškog spola, starih između 15 i 21 godinu, kojima je za dokazano krivično djelo protiv imovine ili protiv života i tijela izrečena bilo koja sankcija predviđena krivičnim zakonom (osim sankcije upućivanja u specijalnu ustanovu i upućivanja u drugu porodicu, koje se praktički ne izriču) primijenjeno je 68 kognitivnih i konativnih testova. Ukupno je bilo ispitano 1342 ispitanika. Ti su ispitanici podijeljeni na devet skupina, prema tome da li su sankcije, koje su im izrečene, izdržali, ili su tek počeli da izdržavaju, i koja im je sankcija izrečena. Obustava postupka sa razloga oportuniteta, iako je krivično djelo dokazano, tretirana je također kao sankcija (kao nulta sankcija). Ukupno je, sa cijelim uzrokom, bilo 10 grupa, od kojih su se neke prepokrivale.

U svakoj grupi izračunane su matrice image kovarijanci, i određeni karakteristični korjenovi i vektori tih matrica. Broj značajnih glavnih osovin (R) za svaku od tih matrica određen je prema kriterijima K1, K2, K3, K4 i K6. Na temelju rezultata u cijelom uzorku fiksiran je k prema kriterijima K1 i K2, i značajne glavne osovine transformirane sukladno varimax i oblimin kriteriju. Interpretirani su dobiveni faktori, i određena sukladnost dobivenih solucija u različitim grupama ispitanika.

3. REZULTATI I DISKUSIJA

U svim grupama, osim grupe 06 u kojoj je broj ispitanika bio nedovoljan za određivanje image varijabli, određen je broj značajnih glavnih osovin sukladno kriterijima K1, K2, K3, K4 i K6. Rezultati su prikazani u tabeli 1.

Tabela 1

BROJ ZNAČAJNIH GLAVNIH OSOVINA IMAGE MATRICA KOVARIJANCI

Grupa	N	K1	K2	K3	K4	K6
01	1342	4	9	5	4	5
02	624	5	11	5	4	5
03	595	5	12	5	4	5
04	123	13	23	6	4	10
05	167	11	21	6	3	7
06	33	—	—	—	—	—
07	161	11	20	6	4	8
08	338	7	16	5	4	6
09	117	15	25	7	4	10
10	214	11	21	6	4	7
11	189	10	20	6	4	6

Kako se vidi, kriteriji koji se temelje na postotku objašnjene varijance daju sasvim inkonzistentne rezultate. Broj faktora, prema K1, varira od 4 — 15, a prema K2 od 9 — 25. Na prvi se pogled čini, da je broj faktora izoliranih po K1 i K2 u nekoj vezi s brojem ispitanika po grupama, što je sukladno rezultatima nekih istraživanja koja su pokazala da je kontaminacija faktora error varijancom utoliko veća, ukoliko je broj ispitanika manji. Međutim, te se grupe međusobno značajno razlikuju i po razini kognitivnih i konativnih dimenzija a i po varijabilitetu i, osobito, kovarijabilitetu tih dimenzija. Manje grupe su u pravilu homogenije, pa je zbog kontrakcije kovarijanci i kovarijabilitet testova snižen. Otuda su i razlike karakterističnih korjenova manje, pa je veći broj korjenova potreban da bi se objasnio neki, unaprijed određeni dio tražene image matrice kovarijanci.

Isti je razlog vjerojatno i u osnovi inkonzistentnog ponašanja K6, kriterija osnovanog na veličini karakterističnih korjenova. Što je, naime, razlika između prvih karakterističnih korjenova manja, veća je vjerojatnost da će veći broj tih korjenova biti veći od 1,00; a korjenovi se jedne matrice utoliko manje razlikuju, što je kovarijabilitet varijabli manji. K6 daje od 5 — 10 faktora, pa je nešto konzistentniji od K1 i K2 ali je očito da kao ni K1 i K2 ne može služiti kao valjani osnov za određivanje statističkog ranga image matrica kovarijanci.

K3 se ponaša prilično konzistentno, producirajući 5 — 7 faktora (7 samo u jednom slučaju). Taj se kriterij očito osniva na nagibu kumulativne distribucije kovarijanci, objašnjenih svakim karakterističnim korjenom, pa vjerojatno ne može biti sasvim neosjetljiv na veličinu generalizirane varijance matrice korelacija među faktorima (koja je utoliko veća, što je kovarijabilitet faktora manji). No osjetljivost toga kriterija na promjene kovarijabiliteta testova ili faktora je očito dosta niska, pa se u svim slučajevima, kada te razlike nisu visoke, K3

može smatrati relativno pouzdanim kriterijem za određivanje broja faktora. Uostalom, K3 ima iste racionalne kao i poznati scree test, koji je također, kako je poznato relativno dobra mjera za određivanje broja faktora iako je dakako njegova primjena podložna brojnim subjektivnim greškama. Inače K3 je u znatnom kovarijabilitetu sa K6, a u izvjesnom kovarijabilitetu i sa K1 i K2, ali u pravilu producira manje faktora od K1 i K6, a naravno i od K2. Malena je vjerojatnost da će, po K3, u analizu biti uključen i neki error faktor, ali nije pouzdano da se tim kriterijem uvijek dobiva dovoljan broj značajnih faktora.

