INSTITUT ZA KINEZIOLOGIJU
FAKULTETA ZA FIZIČKU KULTURU
SVEUČILIŠTA U ZAGREBU
ODJEL ZA INFORMATIKU I STATISTIKU

KOMPJUTERSKI PROGRAMI ZA KLASIFIKACIJU, SELEKCIJU, PROGRAMIRANJE I KONTROLU TRENINGA

1.

SPECIJALNI PROGRAMI

ZAGREB
1983

AUTORI: Prof. dr KONSTANTIN MOMIROVIĆ
JANEZ ŠTALEC
FRANJO PROT
KSENIJA BOSNAR
Mr NATAŠA VISKIĆ
LEO PAVIČIĆ
VESNA DOBRIĆ

Suradnici i konsultanti: RATKO GOSPODNETIĆ
ŽIVAN KARAMAN

Ovaj elaborat izradjen je na temelju ugovora broj 1940 od 22.12.1983. izmedju Republičke samoupravne interesne zajednice za fizičku kulturu u Zagrebu i Fakulteta za fizičku kulturu Sveučilišta u Zagrebu.

Ugovor je zaveden pod brojem 05-5509/1-83, 23.12.1983. na Fakultetu za fizičku kulturu.

SADRŽAJ

1. UVOD

Za realizaciju znanstvenih i stručnih radova i neposrednu primjenu ovih rezultata u praksi neophodna je primjena elektroničkih računala, pogotovo u postupcima klasifikacije, selekcije, programiranja i kontrole treninga. Ovo zbog toga što svi ti postupci zahtijevaju složene operacije sa velikim brojem varijabli i primjenu kompliciranih matematičkih i statističkih postupaka.

Svrha ovog projekta je izrada odredjenog broja programa za analizu podataka na elektroničkom računalu koji su u ovom času najpotrebniji za dalji nastavak znanstvenog i stručnog rada i primjenu rezultata u praksi. Tim je programima omogućena primjena novih metoda za analizu podataka, od kojih su neke posebno razvijene za primjenu na malom broju ispitanika, što je najčešće slučaj u vrhunskom sportu.

Na temelju analize potreba iskazanih u programima znanstvenog i stručnog rada izradjeni su ovi programi:

(1) Program za simultanu kanoničku i kvazikanoničku analizu, kojim se omogućava pouzdanija analiza i predvidjanje višekriterijalnih rezultata u borilačkim sportovima i sportskim igrama (program QCCR);

(2) Program za robustnu diskriminativnu analizu, kojim se omogućava klasifikacija i usmjeravanje polaznika sportskih škola (program SDA);

(3) Program za višekriterijalnu selekciju sa ograničenjima kojim se omogućava planski rad na usmjeravanju i izboru djece i omladine za pojedine sportske discipline (program CORAMIN);

(4) Program za simultanu regresijsku i kvaziregresijsku analizu, kojim se omogućava pouzdanija izrada jednadžbi specifikacija u sportovima sa malo vrhunskih takmičara (program CAOS);

(5) Program za diskriminativnu analizu u Mahalanobisovom prostoru, kojim se omogućava odredjivanje optimalnog psihofizičkog profila sportaša za odredjenu sportsku disciplinu (program DIANA);

(6) Program za analizu klika, kojim se omogućava odredjivanje sociometrijske strukture momčadi (program ITA);

(7) Program za analizu strukturalnih promjena, kojim se omogućava analiza kvalitativnih promjena u toku treninga (program COMPAT-ZZ);

(8) Program za kanoničku analizu strukturalnih promjena, kojim se omogućava utvrdjivanje kvalitativnih promjena nezavisno od kvantitativnih promjena (program LIMAX);

(9) Program za analizu konačnog broja sadržaja tranzitivnih operatora, kojim se omogućava odredjivanje karakteristika treninga u pojedinim fazama trenažnog ciklusa (program QUQULELE);

(10) Program za dekompoziciju krivulja vježbe, kojim se omogućava analiza kvantitativnih promjena u toku treninga (program LAV);

(11) Program za kose Prokrustove transformacije pomoću ortogonalnih rotacija, koji omogućava ispitivanje ispravnosti hipoteza o strukturi psihofizičkih osobina sportaša (program SCHOENEMANN);

(12) Program za usporedbu objektivne strukture psihofizičkih karakteristika i teoretske strukture tih karakteristika, kojim se omogućava ispitivanje teoretskih koncepcija o strukturi nekog skupa karakteristika (program QUAQUA).

Svi su programi napisani u programskom jeziku SS i pohranjeni u programskoj biblioteci FFK*LIB i javnoj, svakom korisniku dostupnoj programskoj biblioteci SRCE*SS-MAKRO.

Svi programi, koji realiziraju nove metode ili algoritme, ili bitnu modifikaciju postojećih algoritama, snabdjeveni su dokumentacijom u vidu članaka, koji su objavljeni ili će biti objavljeni, u 1983. ili 1984. godini u znanstvenim ili stručnim časopisima i zbornicima. Ti su članci pridruženi ovom elaboratu, koji sadržava i listinge programa i upute o njihovom korištenju.

2. PROGRAMI

2.1 ALGORITAM I PROGRAM QCCR
2.2 ALGORITAM I PRO RAM SDA
2.3 ALGORITAM I PROGRAM CORAMIN
2.4 ALGORITAM I PROGRAM CAOS
2.5 ALGORITAM I PROGRAM DIANA
2.6 ALGORITAM I PROGRAM ITA
2.7 ALGORITAM I PROGRAM COMPAT ZZ
2.8 ALGORITAM I PROGRAM LIMAX
2.9 ALGORITAM I PROGRAM QUOULELE
2.10 ALGORITAM I PROGRAM LAV
2.11 ALGORITAM I PROGRAM SCHOENEMANN
2.12 ALGORITAM I PROGRAM QUAQUA

2.1 ALGORITAM I PROGRAM QCCR

Jedan od osnovnih problema kineziologije jest utvrdjivanje relacija izmedju dva skupa varijabli, kao npr. izmedju obilježja trenažnog procesa i efikasnosti u sportskoj aktivnosti, izmedju pokazatelja neke dimenzije psihosomatskog statusa i efekata treninga i sl. Rezultati analize relacija, medjutim, mogu u znatnijoj mjeri ovisiti od modela pod kojim je analiza vršena, pa tako izbor modela može utjecati i na donošenje odluka prilikom primjene rezultata u praksi.

Veći stupanj sigurnosti u zaključivanju i viši nivo kvalitete odluka može se postići analizom podataka pod različitim modelima i usporedbom dobijenih rezultata.

