FAKTORSKA ANALIZA

Začetnikom faktorske analize smatra se Charles Edward Spearman koji je prvi postavio dvofaktorsku teoriju inteligencije, izrazivši je jednadžbom

Charles Edward Spearman
(1863. – 1945.)

$$\boldsymbol{z_j} = a_j\, \boldsymbol{g} + \boldsymbol{s_j}$$

gdje je

- $\boldsymbol{z_j}$ - manifestna varijabla *j*
- a_j - koeficijent utjecaja generalnog faktora $\boldsymbol{g}$ na manifestnu varijablu *j*
- $\boldsymbol{g}$ - generalni faktor
- $\boldsymbol{s_j}$ - faktor specifičan samo za manifestnu varijablu *j*.

Za empirijsku provjeru svog teoretskog koncepta Spearman je 1904. godine razvio prvi model faktorske analize.

FAKTORSKA ANALIZA

Značajnu ulogu u razvoju faktorske analize imao je i američki psihometričar Louis Leon Thurstone koji je u svrhu empirijske provjere svoje teoriju o postojanju više primarnih mentalnih sposobnosti 1931. godine razvio multifaktorsku analizu.

Louis Leon Thurstone
(1887. – 1955.)

FAKTORSKA ANALIZA

Faktorska analiza je zajedničko ime za više metoda kojima je cilj kondenzacija većeg broja ***manifestnih varijabli***, među kojima postoji povezanost (korelacija), na manji broj ***latentnih dimenzija*** (***faktora***) koje su izvor te povezanosti.

Faktorska analiza se, osim za izračunavanje rezultata entiteta u latentnim dimenzijama, koristi i u svrhu uvida u strukturu međusobne povezanosti više manifestnih varijabli.

FAKTORSKA ANALIZA

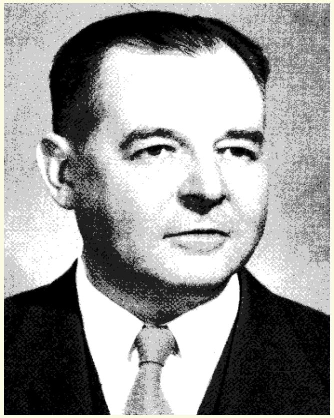
Manifestne varijable su varijable, odnosno obilježja koja se mogu izravno mjeriti postojećim mjernim instrumentima (npr. skok udalj s mjesta, bacanje medicinke iz ležanja na leđima, bacanje kugle).

Latentne dimenzije su varijable, odnosno obilježja koja nisu izravno mjerljiva postojećim mjernim instrumentima, već se izračunavaju kao linearna kombinacija manifestnih varijabli (npr. eksplozivna snaga).

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Komponentni model ili ***metoda glavnih komponenata*** je metoda ekstrakcije, odnosno izračunavanja latentnih dimenzija koju je 1933. godine predložio američki ekonomist i statističar Harold Hotelling.



Harold Hotelling
(1895. – 1973.)

Metodom glavnih komponenata se iz skupa od *m* manifestnih varijabli na temelju nereducirane korelacijske matrice izračuna *m* latentnih dimenzija koje su međusobno linearno nezavisne, a nazivaju se ***glavne komponente***.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Glavne komponente su linearne kombinacije manifestnih varijabli izračunate na način da prva glavna komponenta objašnjava maksimalan moguć dio ukupne varijance manifestnih varijabli te da druga, kao i svaka sljedeća glavna komponenta, objašnjava najveći dio preostale varijance manifestnih varijabli, odnosno najveći dio varijance manifestnih varijabli koji nije objašnjen prethodnim glavnim komponentama.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Glavne komponente izračunavaju se na sljedeći način:

Neka su u matrici **B** podaci skupa $E = \{e_i \,;\, i = 1,\ldots,n\}$ entiteta koji su opisani skupom $V = \{v_j \,;\, j = 1,\ldots,m\}$ manifestnih varijabli. Operacijom

$$\mathbf{Z} = \mathbf{B}_c \mathbf{V}^{-1}$$

gdje je

- $\mathbf{B}_c$ - matrica centriranih podataka dobivenih operacijom $\mathbf{B}_c = \mathbf{B} - \mathbf{1m}^T$
- $\mathbf{m}$ - vektor aritmetičkih sredina varijabli matrice **B**
- $\mathbf{V}^{-1}$ - dijagonalna matrica standardnih devijacija varijabli iz matrice **B**

izračuna se matrica standardiziranih podataka (**Z**).