K4 se ponaša gotovo apsolutno konzistentno producirajući uvijek 4 faktora. Samo u jednom slučaju po tom kriteriju je esktrahirano 3 faktora, ali i tada je to učinjeno samo zbog razlike na četvrtoj decimali. Na temelju takve konzistentnosti a i na temelju logike njegove konstrukcije (predložio ga je J. Štalc) čini se, da je K4 najpouzdanija mjera broja značajnih dimenzija u image matricama. Ipak treba upozoriti, da K4 producira manje faktora od svih ostalih kriterija, pa je moguće, da se po tom kriteriju neće izolirati neki faktori koji nisu kontaminirani error varijancom.

Posebno treba upozoriti na to, da se u grupi 01 (cijeli uzorak) K1 i K4 ponašaju identično i produciraju samo jedan faktor manje od K3 i K6. Osim toga, matrica interkorelacija faktora dobivena po K2 faktorizirana je da bi se dobili faktori drugog reda; tim je postupkom izolirano četiri faktora drugog reda, koji su ne samo bili dovoljni za eksplikaciju kovarijabiliteta faktora, već je četiri faktora bilo dovoljno, na temelju veličine reziduala u rezidualnim matricama u svim grupama, za eksplikaciju kovarijabiliteta image varijabli. Na temelju aproksimativne formule za određivanje standardne devijacije parcijalnih koeficijenata korelacije, ni u jednoj rezidualnoj matrici nije broj značajnih reziduala bio veći od 5%, unaprijed određene kriterija za dopušteni broj značajnih reziduala.

Naime, ako je σ_p , standardna devijacija reziduala, aproksimativno

$$\sigma_p = [N - k - (k^2 - k)/2 - 2]^{-1/2}$$

broj reziduala većih od $t_{05} \sigma_p$ ne smije biti veći od 5% ako je rezidualna matrica statistički jednaka nula matrici.

Varimax i oblimin transformacije značajnih glavnih osovina dobivenih po K1 i K2 dale su dođuše interpretabilne faktore u grupi 01 ali potpuno neinterpretabilne faktore u ostalim grupama. U tim grupama najveći broj faktora bio je u stvari single ili rjeđe dual.

Kada je broj faktora bio fiksiran na četiri i na devet, sukladno rezultatima dobivenim u grupi 01, ti su se faktori mogli mnogo lakše interpretirati. Kod toga su jednostavne strukture dobivene oblimin transformacijama bile mnogo pregnantnije, od jednostavnih struktura dobivenih varimax tranfor-

macijama. Varimax je tendirao produkciji faktora širokog opsega, slabo definiranih grupnih faktora, a osim toga faktori nisu nipošto bili invarijantni što se obično smatra glavnom prednošću varimax solucije.

Zbog toga je sukladno K4 izvršena varimax i oblimin transformacija značajnih glavnih osovina u svim grupama. Rezultati su sintetizirani u Tabeli 2.

Tabela 2

Faktor	Varimax	Oblimin
G	10	10
Y	10	10
E	8	10
η	7	9
R	4	—

Kako se vidi, oblimin transformacija pod ovim kriterijem daje znatno konzistentnije solucije od transformacije sukladno varimax kriteriju. Faktori su interpretirani na slijedeći način:

G = generalni kognitivni faktor

Y = generalni neurotizam

E = ekstraverzija

η = integriranost ličnosti

R = mentalna rigidnost

Očito je, da pod K4 oblimin transformacija značajnih glavnih osovina image matrica kovarijanci, daje dobro definirane faktore širokog opsega.

U nekoliko slučajeva matrica interkorelacija faktora dobivenih po kriterijima koji su izolirali veći broj faktora bila je analizirana kako bi se dobili faktori drugog reda. Redovito je četiri faktora drugog reda bilo interpretirano kao generalni kognitivni faktor, generalni neurotizam, ekstraverzija i integriranost ličnosti. Čini se, dakle, da K4 gotovo sigurno producira ako ne sve značajne primarne faktore, ono bar faktore koji su virtualno faktori drugog reda.

4. ZAKLJUČAK

Ako se značajnim smatraju samo one glavne osovine image matrica kovarijanci koje iscrpljuju najmanje $1/2 \pi$ e 10^{-2} relativne varijance tih matrica, dobiva se relativno malen, ali konzistentan broj faktora, koji su najčešće dovoljni za eksplikaciju image kovarijanci.

Vjerojatno je, da kriteriji koji se osnivaju na komponentama entropije matrice image kovarijanci zaslužuju dalja istraživanja.

5. LITERATURA

1. Browne, M. W.
A Note on Lower Bounds for the Number of Common Factors. Psychometrika, 1968, Vol. 33, No. 2, 233—236.

2. Harman, H. H.
Modern Factor Analysis. The University of Chicago Press, Chicago and London, 1967.
3. Horn, J. L.
A Rationale and Test for the Number of Factors in Factor Analysis. *Psychometrika*, 1965, Vol. 30, No. 2, 179—185.
4. Horst, P.
Factor Analysis of Data Matrices. Holt, Rinehart and Winston, New York, 1965.
5. Jöreskog, K. G.
Testing a Simple Structure Hypothesis in Factor Analysis. *Psychometrika*, 1966, Vol. 31, No. 2, 165—178.
6. Kaiser, H. F.
The Varimax Criterion for Analytic Rotation in Factor Analysis. *Psychometrika*, 1958, Vol. 23, No. 2, 187—200.
7. Veldman, D. J.
Fortran Programming for the Behavioral Sciences. Holt, Rinehart and Winston, New York, 1967.