Stoga, definiran je algoritam i napisan program QCCR u kojem se relacije izmedju dva skupa podataka analiziraju pod klasičnim biortogonalnim kanoničkim korelacijskim modelom (Hotelling, 1935) i pod generalnim modelom za analizu relacija dva skupa kvantitativnih varijabli, tzv. kvazikanoničkim modelom (Momirović, Dobrić i Karaman, 1983), i usporedjuju dobijeni rezultati.

```
    *
    *   NAZIV PROGRAMA
    *
    *
    *
    *           *** G C C R ***
    *
    *
    *   AUTORI
    *
    *           K. MOMIROVIC
    *           V. DOBRIC
    *           F. PROT
    *           K. BOSNAR
    *
    *
    *   FUNKCIJA
    *
    *
    *   OVAJ PROGRAM PROVODI STANDARDNU BIORTOGONALNU KANONICKU KORELACIJSKU
    *   ANALIZU (HOTELLING,1935) I KANONICKU ANALIZU KOVARIJANCE (MOMIROVIC,
    *   DOBRIC I KARAMAN, 1983) TE UTVRDJUJE RELACIJE REZULTATA DOBIJENIH POD
    *   DVA MODELA.
    *
    *
    *   METODA I ALGORITAM OPISANI SU U
    *
    *           BOSNAR, K., F. PROT I K. MOMIROVIC:
    *           SOME RELATIONSHIPS BETWEEN CANONICAL COVARIANCE ANALYSIS
    *           AND CANONICAL CORRELATION ANALYSIS. 6TH INTERNATIONAL SYMPOSIUM
    *           ON COMPUTATIONAL STATISTICS, PRAGUE, 1984.
    *
    *
    *   PRIPREMA PODATAKA
    *
    *
    *   KONTROLNI ZAPIS I PODACI ZA PRVI SKUP VARIJABLI NALAZE SE U DATOTECI
    *   DATA. KONTROLNI ZAPIS I PODACI ZA DRUGI SKUP VARIJABLI NALAZE SE U
    *   DATOTECI ATA. BROJ VARIJABLI DRUGOG SKUPA TREBA BITI JEDNAK ILI MANJI
    *   OD BROJA VARIJABLI U PRVOM SKUPU.
    *
    *
    *   IZVRSNI DIO PROGRAMA
    *
    *
    * BLOK 0.
    * INPUT I CONFORM ULAZNIH PODATAKA
    *
    OUTPUT(DEVICE=PR)
    HEADING(TEXT=G C C R,T)
    HEADING(TEXT=KANONICKE I KVAZIKANONICKE RELACIJE)
    TEXT(TEXT=           G  C  C  R)
    INPUT (DATA=DATA,SCORE=BB1)
    INPUT (DATA=ATA,SCORE=BB2)
    CONFORM (IN1=BB1,IN2=BB2,OUT1=B1,OUT2=B2)
          DELETE (MATRIX=BB1)
          DELETE (MATRIX=BB2)
    *
    * BLOK 1.
    * INTERKORELACIJE I KROSKORELACIJE
    *
    HEADING (TEXT=INTERKORELACIJE I KROSKORELACIJE,D)
    CORRELATION (SCORE=B1,R=R1)
    CORRELATION (SCORE=B2,R=R2)
```

```
CROSSCORRELATION (P1=B1,P2=B2)
PRINT (MATRIX=R1,TEXT=INTERKORTELACIJE U PRVOM SKUPU)
PRINT (MATRIX=R2,TEXT=INTERKORELACIJE U DRUGOM SKUPU)
PRINT (MATRIX=R12,TEXT=KROSKORELACIJE    )
      DELETE (MATRIX=B2)
*
* BLOK 2.
* KANONICKA KORELACIJSKA ANALIZA
*
HEADING(TEXT=KANONICKA KORELACIJSKA ANALIZA,D)
CANCORR(Z=B1,X1=X1,X2=X2,Q=0.01)
      DELETE (MATRIX=B1)
MULT (A=R12,B=X2,TB,M=C12)
MULT (A=X1,B=R12,M=C21)
PRINT (MATRIX=X1,T,TEXT=KANONICKI KOEFICIJENTI PRVOG SKUPA)
PRINT (MATRIX=F1,T,TEXT=KANONICKI FAKTORI PRVOG SKUPA)
PRINT (MATRIX=C12,TEXT=KANONICKI KROSFAKTORI PRVOG SKUPA)
PRINT (MATRIX=X2,T,TEXT=KANONICKI KOEFICIJENTI DRUGOG SKUPA)
PRINT (MATRIX=F2,T,TEXT=KANONICKI FAKTORI DRUGOG SKUPA)
PRINT (MATRIX=C21,T,TEXT=KANONICKI KROSFAKTORI DRUGOG SKUPA,D)
*
* BLOK 3.
* KANONICKA KOVARIJANCNA ANALIZA
*
HEADING(TEXT=KANONICKA KOVARIJANCNA ANALIZA)
MULT (A=R12,TA,B=R12,M=R)
DIAGONALISATION.
HOTELLING.
MULT (A=F,B=F,TB,M=L)
DIAGMULT (A=F,T,D=L,C=-0.5,R,M=Y2)
MULT (A=R12,B=Y2,M=R12Y2)
DIAGMULT (A=R12Y2,D=L,C=-0.5,R,M=Y1)
MULT (A=R1,B=Y1,M=R1Y1)
MULT (A=R2,B=Y2,M=R2Y2)
MULT (A=Y1,TA,B=R1Y1,M=W1)
MULT (A=Y2,TA,B=R2Y2,M=W2)
SCALE (C=W1,R=M1)
SCALE (C=W2,R=M2)
MULT (A=Y1,TA,B=R12Y2,M=PSI)
DIAGMULT (A=PSI,D=W1,C=-0.5,L,M=LP)
DIAGMULT (A=LP, D=W2,C=-0.5,R,M=ETA)
DIAGMULT (A=R1Y1,D=W1,C=-0.5,R,M=S1)
INVERSION (R=M1,RINV=M1I)
MULT (A=S1,B=M1I,M=P1)
DIAGMULT (A=R2Y2,D=W2,C=-0.5,R,M=S2)
INVERSION (R=M2,RINV=M2I)
MULT (A=S2,B=M2I,M=P2)
DIAGMULT (A=R12Y2,D=W2,C=-0.5,R,M=Q12)
MULT (A=R12,TA,B=Y1,M=R12Y1)
DIAGMULT (A=R12Y1,D=W1,C=-0.5,R,M=Q21)
PRINT (MATRIX=PSI,TEXT=KVAZIKANONICKE KOVARIJANCE,D)
PRINT (MATRIX=ETA,TEXT=KVAZIKANONICKE KORELACIJE,D)
PRINT (MATRIX=Y1,TEXT=KVAZIKANONICKI PONDERI PRVOG SKUPA)
PRINT (MATRIX=P1,TEXT=KVAZIKANONICKI SKLOP FAKTORA PRVOG SKUPA,D)
PRINT (MATRIX=W1,TEXT=KOVARIJANCE KVAZIKANONICKIH FAKTORA PRVOG SKUPA,D)
PRINT (MATRIX=M1,TEXT=KORELACIJE KVAZIKANONICKIH FAKTORA PRVOG SKUPA,D)
PRINT (MATRIX=S1,TEXT=STRUKTURA KVAZIKANONICKIH FAKTORA PRVOG SKUPA)
PRINT (MATRIX=Q12,TEXT=KROSSTRUKTURA KVAZIKANONICKIH FAKTORA PRVOG SKUPA)
      DELETE (MATRIX=M2I)
      DELETE (MATRIX=R1Y1)
      DELETE (MATRIX=LAMBDA)
      DELETE (MATRIX=X)
      DELETE (MATRIX=F)
      DELETE (MATRIX=R12Y2)
```

```
      DELETE (MATRIX=LP)
      DELETE (MATRIX=L)
      DELETE (MATRIX=R)
PRINT (MATRIX=Y2,TEXT=KVAZIKANONICKI KOEFICIJENTI DRUGOG SKUPA)
PRINT (MATRIX=P2,TEXT=SKLOP KVAZIKANONICKIH FAKTORA DRUGOG SKUPA,D)
PRINT (MATRIX=W2,TEXT=KOVARIJANCE KVAZIKANONICKIH FAKTORA DRUGOG SKUPA,D)
PRINT (MATRIX=M2,TEXT=KORELACIJE KVAZIKANONICKIH FAKTORA DRUGOG SKUPA,D)
PRINT (MATRIX=S2,TEXT=STRUKTURA KVAZIKANONICKIH FAKTORA DRUGOG SKUPA)
PRINT (MATRIX=Q21,TEXT=KROSSTRUKTURA KVAZIKANONICKIH FAKTORA DRUGOG SKUPA)
      DELETE (MATRIX=R2Y2)
*
* BLOK 4.
* KORELACIJE KANONICKIH I KVAZIKANONICKIH VARIJABLI
*
HEADING(TEXT=KORELACIJE KANONICKIH I KVAZIKANONICKIH VARIJABLI,D)
MULT (A=X1,B=S1,M=MKH1)
MULT (A=X2,B=S2,M=MKH2)
MULT (A=X1,B=Q12,M=CK1H2)
MULT (A=X2,B=Q21,M=CK2H1)
PRINT (MATRIX=MKH1,D,TEXT=KORELACIJE KAN I KVAZIKAN VARIJABLI PRVOG SKUPA)
PRINT(MATRIX=MKH2,D,TEXT=KORELACIJE KAN I KVAZIKAN VARIJABLI DRUGOG SKUPA)
PRINT(MATRIX=CK1H2,TEXT=PRVA KANONICKA I DRUGA KVAZIKANONICKA KORELACIJA)
PRINT(MATRIX=CK2H1,TEXT=DRUGA KANONICKA I PRVA KVAZIKANONICKA KORELACIJA)
*
* BLOK 5.
* KONGRUENCE KANONICKIH I KVAZIKANONICKIH FAKTORA
*
HEADING (TEXT=KONGRUENCE KANONICKIH I KVAZIKANONICKIH FAKTORA,D)
MULT (A=X1,B=X1,TB,M=XX1)
MULT (A=X2,B=X2,TB,M=XX2)
MULT (A=F1,B=F1,TB,M=FF1)
MULT (A=F2,B=F2,TB,M=FF2)
MULT (A=S1,TA,B=S1,M=SS1)
MULT (A=S2,TA,B=S2,M=SS2)
MULT (A=X1,B=Y1,M=XY1)
MULT (A=X2,B=Y2,M=XY2)
MULT (A=F1,B=S1,M=FS1)
MULT (A=F2,B=S2,M=FS2)
DIAGMULT (A=XY1,D=XX1,C=-0.5,L,M=CNGXY1)
DIAGMULT (A=XY2,D=XX2,C=-0.5,L,M=CNGXY2)
DIAGMULT (A=FS1,D=FF1,C=-0.5,L,M=BIK1)
DIAGMULT (A=BIK1,D=SS1,C=-0.5,R,M=CNGFS1)
DIAGMULT (A=FS2,D=FF2,C=-0.5,L,M=BIK2)
DIAGMULT (A=BIK2,D=SS2,C=-0.5,R,M=CNGFS2)
PRINT (MATRIX=CNGXY1,TEXT=KONGRUENCE KOEFICIJENATA PRVOG SKUPA)
PRINT (MATRIX=CNGXY2,TEXT=KONGRUENCE KOEFICIJENATA DRUGOG SKUPA)
PRINT (MATRIX=CNGFS1,TEXT=KONGRUENCE STRUKTURA PRVOG SKUPA)
PRINT (MATRIX=CNGFS2,TEXT=KONGRUENCE STRUKTURA DRUGOG SKUPA)
*
* KRAJ PROGRAMA QCCR
*
```

NEKE RELACIJE IZMEDJU KANONIČKE I KVAZIKANONIČKE KORELACIJSKE ANALIZE*

K. Bosnar
F. Prot
K. Momirović

ODJEL ZA INFORMATIKU I STATISTIKU
FAKULTETA ZA FIZIČKU KULTURU

SAŽETAK

Generalni model za analizu relacija dva skupa kvantitativnih varijabli, predložen u radu Momirovića, Dobrićeve i Karamana (1983) uspoređjen je sa standardnim biortogonalnim kanoničkim korelacijskim modelom (Hotelling,1936). Utvrdjene su formalne relacije izmedju dva modela i napisan je program u SS-jeziku (Zakrajšek, Štalec i Momirović, 1974) za usporedbu dviju solucija.

* *Ovaj rad je izvorno napisan na engleskom jeziku i bit će prezentiran na Medjunarodnom simpoziju računarske statistike COMPSTAT'84, Prag, 1984.*

0. UVOD

Model kvazikanoničke korelacijske analize (QCR) koji su predložili Momirović, Dobrić i Karaman (1983) predstavlja opći model za analizu relacija dva skupa kvantitativnih varijabli pod kojim je moguć tretman biortogonalne kanoničke analize (Hotelling, 1936) i analize redundance (Van der Wollenberg, 1977) kao specijalnih slučajeva s posebnim ograničenjima.

Model je zasnovan na maksimizaciji kovarijanci linearnih kompozita dva skupa varijabli, uz uvjet ortonormalnosti transformacijskih matrica, što je jedino postavljeno ograničenje. Kvazikanonička korelacijska analiza je robustnija od uobičajenih procedura i posjeduje poželjna svojstva manje osjetljivosti na broj stupnjeva slobode, na efekte visoke korelacije samo dviju varijabli iz različitih skupova, a takodjer dozvoljava i neregularnost matrica podataka.

Medjutim, širu primjenu kvazikanoničke korelacijske analize, usprkos svim prednostima, može priječiti nesnalaženje istraživača u interpretaciji rezultata. Algoritam i program QCCR kojim se usporedjuju rezultati biortogonalne kanoničke korelacijske analize i kvazikanoničke korelacijske analize, namijenjen je u prvom redu slučajevima kad nacrtom istraživanja nije dana prednost jednom od modela, ali može takodjer pomoći istraživačima naviklim na interpretaciju rezultata standardne Hotellingove procedure da prihvate novu tehniku analize relacija dva skupa varijabli.

