



Raziskovalna metodologija v družboslovju

Kvantitativne metode –faktorska analiza

Borut Kodrič

Faktorska analiza

- Model, s katerim skušamo pojasniti povezave med večjim številom spremenljivk, z manjšim številom faktorjev, ki predstavljajo skupne razsežnosti (Mardia, Kent, Bibby, 2003);
- Študij povezav med spremenljivkami s katerim poizkušamo najti novo množico spremenljivk (manj kot merjenih spremenljivk), ki predstavljajo kar je skupnega opazovanim spremenljivkam (Ferligoj, 2001);



Splošni faktorški model

Predpostavke splošnega faktorkega modela - 1

- Skupnih faktorjev ne moremo neposredno meriti;
- Na merjene spremenljivke vplivajo poleg skupnih faktorjev še specifični dejavniki;
- Zvezo med spremenljivkami lahko zapišemo kot:

$$X_i = \sum_{r=1}^k a_{ir} F_r + E_i; i = 1, \dots, m; k < m$$

Predpostavke splošnega faktorkega modela - 2

- Neodvisnost specifičnih faktorjev:
 $Cov(E_i, E_j) = 0; i \neq j$
- Neodvisnost med specifičnimi in skupnimi faktorji:
 $Cov(E_i, F_j) = 0; \text{ za vsak } i \text{ in } j$
- Neodvisnost med skupnimi faktorji
 $Cov(F_i, F_j) = 0; i \neq j$
- Spremenljivke X_i, F_i in E_i naj bodo centrirane
 $E(X_i) = E(F_i) = E(E_i) = 0$

Cilj

Ali so zveze med opazovanimi spremenljivkami pojasnljive z manjšim številom posredno opazovanih faktorjev (latentne spremenljivke)?

Namen faktorske analize

- Seznanitev s strukturo v ozadju podatkov - pojasnitev podatkov z manjšim številom skupnih faktorjev (*exploratory factor analysis*); seznanitev z medsebojnimi odnosi med spremenljivkami, pred uporabo ostalih metod analize (regresijska analiza...);
- Preskušanje domnev glede ustreznega števila skupnih faktorjev oziroma glede spremenljivk, ki naj bi bile pod vplivom določenega faktorja (*confirmatory factor analysis*);

Analiza uspeha (primer) – 1/1

- Spearman, 1904: Prvi poskus uporabe faktorске analize za analizo uspeha učencev.
- Skupni uspeh učencev je želel oceniti na podlagi ocen pri treh predmetih:
 - X_1 - klasične vede;
 - X_2 – francoščina;
 - X_3 – angleščina.

Analiza uspeha (primer) – 1/2

Povezanost med ocenami predmetov (korelacijska matrika)

$$\begin{array}{c} X_1 \\ X_2 \\ X_3 \end{array} \begin{array}{c} X_1 \\ X_2 \\ X_3 \end{array} \begin{pmatrix} 1 & 0,83 & 0,78 \\ & 1 & 0,67 \\ & & 1 \end{pmatrix}$$



Analiza uspeha (primer) – 1/3

Faktorski model:

Analiza uspeha (primer) – 1/4

Faktorski model:

$$X_1 = a_{11}F_1 + E_1$$

$$X_2 = a_{21}F_1 + E_2$$

$$X_3 = a_{31}F_1 + E_3$$

Specifični vplivi E_i bodo imeli majhno variabilnost, če so opazovane spremenljivke (ocene) blizu faktorju F_1 (splošna sposobnost učenca).

Analiza zadovoljstva potrošnikov (primer) 2/1

- Na podlagi kvalitativne analize, identificiramo karakteristike prodajaln, ki jih potrošniki smatrajo kot pomemben dejavnik njihovega zadovoljstva
- Med množico posamičnih karakteristik (vzdušje v prodajalni, pestrost ponudbe, kakovost ponudbe, razpoložljivost proizvodov, znanje prodajalcev, prijaznost prodajalcev, pomoč...) skušamo odkriti splošne dimenzije, ki vplivajo na zadovoljstvo potrošnikov (osebje, ponudba, cene).