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Operacijom

$$\mathbf{R} = \mathbf{Z}^{\mathbf{T}}\mathbf{Z}n^{-1} = \begin{vmatrix} 1 & r_{12} & . & . & r_{1m} \\ r_{21} & 1 & . & . & r_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ r_{m1} & r_{m2} & . & . & 1 \end{vmatrix}$$

izračuna se matrica korelacija manifestnih varijabli (***R***) u čijoj su glavnoj dijagonali jedinice, odnosno varijance standardiziranih varijabli. Zbroj varijanci svih standardiziranih varijabli, odnosno ukupna varijanca skupa standardiziranih varijabli jednaka je *m*, odnosno broju manifestnih varijabli.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Operacijom spektralne dekompozicije matrice korelacija

$$\boldsymbol{R = X\,\lambda\,X^{T}}$$

$$\begin{vmatrix} 1 & r_{12} & . & . & r_{1m} \\ r_{21} & 1 & . & . & r_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ r_{m1} & r_{m2} & . & . & 1 \end{vmatrix} = \begin{vmatrix} x_{11} & x_{12} & . & . & x_{1m} \\ x_{21} & x_{22} & . & . & x_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1} & x_{m2} & . & . & x_{mm} \end{vmatrix} \cdot \begin{vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & . & \\ & & & . & \\ & & & & \lambda_m \end{vmatrix} \cdot \begin{vmatrix} x_{11} & x_{21} & . & . & x_{m1} \\ x_{12} & x_{22} & . & . & x_{m2} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{1m} & x_{2m} & . & . & x_{mm} \end{vmatrix}$$

izračunaju se ***matrica svojstvenih vektora*** (***X***) i ***matrica svojstvenih vrijednosti*** (λ).

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Matrica svojstvenih vektora (***X***) je kvadratna matrica reda $m \times m$ za koju vrijedi da je $\boldsymbol{X^T X} = \boldsymbol{X X^T} = \boldsymbol{I}$ tj. da su vektori matrice linearno nezavisni, a u čijim su stupcima ponderi za izračunavanje glavnih komponenata iz standardiziranih varijabli.

Matrica svojstvenih vrijednosti ($\boldsymbol{\lambda}$) je dijagonalna matrica reda $m \times m$ za koju vrijedi da je $\lambda_1 + \lambda_2 + \ldots + \lambda_m = m$, tj. da je zbroj svojstvenih vrijednosti jednak ukupnoj varijanci standardiziranih varijabli, te da je $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m$, tj. da se svojstvene vrijednosti nižu od najveće prema najmanjoj.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Operacijom

$$\boldsymbol{K = Z\ X}$$

$$\begin{vmatrix} k_{11} & k_{12} & . & . & k_{1m} \\ k_{21} & k_{22} & . & . & k_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ k_{n1} & k_{n2} & . & . & k_{nm} \end{vmatrix} = \begin{vmatrix} z_{11} & z_{12} & . & . & z_{1m} \\ z_{21} & z_{22} & . & . & z_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ z_{n1} & z_{n2} & . & . & z_{nm} \end{vmatrix} \cdot \begin{vmatrix} x_{11} & x_{12} & . & . & x_{1m} \\ x_{21} & x_{22} & . & . & x_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1} & x_{m2} & . & . & x_{mm} \end{vmatrix}$$

izračuna se ***matrica glavnih komponenata*** (***K***).

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Matrica glavnih komponenata (***K***) je matrica reda $n \times m$ koju čine rezultati n ispitanika u m linearno nezavisnih (ortogonalnih) glavnih komponenata. Varijanca prve glavne komponente jednaka je λ_1, varijanca druge glavne komponente λ_2, itd.