1. KANONIČKE RELACIJE

Kanoničke relacije invarijantne su na metriku varijabli. Neka je stoga

$$Z_U = (U - U\,11^T \tfrac{1}{n})\,V^{-1}$$

matrica standardiziranih podataka o obilježjima nekog skupa $E = \{e_i;\ i=1,\dots,n\}$ opisanog nad nekim skupom varijabli $V = \{v_j;\ j=1,\dots,g\}$, gdje je

$$V_U^2 = \text{diag } C_U$$

a

$$C_U = (U^TU - U^T11^T \tfrac{1}{n}\,U)\,\tfrac{1}{n},$$

sa matricom korelacija

$$R_U = Z_U^TZ_U\,\tfrac{1}{n} = V_U^{-1}C_UV_U^{-1}$$

i neka je

$$Z_B = (B - B\,11^T \tfrac{1}{n})V_B^{-1}$$

matrica standardiziranih podataka o obilježjima skupa E opisanog nad nekim drugim skupom varijabli $W = \{w_p;\ p=1,\dots,m\}$, $v \quad w = 0$, gdje je

$$V_B^2 = \text{diag } C_B$$

a

$$C_B = (B^TB - B^T11^T \tfrac{1}{n}\,B)\,\tfrac{1}{n}\ ,$$

sa matricom korelacija

$$R_B = Z_B^TZ_B\,\tfrac{1}{n} = V_B^{-1}C_BV_B^{-1},$$

i neka je

$$R_{UB} = R_{BU}^T = Z_U^TZ_B\,\tfrac{1}{n}$$

matrica kroskorelacija izmedju varijabli iz V i W na skupu E.

Nadjimo linearne kompozite

$$L_{U\ell} = Z_U A_{U\ell} \qquad \ell=1,\dots,q$$

i $\qquad q=min(g_p,m_p)$ *

$$L_{B\ell} - Z_B A_{B\ell} \qquad \ell=1,\dots,q$$

koji zadovoljavaju ovaj skup uvjeta:

(1) $L_{U\ell}^T L_{B\ell} \frac{1}{n} = \rho_\ell = \max \qquad \ell=1,\dots,q$

(2) $\rho_\ell \geq \rho_{\ell+1} \qquad \ell=1,\dots,q-1$

(3) $L_{U\ell}^T L_{U\ell} \frac{1}{n} = 1 \qquad \ell=1,\dots,q$

(4) $L_{U\ell}^T L_{U\ell^*} \frac{1}{n} = 0 \qquad \ell\neq\ell^*$

(5) $L_{B\ell}^T L_{B\ell} \frac{1}{n} = 1 \qquad \ell=1,\dots,q$

(6) $L_{B\ell}^T L_{B\ell^*} \frac{1}{n} = 0 \qquad \ell\neq\ell^*$

(7) $L_{U\ell}^T L_{B\ell^*} \frac{1}{n} = 0 \qquad \ell\neq\ell^*$

Ovo svodi problem na biortogonalni kanonički problem, čije je rješenje definirano skupom karakterističnih jednadžbi

$$(R_{BU}R_U^{-1}R_{UB} - \rho_\ell^2 R_B)\, A_{B\ell} = 0 \qquad \ell=1,\dots,q$$

pri čemu je vektore $A_{U\ell}$ moguće odrediti na temelju relacije

$$A_{U\ell} = R_U^{-1}R_{UB}A_{B\ell} \frac{1}{\rho_\ell} \qquad \ell=1,\dots,q.$$

Značajnost relacija izmedju $L_{U\ell}$ i $L_{B\ell}$, tj. značajnost kanoničkih korelacija izmedju varijabli iz V i W testira se, obično, operacijama

---

* *U analizi kinezioloških procesa vrijedi gotovo uvijek $g > m$ ako je V skup varijabli kojima su definirane neke antropološke, a W skup varijabli kojima su definirane neke kriterijske varijable.*

$$\chi_\ell^2 = -a \log \prod_{\ell}^{q} (1 - \rho_\ell^2) \qquad \ell=1,2,\ldots,q$$

gdje je

$$a = (n - \frac{1}{2}(m + q))$$

jer vrijednosti χ_ℓ^2 imaju χ^2 raspodjelu sa stupnjevima slobode $(g - \ell + 1)*(m - \ell + 1)$. Vrijednosti ρ_ℓ smatraju se značajnim ako je ispunjen uvjet

$$P(\chi_\ell^2 \mid (g_\ell - \ell + 1)*(m - \ell + 1)) \leq \alpha,$$

gdje je α pogreška tipa I izabrana za testiranje hipoteza $\rho_\ell^*=0$.

Zadržane vrijednosti ρ_ℓ organizirajmo u matricu ρ, a njima pridružene vektore $A_{U\ell}$ i $A_{B\ell}$ u matrice A_U i A_B. Koeficijenti u matrici A_U proporcionalni su parcijalnom značenju varijabli iz V za karakteristike skupa E definirane varijablama iz $L_B = (L_{B\ell})$. Koeficijenti u matrici A_B proporcionalni su parcijalnom značenju varijabli iz W za karakteristike skupa E definirane varijablama iz $L_U = (L_{U\ell})$.

Direktna veza izmedju varijabli iz V i W može se utvrditi na temelju matrica faktorskih struktura

$$F_U = R_U A_U = Z_U^T L_U \frac{1}{n}$$

i

$$F_B = R_B A_B = Z_B^T L_B \frac{1}{n}$$

i matrica faktorskih krosstruktura

$$F_{UB} = R_{UB} A_B \frac{1}{n} = Z_U^T L_B \frac{1}{n}$$

i

$$F_{BU} = R_{BU} A_U \frac{1}{n} = Z_B^T L_U \frac{1}{n} .$$

Kanoničke relacije izmedju dva različita skupa karakteristika osnov su za utvrdjivanje zakonitosti u većini znanstvenih disciplina i otuda za predlaganje postupaka s pomoću kojih se mogu postići neki društveno relevantni ciljevi*.

---

* *Kanonički model definirao je H. Hotelling (1935, 1936). Dobar formalni opis metoda nalazi se u Anderson, 1958; Morisson, 1967; Wilks, 1962 i Kendall and Stuart, 1973.*

2. KVAZIKANONIČKE RELACIJE

Kanonički model osjetljiv je na broj stupnjeva slobode, regularnost matrica R_U i R_B i umjetno visoke korelacije u matrici R_{UB}. Zbog toga je relacije izmedju dva skupa varijabli korisno analizirati i kanoničkom analizom kovarijanci (Momirović, Dobrić i Karaman, 1983).

Neka su

$$Q_{U\ell} = Z_U \Gamma_{U\ell} \qquad \ell=1,\dots,q \quad q=min(g,m)$$

i

$$Q_{B\ell} = Z_B \Gamma_{B\ell} \qquad \ell=1,\dots,q$$

linearni kompoziti koji zadovoljavaju uvjete

(1) $$Q_{U\ell}^T Q_{B\ell} \frac{1}{n} = \sigma_\ell = \max \qquad \ell=1,\dots,q$$

(2) $$\sigma_\ell \geq \sigma_{\ell+1} \qquad \ell=1,\dots,q-1$$

(3) $$\Gamma_{U\ell}^T \Gamma_{U\ell} = 1 \qquad \ell=1,\dots,q$$

(4) $$\Gamma_{U\ell}^T \Gamma_{U\ell*} = 0 \qquad \ell\neq\ell*$$

(5) $$\Gamma_{B\ell}^T \Gamma_{B\ell} = 1 \qquad \ell=1,\dots,q$$

(6) $$\Gamma_{B\ell}^T \Gamma_{B\ell*} = 0. \qquad \ell\neq\ell*$$

Ovo svodi problem na spektralnu dekompoziciju matrice R_{UB}

$$R_{UB} = \sum_{\ell=1}^{q} \sigma_\ell \Gamma_{U\ell} \Gamma_{B\ell}^T$$

pa, ako definiramo značajne vrijednosti spektra matrice R_{UB}

kao $\sigma_\ell^2 \geq \sum_{\ell=1}^{q} \sigma_\ell^2 \frac{1}{q}$, i organiziramo značajne koeficijente σ_ℓ u dijagonalnu matricu $\Sigma=(\sigma_\ell)$ i njima pridružene vektore u matrice $\Gamma_U=(\Gamma_{U\ell})$ i $\Gamma_B=(\Gamma_{B\ell})$, skup kvazikanoničkih varijabli koje su izvedene iz V je

$$Q_U = Z_U\Gamma_U$$

a skup kvazikanoničkih varijabli koje su izvedene iz W je

$$Q_B = Z_B\Gamma_B$$

sa matricama kovarijanci

$$W_U = Q_U^TQ_U \frac{1}{n} = \Gamma_U^TR_U\Gamma_U$$

i

$$W_B = Q_B^TQ_B \frac{1}{n} = \Gamma_B^TR_B\Gamma_B.$$

Matrica kroskovarijanci izmedju Q_U i Q_B je, naravno

$$Q_U^TQ_B \frac{1}{n} = \Sigma.$$

Korelacije izmedju varijabli iz Q_U su

$$M_U = D_U^{-1}W_UD_U^{-1}$$

$D_U^2 = \mathrm{diag}\, W_U$, a korelacije izmedju varijabli iz Q_B

$$M_B = D_B^{-1}W_BU_B^{-1}$$

pa su kvazikanoničke korelacije izmedju ovako definiranih linearnih kompozita

$$\rho^* = D_U^{-1}\,\Sigma\, D_B^{-1}.$$

Koeficijenti u ρ^* su mjera povezanosti izmedju varijabli iz V i W pod kvazikanoničkim modelom.

Identifikacija ovih dimenzija moguća je na temelju matrica strukture

$$F_U^* = Z_U^T Q_U D_U^{-1} \frac{1}{n} = R_U \Gamma_U D_U^{-1}$$

i

$$F_B^* = Z_B^T Q_B D_B^{-1} \frac{1}{n} = R_B \Gamma_B D_B^{-1},$$

odnosno matrica sklopa

$$P_U = F_U^* M_U^{-1}$$

i

$$P_B = F_B^* M_B^{-1}.$$

U identifikacijske svrhe mogu dobro poslužiti i kros-strukturalne matrice

$$F_{UB}^* = Z_U^T Q_B D_B^{-1} \frac{1}{n} = R_{UB} \Gamma_B D_B^{-1}$$

i

$$F_{BU}^* = Z_B^T Q_U D_U^{-1} \frac{1}{n} = R_{BU} \Gamma_U D_U^{-1}.$$

Spektralna analiza relacija izmedju dva skupa varijabli je konzervativan, ali siguran osnov za odredjivanje zakonomjernosti i formiranje racionalnih kinezioloških postupaka.