Ocenjevanja faktorkega modela - 1

- Varianco spremenljivke X_i lahko razbijemo na del, ki je pojasnjen s skupnimi faktorji, in na specifično varianco:

$$\sigma_i^2 = \sum_{j=1}^k a_{ij}^2 + \psi_{ii}$$

- Delež variance pojasnjen s skupnimi faktorji (h_i^2), imenujemo komunaliteta:

$$h_i^2 = \sum_{j=1}^k a_{ij}^2$$

Ocenjevanja faktorkega modela - 2

- Potreben pogoj za identifikacijo faktorkega modela (k – število faktorjev, m – število merjenih spremenljivk):
- Komunalitete lahko določimo šele tedaj, ko določimo skupne faktorje, ki pa jih lahko izračunamo iz popravljenih vrednosti komunalitet.

Ocenjevanja faktorkega modela - 3

Faktorski model ocenjujemo v dveh korakih:

- ocena deleža variance proučevanih spremenljivk pojasnjenega s skupnimi faktorji (komunalitete) z metodo glavnih osi ali metodo največjega verjetja in
- ocena faktorskih uteži s poševno in pravokotno rotacijo.

Metode ocenjevanja faktorkega modela

- Metoda glavnih osi (principal axis factoring):
 - Iteracijski postopek ocenjevanja skupnih faktorjev in komunalitet.
 - Postopek ne skonvergira vedno
- Metoda največjega verjetja (maximum likelihood):
 - Predpostavka, da so skupni in specifični faktorji normalno porazdeljeni (v tem primeru se normalno porazdeljujejo tudi vrednosti opazovanih spremenljivk)

Praktični napotki za uporabo - 1

- Katere vrste spremenljivk lahko uporabljamo? (*intervalne in razmernostne, izjemoma nominalne*)
- Koliko spremenljivk naj bo vključenih v faktorsko analizo? (*5 ali več spremenljivk na posamezni faktor*)
- Velikost vzorca? (*najmanj 50 enot, sicer priporočljivo 100 in več enot oziroma 5 do 10 enot na vsako spremenljivko*)

Praktični napotki za uporabo - 2

- Analiza korelacijske matrike (kor.koeficient $> 0,3$ oziroma parcialne korelacije $< 0,7$)
- Bartlettov test (preverjamo domnevo, da je korelacijska matrika enotska)
- $MSA > 0,5$ (tako na nivoju modela kot na nivoju posamezne spremenljivke)

Praktični napotki za uporabo - 3

- Ustrezno število faktorjev
 - Vsebina
 - Kriterij lastnih vrednosti (eigenvalues)
 - Kriterij deleža skupne pojasnjene variance (60% ali več)
 - Scree-diagram

Interpretacija faktorkega modela

- Faktorske uteži (najmanj $\pm 0,3$ oziroma $\pm 0,4$ sicer $\pm 0,5$ ali več zaželjeno)
- Komunalitete ($>0,5$)
- Poimenovanje faktorjev