Pošto vrijedi da je $\lambda_1 + \lambda_2 + ... + \lambda_m = m$ te da je $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_m$, može se zaključiti da je ukupna varijanca m standardiziranih varijabli raspodijeljena tako da prva glavna komponenta objašnjava najveći dio ukupne varijance, druga glavna komponenta najveći dio preostale varijance, itd.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

Operacijom

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

$$\begin{vmatrix} h_{11} & h_{12} & . & . & h_{1m} \\ h_{21} & h_{22} & . & . & h_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ h_{m1} & h_{m2} & . & . & h_{mm} \end{vmatrix} = \begin{vmatrix} x_{11} & x_{12} & . & . & x_{1m} \\ x_{21} & x_{22} & . & . & x_{2m} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{m1} & x_{m2} & . & . & x_{mm} \end{vmatrix} \cdot \begin{vmatrix} \sqrt{\lambda_1} & & & & \\ & \sqrt{\lambda_2} & & & \\ & & . & & \\ & & & . & \\ & & & & \sqrt{\lambda_m} \end{vmatrix}$$

izračuna se ***matrica glavnih osovina*** (***H***).

FAKTORSKA ANALIZA

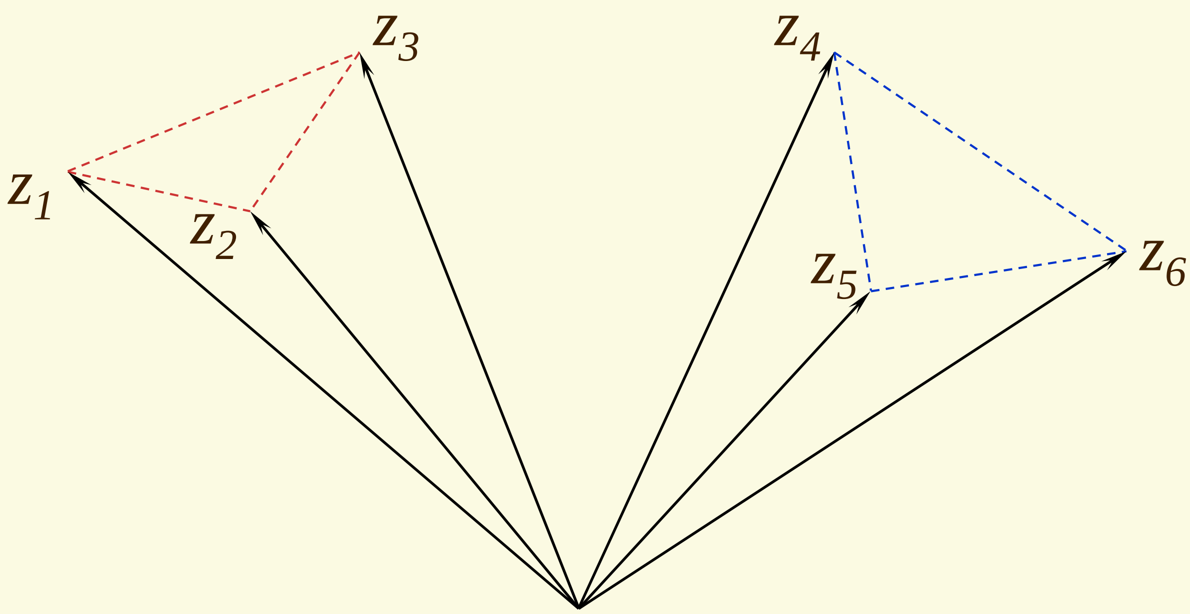
Komponentni model faktorske analize

Matrica glavnih osovina (***H***) je matrica korelacija manifestnih varijabli i glavnih komponenata. Elementi ove matrice ukazuju koliki je doprinos svake pojedine manifestne varijable pri formiranju svake pojedine glavne komponente.