3. ODNOSI IZMEDJU KANONIČKIH I KVAZIKANONIČKIH LINEARNIH KOMPOZITA

Kako su kanonička i kvazikanonička analiza osnov za donošenje zaključaka o prirodi relacija izmedju dva skupa varijabli, važne su i relacije izmedju rezultata dobijenih ovim postupcima.

Kroskorelacije kanoničkih i kvazikanoničkih varijabli kojima su karakterizirane relacije su

$$L_U^T Q_U D_U^{-1} \frac{1}{n} = A_U^T F_U^*$$

i

$$L_B^T Q_B D_B^{-1} \frac{1}{n} = A_B^T F_B^*$$

a kroskorelacije ovih varijabli su

$$L_U^T Q_B D_B^{-1} \frac{1}{n} = A_U^T F_{UB}^*$$

i

$$L_B^T Q_U D_U^{-1} \frac{1}{n} = A_B^T F_{BU}^* .$$

Od izvjesne koristi za usporedbu rezultata dobijenih ovim postupcima je i odredjivanje kosinusa kuteva vektora iz A_U i Γ_U, odnosno A_B i Γ_B.

4. PROGRAM QCCR

Program QCCR napisan je u meta jeziku SS (Zakrajšek, Štalec i Momirović, 1974) u verziji 5.2/M. Podijeljen je u šest sekcija:

0. učitavanje i sredjivanje podataka
1. korelacije unutar skupova varijabli i izmedju dva skupa varijabli
2. kanonička korelacijska analiza
3. kvazikanonička korelacijska analiza
4. korelacije kanoničkih i kvazikanoničkih varijabli
5. kongruence kanoničkih i kvazikanoničkih faktora.

Programom je moguća analiza do 10.000 entiteta opisanih nad do 250 varijabli u pojedinom skupu. Program zahtijeva da podaci budu pripremljeni tako da je broj varijabli prvog učitanog skupa veći ili jednak broju varijabli u drugom skupu.

Program sadrži 109 linija izvršnog koda i 67 linija komentara koji daju sve neophodne informacije o korištenju ovog proizvoda.

Program QCCR je dostupan svim korisnicima računala UNIVAC Sveučilišnog računskog centra u Zagrebu gdje je pohranjen u programskoj biblioteci SRCE*SS-MAKRO.

5. NUMERIČKI PRIMJER

Efikasnost algoritma i programa provjerena je na primjeru skupova antropometrijskih varijabli, namjerno formiranih tako da postoji samo jedna visoka korelacija izmedju dvije varijable iz dva različita skupa.

Prvi skup čine varijable longitudinalne dimenzionalnosti tijela

1. visina (VISINA)
2. dužina noge (DUZNOG)
3. dužina stopala (DUZSTO)

kojima je pridružena varijabla volumena i mase tijela:

3. opseg podlaktice (OPPODL).

Drugi skup čine varijable za ocjenu količine potkožnog masnog tkiva, za koje je poznato iz dosadašnjih istraživanja na uzorcima iste populacije kao i u ovom primjeru, da su bez supstancijalnih korelacija sa mjerama longitudinalne dimenzionalnosti tijela:

1. nabor na pazuhu (NAPAZU)
2. nabor na nadlaktici (NANADL)
3. nabor na trbuhu (NATRBU)

kojima je takodjer dodana jedna varijabla volumena i mase tijela:

4. opseg natkoljenice (OPNATK).

Kroskorelacije varijabli, dobijene na uzorku 737 muška ispitanika, nalaze se u tabeli 1. Rezultati kanoničke i kvazikanoničke korelacijske analize izloženi su u tabeli 2. Vidljivo je da je u ovom primjeru, prva kvazikanonička korelacija bliža procjeni realne povezanosti dva skupa, tj. da je kvazikanonička analiza prikladnija za ovako formuliran problem.

Tabela 1 - KROSKORELACIJE VARIJABLI DVA SKUPA

	NAPAZU	NANADL	NATRBU	OPNATK
VISINA	.07	.10	.05	.32
DUZNOG	.11	.16	.02	.25
DUZSTO	.06	.09	.05	.29
OPPODL	.35	.33	.35	.70

Tabela 2 - KANONIČKI (CW) I KVAZIKANONIČKI (QW) KOEFICIJENTI, KANONIČKI (CF) I KVAZIKANONIČKI FAKTOR, KANONIČKI (CC) I KVAZIKANONIČKI (QC) KROSFAKTOR, KANONIČKA (r_c) I KVAZIKANONIČKA (r_q) KORELACIJA VARIJABLI DVA SKUPA

	CW	QW	CF	QF	CC	QC
VISINA	.32	.31	.46	.76	.33	.22
DUZNOG	-.10	.28	.36	.72	.26	.20
DUZSTO	-.05	.28	.41	.74	.29	.20
OPPODL	.93	.86	.98	.82	.70	.62
NAPAZU	-.03	.36	.48	.81	.34	.29
NANADL	.06	.37	.45	.73	.32	.28
NATRBU	-.06	.32	.48	.73	.34	.25
OPNATK	1.00	.80	.99	.87	.71	.63

$r_c = 0.81$

$r_q = 0.53$

LITERATURA

1. Hotelling, H.:
   Relations between two sets of variates. Biometrika (1936), 28. 321-377.

2. Momirović, K., V. Dobrić and Ž. Karaman:
   Canonical covariance analysis. Proceedings of 5th International symposium "Computer at the University", Cavtat, 1983, 463-473.

3. Van Der Wollenberg, A.L.:
   Redundancy analysis - an alternative model for canonical correlation analysis. Psychometrika (1977), 42, 207-219.

4. Zakrajšek, E., J. Štalec i K. Momirović:
   SS - Programski sistem za multivarijatnu analizu podataka, Zbornik radova 1. Medjunarodnog simpozija "Kompjuter na Sveučilištu", Zagreb, 1974, CB. 1-16.

2.2 ALGORITAM I PROGRAM SDA

Istraživanja u području vrhunskog sporta redovito se vrše na malom broju ispitanika. Razlog tome nisu samo poteškoće prilikom prikupljanja uzorka već i realno mali broj vrhunskih sportaša u nekim sportskim disciplinama, čak i u svjetskim razmjerima.

Podaci sakupljeni na vrhunskim sportašima su, pored toga, još i često nenormalno distribuirani a ponekad i linearno zavisni, što sve onemogućuje primjenu čitavog niza standardnih procedura za analizu podataka, i primjenu klasičnih metoda diskriminativne analize. Problem razlikovanja grupa, medjutim, jedan je od najvažnijih u procesu selekcije i orijentacije u području sporta. Funkcije razlika izmedju vrhunskih sportaša - pripadnika raznih sportskih disciplina su nezamjenjiv podatak za odabir ili upućivanje kandidata u pojedinu sportsku disciplinu.