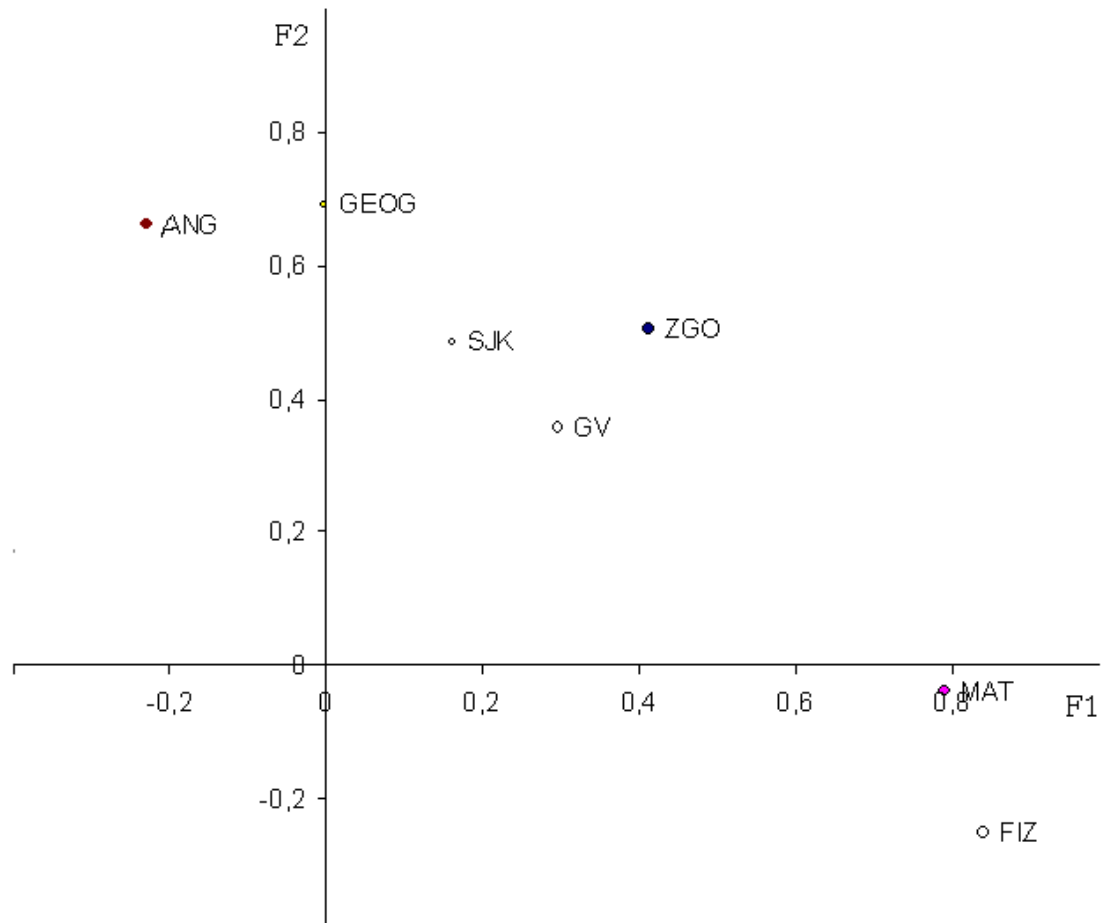
Analiza uspešnosti - primer

- V okviru analize uspešnosti dijakov smo opazovali vzorec 45 dijakov. Glede na dejstvo, da uspešnosti ne moremo neposredno meriti smo si za indikatorje uspešnosti izbral ocene pri izbranih predmetih.
 - a) Prikažite porazdelitve dijakov glede na njihove ocene s histogrami.
 - b) Oblikujte korelacijsko matriko in komentirajte odvisnosti med ocenami posameznih predmetov.
 - c) S pomočjo faktorjske analize ocenite skupne dejavnike, ki vplivajo na ocene dijakov pri posameznih predmetih. Uporabite metodo glavnih osi in metodo največjega verjetja. Koliko skupnih faktorjev ste predpostavili? Zakaj? Primerjava rezultatov obeh modelov...
 - d) Interpretirajte dobljene rezultate (komunalitete, faktorjske uteži) posameznega modela.