FAKTORSKA ANALIZA

Komponentni model faktorske analize

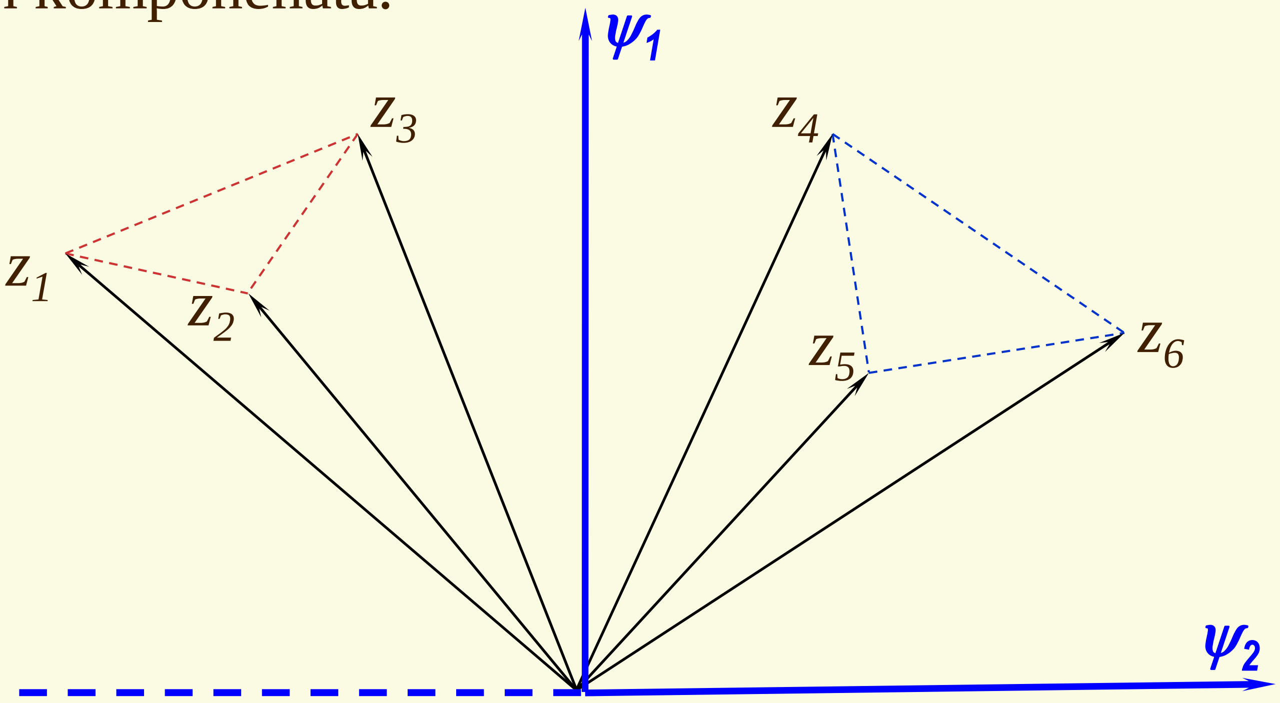
Neka je skup od *n* entiteta opisan sa 6 manifestnih varijabli (v_1,...,v_6). Na temelju korelacijske matrice može se utvrditi odnos odgovarajućih standardiziranih vektora u *n*-dimenzionalnom prostoru.



FAKTORSKA ANALIZA

Komponentni model faktorske analize

Na temelju matrice glavnih osovina može se utvrditi prostorni odnos standardiziranih manifestnih vektora i standardiziranih glavnih komponenata.



FAKTORSKA ANALIZA

Kriteriji za odabir značajnih glavnih komponenata

Pošto je osnovni cilj faktorske analize kondenzacija većeg broja manifestnih varijabli na manji broj latentnih dimenzija, potrebno je izvršiti ***redukciju broja glavnih komponenata***.

Redukcija broja glavnih komponenata odnosno određivanje broja značajnih glavnih komponenata vrši se putem različitih kriterija kao što su ***GK-kriterij***, ***PB-kriterij*** i ***Scree-test***.

FAKTORSKA ANALIZA

Kriteriji za odabir značajnih glavnih komponenata

Prema ***GK-kriteriju*** glavna komponenta *j* je značajna ako je njena varijanca (λ_j) veća ili jednaka *1*.