Da bi se omogućilo utvrdjivanje razlika medju grupama usprkos nepovoljnih karakteristika podataka, definiran je algoritam i napisan program SDA za robustnu diskriminativnu analizu, koji je manje osjetljiv na broj ispitanika i koji dozvoljava korištenje linearno zavisnih i nenormalno distribuiranih varijabli.

```
*
*   NAZIV PROGRAMA
*
*
*                 *** S D A ***
*
*
*
*   AUTORI
*
*           V. DOBRIC
*           K. MOMIROVIC
*
*
*   FUNKCIJA
*
*
*      S(TUPIDNA) D(ISKRIMINATIVNA A(NALIZA) JE MODIFIKACIJA ALGORITMA
*      ROBUSTNE REGRESIJSKE ANALIZE (STALEC I MOMIROVIC, 1984).
*      OSNOV OVE METODE JE ODREDJIVANJE DISKRIMINANTIVNIH FUNKCIJA NA BAZI
*      GLAVNIH KOMPONENATA STANDARDIZIRANIH VEKTORA SREDINA PO GRUPAMA .
*
*
*   METODA I ALGORITAM OPISANI SU U
*
*           DOBRIC, V. I K. MOMIROVIC:
*           AN ALGORITHM AND PROGRAM FOR STUPID DISCRIMINANT ANALYSIS.
*           NEPUBLICIRANI RAD
*
*
*   PRIPREMA PODATAKA
*
*      OVAJ PROGRAM DOZVOLJAVA ANALIZU DO 20 GRUPA SA TOTALOM DO
*      10000 ENTITETA OPISANIH NAD DO 250 KVANTITATIVNIH, NE NUZNO
*      NORMALNO DISTRIBUIRANIH VARIJABLI.
*
*      PROGRAM PRETPOSTAVLJA DA JE MATRICA PODATAKA U DATOTECI DATA UZ PRIPADNOST
*      GRUPI DEFINIRANOJ U 1. POZICIJI IMENA ENTITETA, I DA JE SELEKTORSKA MATRICA
*      (EVENTUALNO PRIPREMLJENA PROGRAMOM BINAR OD V.DOBRIC ILI NEKIM DRUGIM
*      PRIKLADNIM PROGRAMOM) U DATOTECI ATA.
*
*     *UPOZORENJE: U NAREDBI 'HOTELLING', DEFINIRAJ BROJ K
*            DISKRIMINATIVNIH FAKTORA KAO
*            K=MIN(G-1,M)         GDJE JE
*            G=BROJ GRUPA
*            M=BROJ VARIJABLI
*            ZAMJENOM K U CIJELI BROJ STVARNOG
*            BROJA FAKTORA
*
*
*   IZVRSNI DIO PROGRAMA
*
*     BLOK 0.
*     UCITAVANJE I USKLADJIVANJE PODATAKA
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=S D A,T)
HEADING(TEXT=STUPIDNA DISKRIMINATIVNA ANALIZA)
HEADING(TEXT=MODIFCIRANI ALGORITAM STALECA I MOMIROVICA,D)
TEXT(TEXT=              S D A)
INPUT(DATA=DATA)
INPUT(DATA=ATA,SCORE=SS)
CONFORM(IN1=SCORE,IN2=SS, OUT1=B, OUT2=S)
      DELETE(MATRIX=SCORE)
```

```
      DELETE(MATRIX=SS)
*
*     BLOK 1.
*     DISTRIBUCIJE, PARAMETRI, KORELACIJE I
*     PRELIMINARNA ANALIZA VARIJANCE
*
HEADING(TEXT=DISTRIBUCIJE I ANALIZA VARIJANCE,D)
STATISTICS(SCORE=B,V=1,CLASS=9,S)
CORRELATION(SCORE=B)
HEADING(TEXT=KORELACIJE VARIJABLI,D)
PRINT(MATRIX=R,TEXT=INTERKORELACIJE VARIJABLI)
      DELETE (MATRIX=B)
*
*     BLOK 2.
*     STUPIDNE DISKRIMINATIVNE FUNKCIJE
*
HEADING(TEXT=STUPIDNE DISKRIMINATIVNE FUNKCIJE,D)
MULT(A=S,TA,B=S,M=SS)
DIAGMULT(A=S,D=SS,C=-1.0,L,T,M=SSS)
      DELETE(MATRIX=SS)
MULT(A=SSS,B=Z,M=M)
MULT(A=M,TA,B=M,M=G)
DIAGONALISATION(R=G)
*
*  *UPOZORENJE: ZAMIJENI U SLIJEDECOJ NAREDBI
*             HOTELLING(NUM=K,F=HT)
*             SLOVO K SA STVARNIM BROJEM FAKTORA
*
HOTELLING(NUM=K,F=HT)
      DELETE(MATRIX=LAMBDA)
      DELETE(MATRIX=X)
MULT(A=HT,B=HT,TB,M=LAMBDA)
DIAGMULT(A=HT,T,D=LAMBDA,C=-0.5,R,M=XX)
TRANSPOSE(OLD=XX,NEW=X)
PRINT(MATRIX=LAMBDA,TEXT=DISKRIMINATIVNE VRIJEDNOSTI,D)
PRINT(MATRIX=X,T,TEXT=DISKRIMINATIVNI KOEFICIJENTI)
MULT(A=M,B=X,TB,M=Q)
PRINT(MATRIX=Q,TEXT=CENTROIDI GRUPA NA DISKRIMINATIVNIM FUNKCIJAMA,D)
      DELETE(MATRIX=M)
      DELETE(MATRIX=G)
*
*     BLOK 3
*     STUPIDNI DISKRIMINATIVNI FAKTORI
*
HEADING(TEXT=STUPIDNI DISKRIMINATIVNI FAKTORI,D)
MULT(A=R,B=X,TB,M=RX)
MULT(A=X,B=RX,M=W)
SCALE(C=W,R=RHO)
DIAGMULT(A=RX,D=W,C=-0.5,R,M=F)
INVERSION(R=RHO)
MULT(A=F,B=RINV,M=A)
MULT(A=A,B=F,TB,M=RZ)
LINEAR(A=R,B=RZ,CB=-1.0,M=REZ)
PRINT(MATRIX=A,TEXT=SKLOP DISKRIMINATIVNIH FAKTORA,D)
PRINT(MATRIX=RHO,TEXT=INTERKORELACIJE DISKRIMINATIVNIH FAKTORA,D)
PRINT(MATRIX=REZ,TEXT=RESIDUALNE KORELACIJE VARIJABLI,D)
      DELETE(MATRIX=RX)
      DELETE(MATRIX=W)
      DELETE(MATRIX=RINV)
      DELETE(MATRIX=RZ)
*
*     BLOK 4.
*     DISKRIMINATIVNE VARIJABLE I TESTIRANJE HIPOTEZA
*
```

```
HEADING(TEXT=DISTRIBUCIJE DIS VARIJABLI I ANALIZA VARIJANCE.D)
MULT(A=Z,B=X,TB,M=D)
      DELETE(MATRIX=Z)
STATISTICS(SCORE=D'S'CLASS=9,V=1)
*
* KRAJ PROGRAMA SDA
*
HEADING(TEXT=KRAJ SDA)
*
```

ALGORITAM I PROGRAM ZA ROBUSTNU DISKRIMINATIVNU ANALIZU*

V. Dobrić
K. Momirović

SVEUČILIŠNI RAČUNSKI CENTAR

SAŽETAK

U radu je predloženo pojednostavljenje algoritma SDA robustne diskriminativne analize Štaleca i Momirovića (1984). Osnovne karakteristike SDA modela su zadržane i implementirane u kratkom i efikasnom programu, pisanom u SS jeziku, koji uključuje procjenu diskriminativnih varijabli i izračunavanje sklopa, strukture i matrice interkorelacija značajnih diskriminativnih faktora.

** Ovaj rad je izvorno napisan na engleskom jeziku i biti će prezentiran na VIII Bosansko-Hercegovačkom simpoziju iz informatike, Jahorina, 1984.*

1. UVOD

Štalec i Momirović su 1984 predložili vrlo jednostavnu metodu robustne diskriminativne analize, koja se bazira na maksimiziranju varijance aritmetičkih sredina grupa na mutualno ortogonalnim latentnim dimenzijama, definiranim u prostoru standardiziranih aritmetičkih sredina grupa opisanih na eventualno singularnom skupu ne nužno normalno distribuiranih varijabli.

Ova metoda proizvodi korelirane diskriminativne funkcije u prostoru razapetom vektorima entiteta, koju slijedi faktorska analiza sa orthoblique rotacijom diskriminativnih faktora.

Premda je algoritam iz originalnog rada Štaleca i Momirovića do kraja izveden, još nije napisan program kojim bi se testiralo njegovo ponašanje na realnim podacima; ovaj važni dio svake nove metode tek se očekuje.

U ovom radu je predložena modifikacija originalnog algoritma Štaleca i Momirovića i napisan vrlo jednostavan i efikasan program u verziji 5.2 SS jezika (Zakrajšek, Štalec i Momirović, 1974), u kome je izostavljena transformacija u parsimonijsku poziciju i ne maksimizira se varijanca, već disperzija sredina odstupanja izmedju grupa.

2. ALGORITAM

Neka je $B = (b_{ij})$; $i=1,...,n$; $j=1,...,m$ matrica podataka dobivena opisom rezultata n entiteta iz slučajnog uzorka $E = \{e_i;\ i=1,...,n\} \subset P$ iz populacije P nad skupom $V=\{v_j; j=1,...,m\}$ kvantitativnih, ne nužno multivarijatno normalno distribuiranih varijabli. Nadalje, neka je $S=(s_{ik})$; $i=1,...,n$; $k=1,...,g$ selektorska matrica s elementima 0 i 1 koja definira, zavisno od izabranog postupka selekcije, pripadnost svakog entiteta $e_i \in E$ jednom i samo jednom podskupu (ili grupi) E_k od E, pri čemu je $\bigcup_{k=1}^{g} E_k=E$, $E_k \cap E_k = 0$, $k \neq k$. Kovarijance izmedju varijabli V, procijenjene pod modelom maksimalne vjerodostojnosti sadržane su u matrici $C = B^TB - B^T 1 1^T \frac{1}{n} B) \frac{1}{n}$ tipa (m,m), pri čemu je sa 1 označen n-dimenzionalni sumacioni vektor, a njihove varijance u matrici $V = \mathrm{diag}\, C$ tipa (m,m). Standardizirani rezultati ispitanika Z dobijeni su iz matrice B operacijom

$$Z = (B - 11^T \tfrac{1}{n} B)V^{-1/2} \qquad (m,n)$$

Maksimalnom vjerodostojnošću procijenje korelacije varijabli su

$$R = Z^TZ\tfrac{1}{n} = V^{-1/2}CV^{-1/2}.$$

Operacijom

$$(S^TS)^{-1}S^TZ = M \qquad (k,m)$$

odredjena je matrica M standardiziranih aritmetičkih sredina varijabli iz V po grupama E_k, centriranih na zajedničku aritmetičku sredinu 0 i sa matricom disperzije

$$G = M^TM \qquad (m,m).$$

pa je $G\frac{1}{g}$ matrica varijanci odstupanja izmedju grupa ranga $q=\min(g,m-1)$. Asimptotska distribucija varijabli iz M je $N(0,G\frac{1}{n})$.

Stupidne diskriminativne funkcije koje maksimalno separiraju standardizirane sredine grupa dobijjene su sukcesivnim traženjem rezidualnih linearnih kombinacija Q_p, $p=1,\dots,m$

$$Q_p = MX_p$$

tako da

$$Q_p^T Q_p = X_p^T G X_p = \lambda_p \rightarrow \max \qquad p=1,\dots,n$$

uz uvjete

$$X_p^T X_p = 1\ , \quad X_p^T X_r = 0 \quad p\neq r$$

i

$$\lambda_p \geq \lambda_{p+1}.$$

Rješenje problema se na jednostavan način izvodi iz teorije svojstvenih vrijednosti (Green i Carrol, 1976) preko karakteristične jednadžbe

$$GX = X\Lambda$$

pod uvjetom ortogonalnosti transformacijskih vektora iz X. Da bi se omogućila interpretacija u prostoru relevantnih informacija, ne zadržavaju se svi diskriminativni faktori, već se njihov broj odredjuje na temelju prosjeka nenultih svojstvenih vrijednosti matrice G:

$$\ell = \text{num}\ (\lambda_p > \sum_{p=1}^{m} \lambda_p \cdot \frac{1}{q}).$$

Matrica Q tipa (g,ℓ) je matrica projekcija standardiziranih sredina grupa na diskriminativne varijable; nadalje matrica

$$D = ZX \qquad (n,\ell)$$

je matrica projekcija standardiziranih vektora entiteta na diskriminativne varijable.

Očito, diskriminativne varijable iz Q su ortogonalne u M prostoru jer je $Q^TQ = \Lambda$, ali nisu ortogonalne u Z prostoru, jer imaju nenulte kovarijance sadržane u matrici

$$W = X^TRX$$

sa varijancama U u glavnoj dijagonali matrice W, $U = \text{diag}W$, tako da se standardizirane diskriminativne varijable za entitete iz D odredjuju operacijom

$$K = DU^{-1/2}$$

i imaju matricu korelacija $\rho = U^{-1/2}WU^{-1/2}$.

Relacije početnih i diskriminativnih varijabli sadržane su u matrici strukture

$$F = Z^ZK\frac{1}{n} = RXU^{-1/2}$$

a koordinate varijabli u diskriminativnom prostoru sadržane su u matrici sklopa

$$A = F\rho^{-1} = RXW^{-1}U^{1/2}.$$

Matrice F i A su faktorske matrice za R jer je

$$AF^T = RX(X^TRX)^{-1}X^TR$$

za ℓ zadržanih faktora i, kada je rang od G jednak m i kada su zadržani svi diskriminativni faktori, A i F točno reproduciraju R.