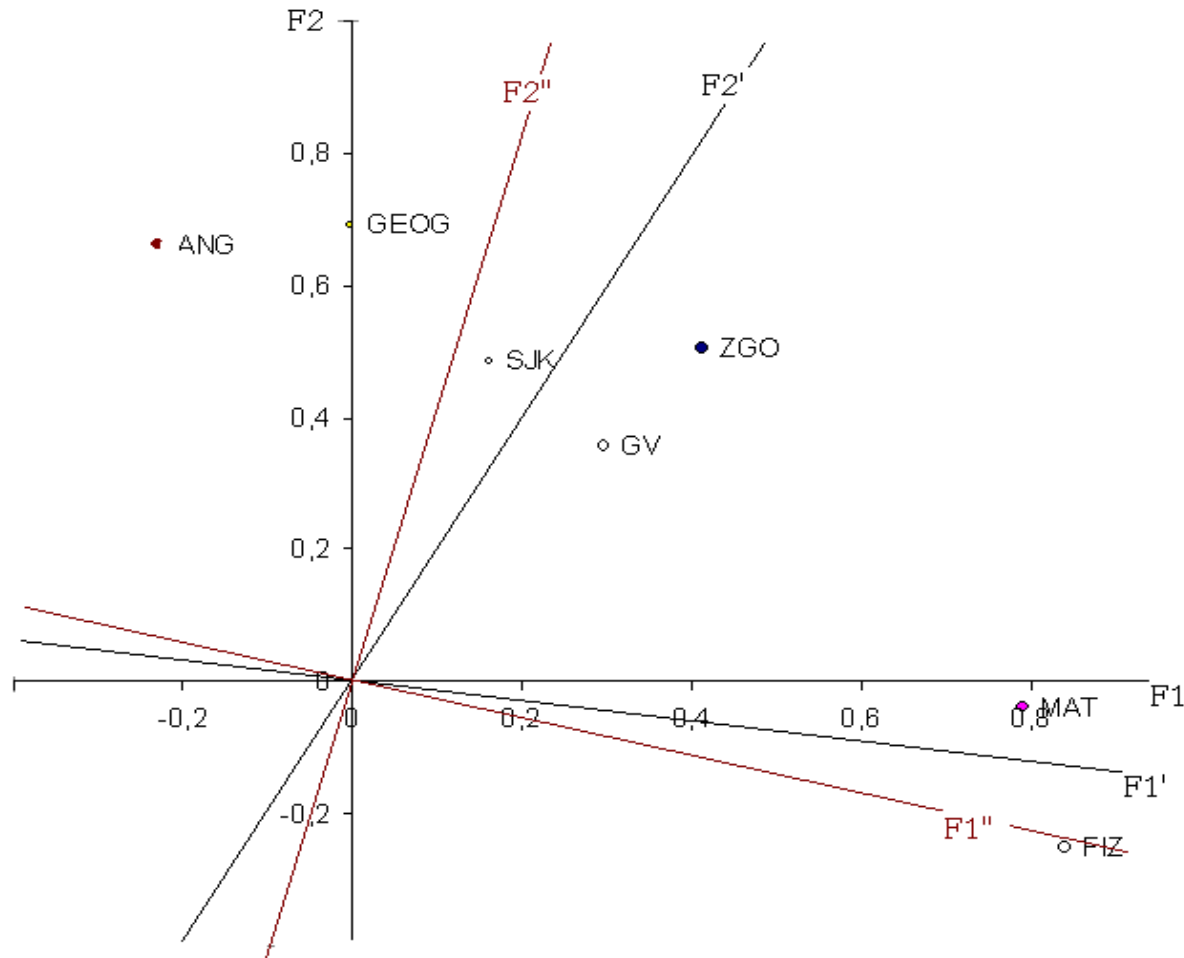
Rotacije faktorjev

- V primeru, da dobljenih skupnih faktorjev ne znamo smiselno interpretirati se odločimo za rotacijo faktorjev (če so projekcije določene spremenljivke precejšnje na več faktorjih, če imamo splošen faktor...).
- Želimo izrazitejše factorske uteži in čim enostavnejšo factorsko strukturo.

Osnovna faktorska rešitev



Rotirani rešitvi (pravokotna in poševna)



Načela rotiranja

- Thurston-ova načela enostavne faktorске strukture:
 - Vsaka spremenljivka naj ima vsaj pri enem faktorju ničlo (utež med 0,2 in -0,2);
 - Če je k -skupnih faktorjev, naj ima vsak faktor v matriki vsak k ničel (utež med 0,2 in -0,2);
 - Za vsak par faktorjev naj bo več spremenljivk, ki imajo močne uteži v enem stolpcu in majhne na ostalih;
 - Za vsak par faktorjev naj ima velik del spremenljivk majhne uteži na obeh faktorjih;
 - Za vsak par faktorjev naj bo le majhen del spremenljivk z utežmi različnimi od 0 na obeh faktorjih;

Pravokotne rotacije

- Uporabljamo, kadar faktorji niso korelirani med seboj;
- QURTIMAX, VARIMAX in EQIMAX;
- Najbolj poznana je VARIMAX metoda, pri kateri skušamo doseči, da ima vsaka opazovana spremenljivka visoko faktorsko utež le pri enem faktorju, pri drugih faktorjih pa so uteži skoraj enake 0.

Poševne rotacije

- Uporabljamo v primeru, ko so faktorji korelirani med seboj;
- OBLIMIN, OBLIMAX, QUARTIMIN, COVARMIN IN BIQUARTIMIN;
- Dobimo dvoje uteži:
 - **'pattern' uteži:** regresijski koeficienti, ki so običajno manjši od strukturnih uteži in bolj jasno pokažejo sliko kaj je tipično za posamezni faktor;
 - **strukturne uteži:** koeficienti korelacije med spremenljivko in faktorjem.

Analiza uspešnosti - primer

- Faktorje, ki ste jih pri analizi uspešnosti dijakov ocenili po metodi glavnih osi oziroma metodi največjega verjetja rotirajte s:
 - Pravokotno rotacijo (VARIMAX);
 - Poševno rotacijo (OBLIMIN).
- Primerjajte uteži dobljene z različnimi metodami oziroma rotacijami. Podajte sklepne ugotovitve.