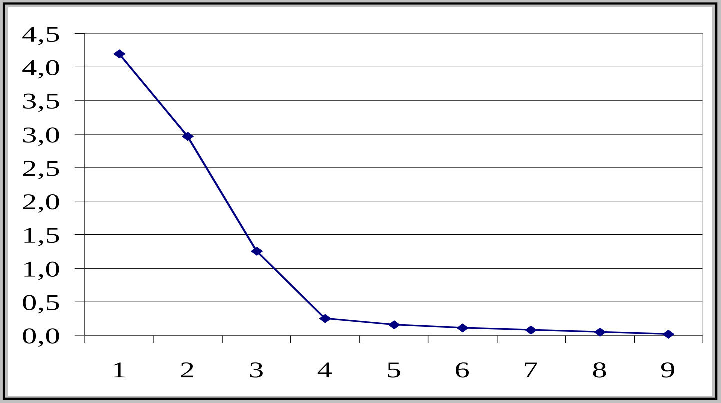
Prema ***PB-kriteriju*** broj značajnih glavnih komponenata jednak je broju svojstvenih vrijednosti poredanih po veličini čiji zbroj ne prelazi Σsmc (sumu kvadrata multiplih korelacija).

Scree-test je grafički kriterij. Na *scree plotu* se subjektivnom procjenom odredi točka nakon koje se svojstvene vrijednosti smanjuju u skladu s blagim linearnim trendom. Značajnima se smatraju sve prethodne glavne komponente.

FAKTORSKA ANALIZA

Kriteriji za odabir značajnih glavnih komponenata

Zadatak - Na 9 manifestnih varijabli utvrđene su sljedeće svojstvene vrijednosti: $\lambda_1=4,18$, $\lambda_2=2,96$, $\lambda_3=1,27$, $\lambda_4=0,22$, $\lambda_5=0,15$, $\lambda_6=0,11$, $\lambda_7=0,07$, $\lambda_8=0,03$ i $\lambda_9=0,01$. $\Sigma smc=8,14$, a odgovarajući scree plot izgleda ovako:



Koliko je značajnih glavnih komponenata prema GK-kriteriju, prema PB-kriteriju, a koliko prema Scree-testu?

FAKTORSKA ANALIZA

Komunaliteti i unikviteti

značajne glavne komponente | glavne komponente koje nisu značajne

$$\boldsymbol{H} = \left| \begin{array}{cccc:cccc} h_{11} & . & . & h_{1k} & . & . & . & h_{1m} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ h_{m1} & . & . & h_{mk} & . & . & . & h_{mm} \end{array} \right|$$

(Matrica ***H*** prije i nakon redukcije broja glavnih komponenata)

FAKTORSKA ANALIZA

Komunaliteti i unikviteti

Varijancu svake manifestne varijable moguće je dekomponirati na ***komunalitet*** (h^2) i ***unikvitet*** (u^2) pri čemu je

$$s^2 = 1 = h^2 + u^2$$

Komunalitet je dio varijance manifestne varijable *j* koji je moguće objasniti s *k* značajnih glavnih komponenata.

$$h_j^2 = \sum_{p=1}^{k} h_{jp}^2 = h_{j1}^2 + h_{j2}^2 + \ldots + h_{jk}^2$$

Unikvitet je dio varijance manifestne varijable *j* koji nije moguće objasniti s *k* značajnih glavnih komponenata.

$$u_j^2 = 1 - h_j^2$$

STATISTICA 7

Komponentni model faktorske analize

Komponentni model faktorske analize provodi se slijedom koraka: padajući izbornik *Statistics* → *Multivariate Exploratory Techniques* → *Factor analysis*. U dijaloškom okviru koji se pokreće odabirom opcije *Variables* potrebno je označiti manifestne varijable. Nakon odabira varijabli u dijaloškom okviru *Advanced* potrebno je označiti opciju *Principal components* i definirati kriterij za redukciju broja glavnih komponenata. Pri tome je moguće iskoristiti opcije *Max. no. of factors* (maksimalan broj faktora) i *Mini. eigenvalue* (minimalna svojstvena vrijednost).