3. PROGRAM SDA

SDA je napisan u 5.2 verziji SS jezika i sastoji se od 56 linija izvršnog koda i 90 linija komentara. SDA je podijeljen u slijedeće blokove:

0. učitavanje i sredjivanje podataka
1. distribucije, parametri i korelacije varijabli, te preliminarna analiza varijance
2. procjena stupidnih diskriminativnih funkcija
3. stupidni diskriminativni faktori
4. diskriminativne varijable i testovi hipoteza.

U ovoj verziji SDA dozvoljava obradu do 10.000 entiteta, do 250 varijabli i do 20 grupa. Posljednja restrikcija se može ublažiti na obradu do 250 grupa ako se iz programa izostavi univarijatna analiza varijance.

Pripremljen je numerički primjer i pohranjen, zajedno sa SDA Programom, u SRCE*SS-MAKRO biblioteci SRCA*.

BILJEŠKA:

1. Štalec, J. and K. Momirović:
   On a very simple method for robust discriminant analysis.
   Neobjavljeni rukopis, 1984.

LITERATURA:

1. Green, P.E. and J.D. Carrol:
   Mathematical tools for applied multivariate analysis.
   Academic Press, New York, 1976.

2. Zakrajšek, E., J. Štalec and K. Momirović:
   SS-Programming system for multivariate data analysis. Zbornik radova I medjunarodnog simpozija Kompjuter na Sveučilištu, 1974, C8.1-16.

---

* *Broj štampanih stranica test primjera iznosi 56 stranica (18 varijabli, 5 grupa, 219 entiteta).*

2.3 ALGORITAM I PROGRAM CORAMIN

Klasifikacija i selekcija u sportu najčešće se vrši uz pomoć više različitih podataka o kandidatima na temelju kojih se predvidja rezultat pojedinca u više različitih kriterija - pokazatelja uspjeha u više sportskih disciplina ili više kriterija uspješnosti u jednoj disciplini. Pri tome nije rijedak slučaj da materijalni uvjeti ili kadrovska ekipiranost postavljaju ograničenja, pa, na primjer, u nekoj sportskoj disciplini mogu postojati uvjeti za trening samo za osobe jednog spola, u drugoj disciplini samo za odredjene dobne kategorije, itd.

Optimalna klasifikacija i selekcija uz uvažavanje ograničenja je složen zadatak čijem rješavanju je namijenjen CORAMIN algoritam i program za višekriterijalno odabiranje s konzistentnim linearnim ograničenjima.

```
*
*   NAZIV PROGRAMA
*
*
*                  ***  CORAMIN ***
*
*
*   AUTORI
*
*            K. MOMIROVIC
*            V. DOBRIC
*            Z. KARAMAN
*
*
*   FUNKCIJA
*
*     PROGRAM ZA VISEKRITERIJALNO ODABIRANJE SA KONZISTENTNIM
*     LINEARNIM OGRANIČENJIMA
*
*
*   METODA I ALGORITAM OPISANI SU U
*
*            MOMIROVIC, K., V. DOBRIC I Z. KARAMAN:
*            ALGORITHM AND PROGRAM FOR MULTICRITERIAL SELECTION
*            WITH CONSISTENT LINEAR CONSTRAINTS.INTERNATIONAL CONFERENCE
*            "GENETICS OF PSYCHOMOTOR PROPERTIES IN MAN", JABLONNA, 1983.
*
*
*   PRIPREMA PODATAKA
*
*            KONTROLNI ZAPIS I PODACI ZA OBA SKUPA VARIJABLI NALAZE
*            SE U DATOTECI DATA.
*
*
*   IZVRSNI DIO PROGRAMA
*
*     BLOK 0.
*     NASLOV. UČITAVANJE I KONTROLA PODATAKA
*
OUTPUT(DEVICE=PR)
TEXT(TEXT=             CORAMIN)
HEADING(TEXT=CORAMIN,T)
HEADING(TEXT=VISEKRITERIJALNO ODABIRANJE SA LINEARNIM OGRANIČENJIMA)
INPUT(DATA=DATA,SCORE=B1)
REWIND(FILE=DATA)
INPUT(DATA=DATA,SCORE=B2)
CONFORM(IN1=B1,IN2=B2,OUT1=S1,OUT2=S2)
            DELETE(MATRIX=B1)
            DELETE(MATRIX=B2)
*
*     BLOK 1.
*     DISTRIBUCIJE,PARAMETRI I INTERKORELACIJE
*     PRVOG SKUPA VARIJABLI
*
HEADING(TEXT=DISTRIBUCIJE I PARAMETRI VARIJABLI PRVOG SKUPA,D)
STATISTICS(SCORE=S1,S,CLASS=9,Z=Z1)
CORRELATION(SCORE=S1,R=R11,C=C11)
DIAGONALISATION(R=R11,X=XP,LAMBDA=LP)
HOTELLING(X=XP,LAMBDA=LP,NUM=1,K1=K1P,Z=Z1)
HADMULT(A=K1P,C=0.0,M=JEDP)
MULT(A=JEDP,TA,B=JEDP,M=NP)
DIAG(A=NP,C=-1.0,D=NPINV)
MULT(A=JEDP,TA,B=S1,M=SUM1)
MULT(A=NPINV,B=SUM1,M=MEAN)
```

```
DIAG(A=C11,D=SIGMA,C=0.5)
DUMP(OUTPUT=TEMP,MATRIX=MEAN,F)
DUMP(OUTPUT=TEMP,MATRIX=SIGMA,F)
HEADING(TEXT=INTERKORELACIJE VARIJABLI PRVOG SKUPA,D)
PRINT(MATRIX=R11,TEXT=MATRICA KORELACIJA VARIJABLI PRVOG SKUPA)
              DELETE(MATRIX=S1)
              DELETE(MATRIX=SUM1)
              DELETE(MATRIX=C11)
              DELETE(MATRIX=XP)
              DELETE(MATRIX=LP)
              DELETE(MATRIX=K1P)
              DELETE(MATRIX=JEDP)
              DELETE(MATRIX=NP)
              DELETE(MATRIX=NPINV)
              DELETE(MATRIX=MEAN)
              DELETE(MATRIX=F)
              DELETE(MATRIX=SIGMA)
*
*     BLOK 2.
*     DISTRIBUCIJE, PARAMETRI I INTERKORELACIJE
*     DRUGOG SKUPA VARIJABLI
*
HEADING(TEXT=DISTRIBUCIJE I PARAMETRI VARIJABLI DRUGOG SKUPA,D)
STATISTICS(SCORE=S2,S,CLASS=9,Z=Z2)
              DELETE(MATRIX=S2)
HEADING(TEXT=INTERKORELACIJE VARIJABLI DRUGOG SKUPA,D)
CORRELATION(SCORE=Z2,R=R22)
PRINT(MATRIX=R22,TEXT=MATRICA KORELACIJA VARIJABLI DRUGOG SKUPA)
*
*     BLOK 3.
*     KROSKORELACIJE PRVOG I DRUGOG SKUPA VARIJABLI
*
HEADING(TEXT=KROSKORELACIJE PRVOG I DRUGOG SKUPA VARIJABLI,D)
CROSSCORRELATION(P1=Z1,P2=Z2,R12=R12)
PRINT(MATRIX=R12,TEXT=MATRICA KROSKORELACIJA)
*
*     BLOK 4.
*     FAKTORI REDUNDANCIJE PRVOG SKUPA VARIJABLI
*
HEADING(TEXT=FAKTORI REDUNDANCIJE PRVOG SKUPA VARIJABLI,D)
MULT(A=R12,B=R12,TB,M=B)
DIAGONALISATION(R=B)
HOTELLING,
              DELETE(MATRIX=LAMBDA)
              DELETE(MATRIX=X)
MULT(A=F,B=F,TB,M=LAMBDA)
DIAGMULT(A=F,T,D=LAMBDA,C=-0.5,R,M=X)
              DELETE(MATRIX=B)
              DELETE(MATRIX=F)
MULT(A=R11,B=X,M=R11X)
MULT(A=X,TA,B=R11X,M=W)
DIAGMULT(A=R11X,D=W,C=-0.5,R,M=F)
MULT(A=R12,TA,B=X,M=C)
MULT(A=Z1,B=X,M=Z1X)
DIAGMULT(A=Z1X,D=W,C=-0.5,R,M=L)
DIAGMULT(A=W,D=W,C=-0.5,L,R,M=M)
INVERSION(R=W,RINV=WINV)
MULT(A=R11X,B=WINV,M=R11XWI)
DIAGMULT(A=R11XWI,D=W,C=0.5,R,M=A)
MULT(A=R11XWI,B=R11X,TB,M=R11T)
PRINT(MATRIX=W,TEXT=KOVARIJANCE FAKTORA REDUNDANCIJE PRVOG SKUPA)
PRINT(MATRIX=M,TEXT=KORELACIJE FAKTORA REDUNDANCIJE PRVOG SKUPA)
PRINT(MATRIX=F,TEXT=STRUKTURA FAKTORA REDUNDANCIJE PRVOG SKUPA)
PRINT(MATRIX=A,TEXT=SKLOP FAKTORA REDUNDANCIJE PRVOG SKUPA)
```

```
PRINT(MATRIX=R11T,TEXT=TEORIJSKE KORELACIJE PRVOG SKUPA)
PRINT(MATRIX=C,TEXT=KROSFAKTORI REDUNDANCIJE DRUGOG SKUPA)
*
*       BLOK 5.
*       PROMAX TRANSFORMACIJA FAKTORA REDUNDANCIJE
*       PRVOG SKUPA
*
HEADING(TEXT=PROMAX TRANSFORMACIJA FAKTORA REDUNDANCIJE PRVOG SKUPA,D)
INPUT(DATA=TARGET,SCORE=TT)
MULT(A=X,B=X,TB,M=XXT)
DIAGMULT(A=TT,D=XXT,L,C=0.5,M=T)
MULT(A=X,TA,B=T,M=Q)
MULT(A=X,B=Q,M=Y)
MULT(A=Z1,B=Y,M=KSTAR)
MULT(A=R12,TA,B=Y,M=CSTAR)
DIAGMULT(A=Q,T,R,D=LAMBDA,M=NISTA)
MULT(A=NISTA,B=Q,M=CSTCST)
              DELETE(MATRIX=NISTA)
DUMP(MATRIX=Y,OUTPUT=TEMP,F)
MULT(A=Q,TA,B=W,M=NISTA)
MULT(A=NISTA,B=Q,M=WSTAR)
              DELETE(MATRIX=NISTA)
DIAGMULT(A=KSTAR,R,D=WSTAR,C=-0.5,M=LSTAR)
DIAGMULT(A=WSTAR,L,R,D=WSTAR,C=-0.5,M=MSTAR)
MULT(A=R11,B=Y,M=NISTA)
DIAGMULT(A=NISTA,R,D=WSTAR,C=-0.5,M=FSTAR)
TRANSPOSE(OLD=Q,NEW=QT)
GAUSSJORDAN(M=QT,MINV=QTINV)
DIAGMULT(A=QTINV,R,D=WSTAR,C=0.5,M=NESTO)
MULT(A=R11XWI,B=NESTO,M=ASTAR)
PRINT(MATRIX=CSTAR,TEXT=TRANSFORMIRANI KROSFAKTORI REDUNDANCIJE)
PRINT(MATRIX=CSTCST,TEXT=KVADRIRANE KOVARIJANCE KROSFAKTORA REDUNDANCIJE)
PRINT(MATRIX=WSTAR ,TEXT=KOVARIJANCE TRANSFORMIRANIH FAKTORA REDUNDANCIJE)
PRINT(MATRIX=MSTAR ,TEXT=KORELACIJE TRANSFORMIRANIH FAKTORA REDUNDANCIJE)
PRINT(MATRIX=FSTAR ,TEXT=STRUKTURA TRANSFORMIRANIH FAKTORA REDUNDANCIJE)
PRINT(MATRIX=ASTAR ,TEXT=SKLOP TRANSFORMIRANIH FAKTORA REDUNDANCIJE)
*
*       BLOK 6.
*       ANALIZA REZIDUALA DRUGOG SKUPA VARIJABLI
*
HEADING(TEXT=REZIDUALI U DRUGOM SKUPU VARIJABLI)
DIAGMULT(A=CSTAR,D=WSTAR,R,C=-0.5,M=G)
MULT(A=C,B=WINV,M=PAS)
MULT(A=PAS,B=NESTO,M=P)
MULT(A=P,B=G,TB,M=PGT)
LINEAR(A=R22,B=PGT,CB=-1.0,M=N)
PRINT(MATRIX=G,TEXT=KROSSKLOP VARIJABLI DRUGOG SKUPA)
PRINT(MATRIX=P,TEXT=KROSSTRUKTURA VARIJABLI DRUGOG SKUPA)
PRINT(MATRIX=N,TEXT= KORELACIJE REZIDUALA DRUGOG SKUPA VARIJABLI)
*
HEADING(TEXT= KRAJ PROGRAMA CORAMIN ,T,D)
*
```