STATISTICA 7

Komponentni model faktorske analize

Nakon ekstrakcije glavnih komponenata moguće je pregledati svojstvene vrijednosti (*Explained variance* → *Eigenvalues*), komunalitete (*Explained variance* → *Communalities*), matricu glavnih osovina (*Loadings* → *Summary: Factor loadings*) i matricu glavnih komponenata (*Scores* → *Factor scores*).

Zadatak - Na varijablama matrice *Judo.sta* provedite komponentni model faktorske analize uz GK-kriterij za redukciju broja glavnih komponenata! Izračunajte svojstvene vrijednosti, komunalitete, matricu glavnih osovina i matricu glavnih komponenata!

FAKTORSKA ANALIZA

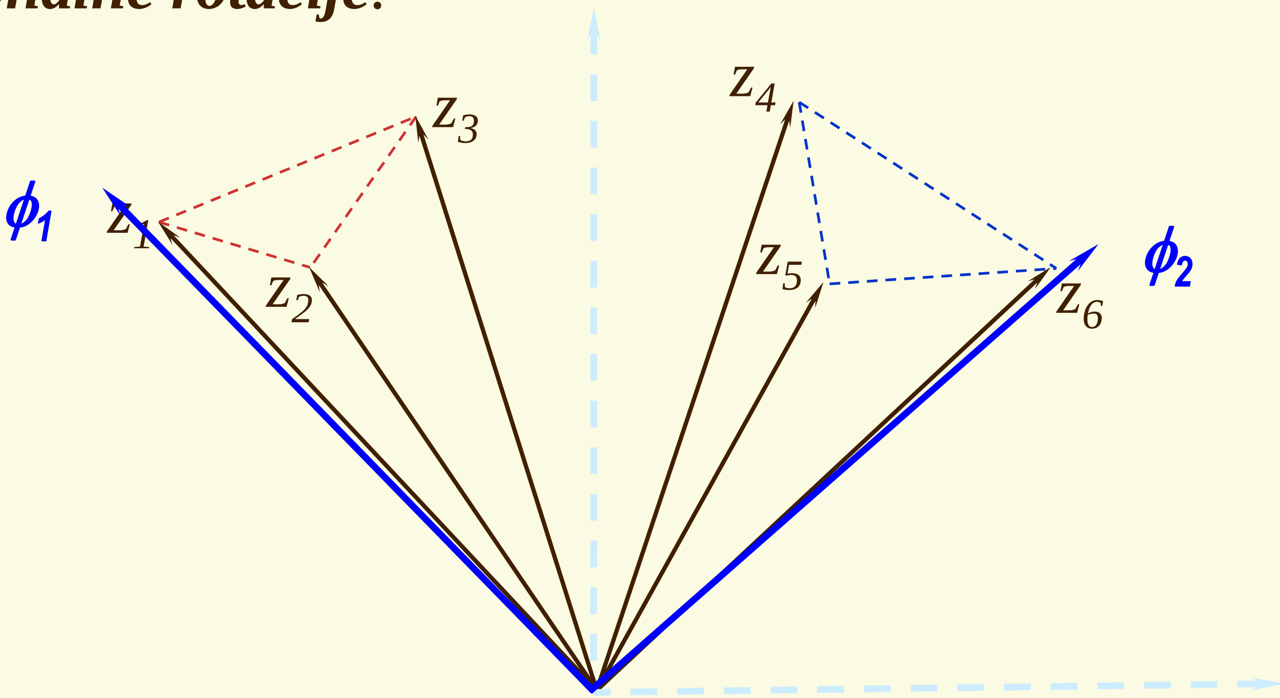
Rotacije

Pravi uvid u strukturu međusobnih odnosa manifestnih varijabli često nije moguće steći putem glavnih komponenata. U takvim se slučajevima koriste transformacije glavnih komponenata čija je svrha postizanje jednostavne faktorske strukture, a koje se nazivaju ***rotacije***.