ALGORITAM I PROGRAM ZA VIŠEKRITERIJALNO ODABIRANJE S KONZISTENTNIM LINEARNIM OGRANIČENJIMA*

Momirović, K., V. Dobrić i Ž. Karaman

1. UVOD

Baza za višekriterijalno odabiranje entiteta iz skupa entiteta, na osnovu skupa kvantitativnih, multivarijatno normalno distribuiranih, genetički generiranih varijabli, su obično rezultati niza regresijskih analiza pod vidom najmanjih kvadrata, primijenjenih za svaku kriterijsku varijablu posebno (Gauss, 1809; Markov, 1900; Wilks, 1962), ili rezultati standardne biparcijalne kanoničke korelacijske analize (Hotelling, 1935; 1936; Anderson, 1938); medjutim, u nekim slučajevima se mogu u tu svrhu upotrijebiti i rezultati nedavno razvijenih metoda analize redundancije (Van den Wollenberg, 1977; Johansson, 1981).

Da bi se izbjegli neki nedostaci multivarijatne regresijske analize pod vidom najmanjih kvadrata, kanoničke korelacijske analize i analize redundancije (osjetljivost na broj stupnjeva slobode, osjetljivost na neregularnost matrice kovarijanci prediktora, osjetljivost na ekstremne rezultate i lažne efekte visoke korelacije izmedju dvije varijable koje pripadaju različitim skupovima, nedavno je predložena klasa metoda baziranih na maksimiziranju kovarijance, a ne korelacija, linearnih kombinacija prediktorskih i kriterijskih varijabli. Kanonička analiza kovarijance (Momirović, Dobrić i Karaman, 1983) je model za analizu odnosa izmedju dva skupa kvantitativnih multivarijatno normalno distribuiranih varijabli uz nešto oslabljene uvjete. Osnov modela je maksimiziranje kovarijanci linearnih kombinacija dvaju skupova varijabli pod uvjetom ortonormalnosti transformacijskih matrica. Specijalni slučaj ove metode

---

* *Ovaj rad izvorno je napisan na engleskom jeziku i izložen je na Medjunarodnoj konferenciji "Genetics of psychomotor properties in man", Jabłonna, 1983.*

je SRA regresijska analiza (Štalec i Momirović, 1983) koja se bazira na maksimiziranju kovarijanci linearne kombinacije prediktorskih varijabli i kriterijske varijable pod uvjetom da je norma vektora regresijskih koeficijenata jednaka 1. Praktički isti pristup daje novu varijantu analize redundancije (Prot, Bosnar i Momirović, 1983); predložena procedura se bazira na maksimiziranju sume kvadriranih kovarijanci varijabli redundancije prediktorskog skupa sa kriterijskim skupom varijabli, pod uvjetom da je transformacijska matrica koja definira varijable redundancije ortonormalna. U ovom radu je predložen jedan drugi model asimetrične analize redundancije. Osnova modela je sukcesivno maksimiziranje sume kvadrata kovarijanci kriterijskih varijabli sa linearnim kombinacijama prediktorskih varijabli pod uvjetom ortogonalnosti transformacijskih vektora normiranih na 1, nakon čega slijedi transformacija ovako dobijene transformacijske matrice da bi se zadovoljili dodatni linearni uvjeti. Opisani algoritam je implementiran u SS jeziku na računalu UNIVAC 1100/42; da bi se ispitalo ponašanje ove metode pripremljen je numerički primjer u kojem se analiziraju odnosi kognitivnih i konativnih latentnih dimenzija sa skupom mjera motoričkih sposobnosti.

2. MODEL I ALGORITAM

Neka je $P = \{p_j;\ j=1,\dots,m\}$ skup prediktorskih varijabli, $K = \{k_\ell;\ \ell=1,\dots,q\}$ skup kriterijskih varijabli i $E = \{e_i;\ i=1,\dots,n\}$ skup entiteta dobijenih nekim slučajnim postupkom iz populacije Π. Pretpostavimo da su varijable iz P i K tako standardizirane ili prethodno transformirane nekim monotonim transformacijskim postupkom da matrice $Z_1 = (z_{ij})$; $i=1,\dots,n$; $j=1,\dots,m$ i $Z_2 = (z'_{i\ell})$; $i=1,\dots,n$; $\ell=1,\dots,q$ dobivene opisom skupa E nad P i K imaju multivarijatne normalne distribucije $f(P)=N(0,\tilde{R}_{11})$ i $\bar{f}(K)=N(0,\tilde{R}_{22})$, diag $\tilde{R}_{11}=I$, diag $\tilde{R}_{22}=I$. Naravno, maksimalnom vjerodostojnošću procijenjene korelacijske matrice su $R_{11}=Z_1^TZ_1\frac{1}{n}$ i $R_{22}=Z_2^TZ_2\frac{1}{n}$, a maksimalnom vjerodostojnošću procijenjene kroskorelacije izmedju varijabli iz *P* i *K* nad *E* su sadržane u matrici $R_{12} = Z_1^T Z_2 \frac{1}{n}$.

U prvom dijelu algoritam je definiran jednom varijantom Van den Wollenbergove analize redundancije sa oslabljenim uvjetima (Prot, Bosnar i Momirović, 1983). Traže se takve linearne kombinacije K_p prediktorskih varijabli

$$K_p = Z_1 X_p \; , \qquad X_p \neq 0, \qquad\qquad p=1,\dots,m$$

koje će sukcesivno maksimizirati sumu kvadriranih kovarijanci

$$C_p = Z_2^T K_p \frac{1}{n}$$

$$C_p^T C_p \rightarrow \max \qquad\qquad p=1,\dots,m$$

izmedju kriterijskih varijabli Z_{2p} i varijabli redundancije K_p, $p=1,\dots,m$ prediktorskih varijabli, pod uvjetom ortonormalnosti transformacijskih vektora.

Rješenje ovog problema je dobro poznato iz teorije svojstvenih vrijednosti matrica i može se sukcesivno izvesti u p koraka, $p=1,\dots,m$. U prvom koraku uvodi se Lagrangeov multiplikator λ_1 i maksimizira se funkcija f_1

$$f_1 = C_1^T C_1 = X_1^T R_{12} R_{12}^T X_1 - \lambda_1 (X_1^T X_1 - 1).$$

Izjednačavanjem svih parcijalnih derivacija od f_1 sa 0, dobijene su jednadžbe

$$\partial f_1 / \partial X_1 = 2R_{12} R_{12}^T X_1 - 2\lambda_1 X_1 = 0 \qquad (1)$$

tj.

$$R_{12} R_{12}^T X_1 = \lambda_1 X_1 \qquad (1')$$

$$\partial t / \partial \lambda_1 = X_1^T X_1 - 1 = 0 \qquad (2)$$

tj.

$$X_1^T X_1 = 1 \qquad (2').$$

U drugom koraku moraju biti zadovoljeni uvjeti $X_2^T X_1 = 0$ i $X_2^T X_2 = 1$, pa je funkcija koju treba maksimizirati oblika

$$f_2 = c_2^T c_2 = X_2^T R_{12} R_{12}^T X_2 - \lambda_2 (X_2^T X_2 - 1) - \eta_1 (X_2^T X_1).$$

Sada je

$$\partial f_2 / \partial X_2 = 2 R_{12} R_{12}^T X_2 - 2\lambda_2 X_2 - \eta_1 X_1 = 0 \qquad (3)$$

$$\partial f_2 / \partial \lambda_2 = X_2^T X_2 - 1 = 0 \qquad (4)$$

$$\partial f_2 / \partial \eta_1 = X_2^T X_1 = 0. \qquad (5)$$

Pomoži li se (3) s lijeva s X_1^T, dobija se

$$2 X_1^T R_{12} R_{12}^T X_2 = \eta_1 .$$

Uvrštavanjem jednakosti (1´) i (5) u gornju jednadžbu, $\eta_1 = 0$, pa je rezultat karakteristična jednadžba

$$R_{12} R_{12}^T X_2 = \lambda_2 X_2$$

ili, u p-tom koraku

$$R_{12} R_{12}^T X_p = \lambda_p X_p$$

uz uvjete

$$X_p^T X_p = 1, \quad X_p^T X_q = 0 \qquad p \neq q; \qquad p, q = 1, \ldots, m.$$