FAKTORSKA ANALIZA

Rotacije

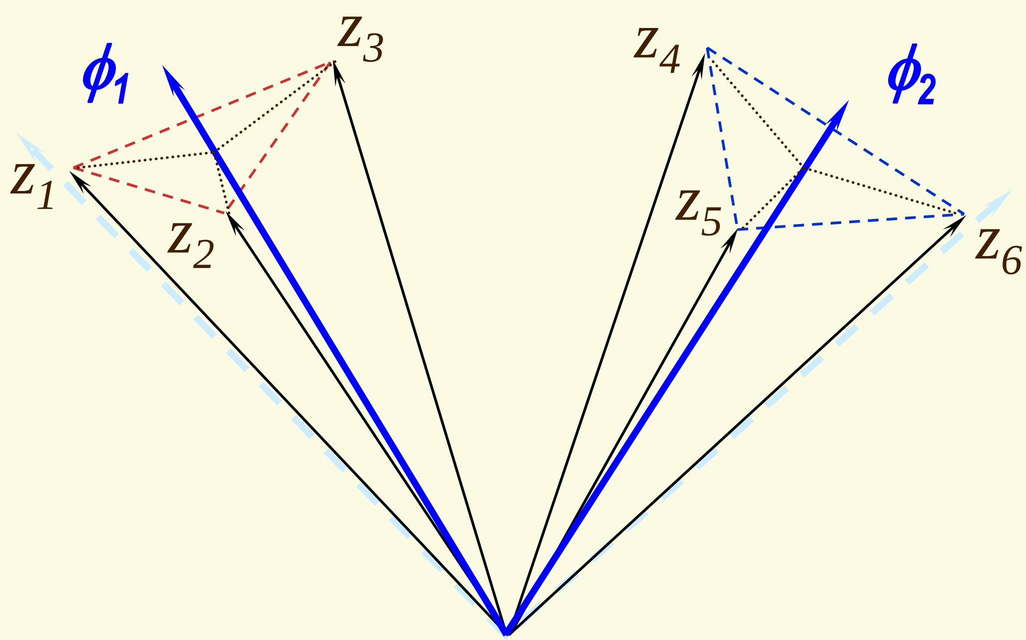
Transformacije glavnih komponenata koje se provode uz uvjet zadržavanja njihove linearne nezavisnosti nazivaju se ***ortogonalne rotacije***.



FAKTORSKA ANALIZA

Rotacije

Transformacije glavnih komponenata koje se provode bez uvjeta zadržavanja njihove linearne nezavisnosti nazivaju se ***neortogonalne*** ili ***kosokutne rotacije***.



FAKTORSKA ANALIZA

Rotacije

Interpretacija faktora nakon rotacije vrši se putem ***matrice strukture*** (***F***), ***matrice sklopa*** (***A***), i ***matrice korelacija među faktorima*** (***M***).

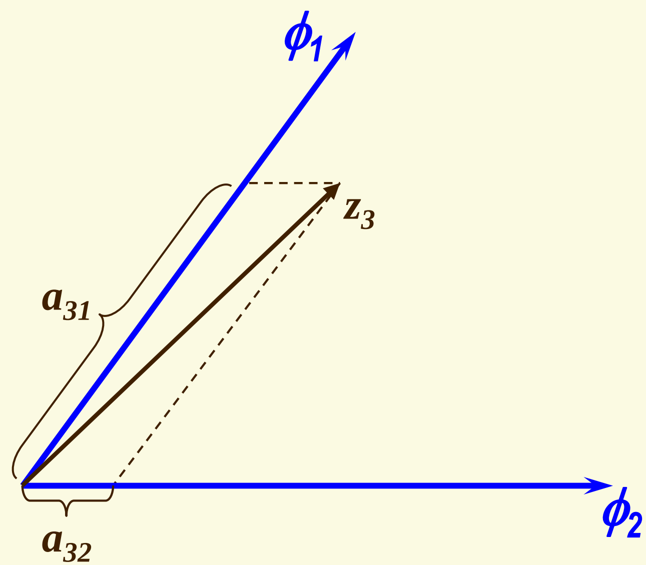
Matricu strukture čine korelacije manifestnih varijabli s faktorima.

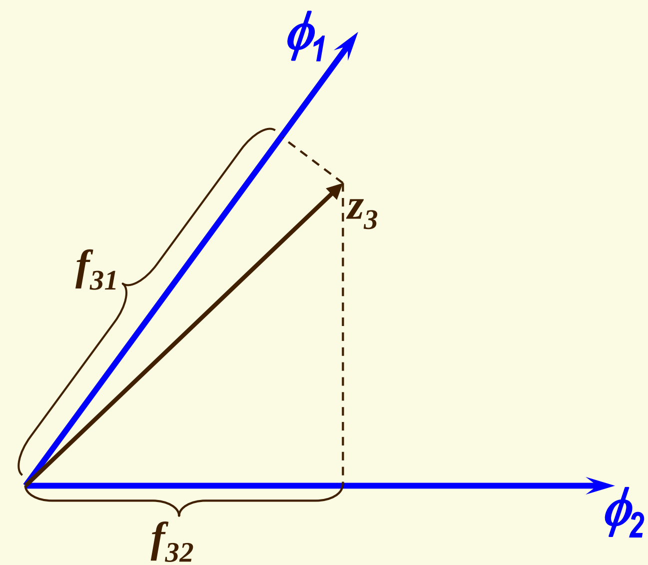
Matricu sklopa čine paralelne projekcije manifestnih varijabli na faktore.

FAKTORSKA ANALIZA

Rotacije

Matrica sklopa (paralelne projekcije manifestnih varijabli na faktore) često jasnije pokazuje koje varijable određuju pojedine faktore nego matrica strukture (ortogonalne projekcije manifestnih varijabli na faktore).

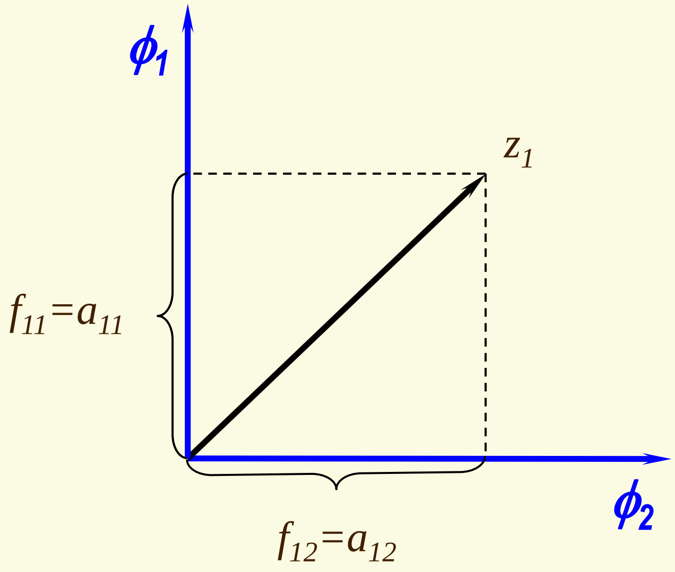




FAKTORSKA ANALIZA

Rotacije

Paralelne projekcije i ortogonalne projekcije manifestnih varijabli na faktore bit će sličnije što su korelacije među faktorima manje. Matrica strukture bit će jednaka matrici sklopa ako su faktori potpuno nezavisni.



STATISTICA 7

Rotacije

Nakon ekstrakcije glavnih komponenata moguće je izvršiti njihovu transformaciju odabirom neke od ponuđenih rotacija u padajućem izborniku *Loadings* → *Factor rotation*. Nakon rotacije glavnih komponenata moguće je pregledati matricu strukture (*Loadings* → *Summary: Factor loadings*) i matricu rezultata entiteta na faktorima (*Scores* → *Factor scores*).

Zadatak - Na varijablama matrice *Judo.sta* provedite komponentni model faktorske analize uz GK-kriterij redukcije i normaliziranu Varimax rotaciju! Izračunajte matricu faktorske strukture i matricu rezultata entiteta na faktorima!

FAKTORSKA ANALIZA

Literatura za pripremanje kolokvija

Dizdar, D. (2006). *Kvantitativne metode*. Zagreb: Kineziološki fakultet, str. 214-237.

Mejovšek, M. (2003). Uvod u metode znanstvenog istraživanja u društvenim i humanističkim znanostima. Jastrebarsko: Naklada Slap, str. 153-181.