Broj k zadržanih svojstvenih vektora se odredjuje na osnovu nenultih svojstvenih vrijednosti λ_p. Ako je $r = \mathrm{num}(\lambda_p > 0)$,

$$k = \mathrm{num}\ (\lambda_p \geq \frac{1}{r} \sum_{p'=1}^{m} \lambda_{p'}).$$

Rješenje problema maksimiziranja sume kvadriranih kovarijanci c_p se može matrično pisati u obliku

$$C^T C = X^T R_{12} R_{12}^T X = \Lambda$$

uz

$$C = Z_2^T K \frac{1}{n},$$

$$K = Z_1 X$$

gdje je $X=(x_{ij})$; $i=1,\ldots,m$; $j=1,\ldots,k$ matrica zadržanih svojstvenih vektora sa svojstvom da je $X^T X = I$, $XX^T \neq I$ a K je matrica nestandardiziranih varijabli redundancija. Kako vektori u K nisu mutualno ortogonalni, interne kovarijance varijabli u K su

$$W = K^T K \frac{1}{n} = X^T R_{11} X$$

sa varijancama

$$S^2 = \text{diag } W.$$

Vektori iz K se mogu normirati na jediničnu varijancu operacijom

$$L = KS^{-1} = Z_1 X S^{-1}$$

pa su korelacije ovako normiranih varijabli jednake

$$M = L^T L \frac{1}{n} = S^{-1} X^T R_{11} X S^{-1}.$$

Kako bi se olakšala interpretacija rezultata, vrlo je važno naći matricu strukture faktora redundancije prediktorskog skupa

$$F = Z_1^T L \frac{1}{n} = R_{11} X S^{-1}.$$

Faktori redundancije nisu ortogonalni jer

$$F^TF = S^{-1}X^TR_{11}^2XS^{-1} \neq \text{diag},$$

pa se matrica sklopa može izvesti kao

$$A = FM^{-1} = R_{11}X(X^TR_{11}X)^{-1}S.$$

Očito, F i A su faktorske matrice za R_{11} jer je

$$AF^T = R_{11}X(X^TR_{11}X)^{-1}X^TR_{11} = R_{11}^*$$

što predstavlja opći Guttmanov izraz (1949; 1957) za "faktorizaciju matrice", u smislu odredjivanja k komponenti korelacijske matrice R_{11}.

U drugoj fazi algoritam traži takvu linearnu transformaciju Q matrice X kojom se postiže na neki način najbolje slaganje transformacijske matrice X i neke matrice cilja T, prethodno specificirane na osnovu neke hipoteze o matrici sklopa ili strukture prediktorskih varijabli.

Neka je $T = (t_{ij})$; $i=1,\ldots,m$; $j=1,\ldots,k$ matrica cilja, jednostavne strukture, istog tipa kao X, sa elementima koji opisuju hipotetsku strukturu ili sklop prediktorskih varijabli i neka T ima svojstvo da je $\text{diag}(TT^T) = \text{diag}(XX^T)$. Aproksimativni model se u matričnoj notaciji može napisati u obliku

$$T = XQ + E \tag{6}$$

sa takvom matricom Q da je

$$\text{tr}(E^TE) = \min \tag{7}$$

gdje je E matrica pogreške. Bez drugih uvjeta na Q rješenje je vrlo jednostavno; izjednačivši parcijalne derivacije od (7) obzirom na T s nulom,

$$\frac{\partial}{\partial T}(\mathrm{tr}(E^TE)) = -2X^TT + 2X^TXQ = 0,$$

slijedi da je

$$X^TXQ = X^TT$$

odnosno

$$Q = X^TT$$

jer je $X^TX=I$. Rješenje dobijeno ovakvom aproksimacijom je dobro poznato iz literature kao bezuvjetna baza PROMAX rotacija (Mulaik, 1972). Očito da transformacijska matrica Q ima vrlo jednostavan oblik.

Slijedeći korak algoritma je ponovno razmatranje prve faze s matricom

$$Y = XQ = XX^TT$$

umjesto matrice X; to znači da algoritam izvodi nove nestandardizirane linearne kombinacije prediktorskih varijabli

$$K^* = Z_1Y = KQ$$

koje imaju kovarijance s kriterijskim varijablama jednake

$$C^* = Z_2^TK^*\frac{1}{n} = R_{12}^TY = R_{12}^TXQ$$

pri čemu je suma kvadrata kovarijanci jednaka

$$C^{*T}C^* = Q^T\Lambda Q.$$

Kovarijance ovako dobivenih varijabli redundancije su

$$W^* = K^{*T}K\frac{1}{n} = Q^TWQ$$

s varijancama

$$S^{*2} = \text{diag}W^* = (s_p^{*2}) \qquad p=1,\ldots,k$$

tako da su nove standardizirane varijable redundancije jednake

$$L^* = K^*S^{*-1} = Z_1YS^{*-1} = KQS^{*-1}$$

sa interkorelacijama dobijenim operacijom

$$M^* = L^{*T}L^*\frac{1}{n} = S^{*-1}W^*S^{*-1}.$$

Ovi rezultati su očito vrlo slični rezultatima dobijenim u prvoj fazi algoritma jer su odredjeni transformacijom rezultata iz prve faze putem transformacijske matrice Q.

Još jednom, za interpretaciju je vrlo korisno ispitati matricu strukture kosih faktora redundancije

$$F^* = Z_1^TL^*\frac{1}{n} = R_{11}YS^{*-1} = R_{11}XQS^{*-1}$$

kao i njihov sklop

$$A^* = F^*M^{*-1} = R_{11}X(X^TR_{11}X)^{-1}Q^{-T}S^*.$$

Interesantno je da A^* i F^* egzaktno reproduciraju R_{11}^*, što znači da nisu samo A i F, već i A^* i F^* faktorske matrice za R_{11}:

$$A^*F^{*T} = R_{11}X(X^TR_{11}X)^{-1}X^TR_{11} = R_{11}^*.$$

Slijedeći korak algoritma je višekriterijalno odabiranje entiteta na bazi promax rotiranih varijabli redundancije

$$K_p^* = Z_1 Y_p \qquad p=1,\dots,k$$

čije su varijance (s_p^{*2}), $p=1,\dots,k$ i koje su normalno distribuirane, jer su dobivene kao linearne kombinacije inicijalnih normalno distribuiranih varijabli:

$$g(K_p^*) = (2\Pi)^{-1/2}\ s_p^{*-1}\ e^{-1/2\ k_p^{*2}/s_p^{*2}},$$

ili na bazi standardiziranih varijabli redundancije

$$L_p^* = K_p^* s_p^{*-1} = (\ell_i)_p \qquad i=1,\dots,n \quad p=1,\dots,k$$

sa funkcijama gustoće

$$\xi(L_p^*) = (2\Pi)^{-1/2} e^{-1/2} L_p^{*2}.$$

Ako se za svaki L_p^*, $p=1,\dots,k$ definiraju kriterijske granice $\tilde{\ell}_p$, vjerojatnost da se neki L_p^* nalazi u intervalu $(-\infty,\tilde{\ell}_p)$ je

$$p(L_p^* \leq \tilde{\ell}_p) = (2\Pi)^{-1/2} \int_{\infty}^{\tilde{\ell}_p} e^{-1/2\ L_p^{*2}}\ dL_p^* .$$

Vektor kriterijskih granica je

$$\tilde{L}^T = (\ell_1,\dots,\ell_k)$$

i funkcija odabiranja σ_i za pojedini objekt e_i. $i=1,\dots,n$ se definira kao

$$\sigma_i = \begin{cases} 1, & \text{za } \ell_{pi}^* > \tilde{\ell}_p\ \forall p \\ 0, & \text{inače} \end{cases} \qquad i=1,\dots,n.$$

U matričnoj notaciji, funkcija gustoće za standardizirane varijable redundancije iz matrice L^* je

$$h(L^*) = \det(M^*)^{1/2} (2\Pi)^{-1/2\,k} e^{-1/2\,\xi}$$

gdje je $\det(M^*)$ determinanta matrice interkorelacija M^* i

$$\xi = L^{*T}M^{*-1}L^*;$$

vjerojatnost za definirane granice u $\tilde{L}$ je

$$p(L^* \leq \tilde{L}) = \det(M^*)^{1/2} (2\Pi)^{-1/2\,k} \int_{-\infty}^{\tilde{\ell}_1} \dots \int_{-\infty}^{\tilde{\ell}_k} e^{-1/2\,\xi_j}\, dL_1^* \dots dL_k^*$$

$$\xi_j = L_j^{*T}M^{*-1}L_j^* \qquad j=1,\dots,k\,.$$

Metoda višestrukog odsijecanja je razumni postupak u većini slučajeva kada je potrebno otkriti mali broj entiteta sposobnih za obavljanje neke složene psihomotorne aktivnosti.

U posljednjem koraku algoritam definira matrice strukture i sklopa krosfaktora redundancije kriterijskih varijabli:

$$G = Z_2^T L^* \frac{1}{n} = R_{12}^T Y S^{*-1}$$

i

$$P = GM^{*-1} = R_{12}^T X (X^T R_{11} X)^{-1} Q^{-T} S^*\,.$$

Matrice P i G su faktorske matrice za R_{22} jer je

$$PG^T = R_{12}^T X (X^T R_{11} X)^{-1} X^T R_{12}$$

projekcija drugog skupa varijabli u prostor razapet zadržanim faktorima redundancije prvog skupa varijabli.

Algoritam odredjuje i matricu pogreške da bi se ustanovilo koja se količina informacija iz R_{22} gubi takvim projiciranjem

$$N = R_{22} - PG^T = R_{22} - R_{12}^T X(X^T R_{11} X)^{-1} X^T R_{12} .$$

Ako je k=m, $N = R_{22} - R_{12}^T R_{11}^{-1} R_{12}$, tj. poznata matrica varijanci - - kovarijanci reziduala iz multivarijatne regresijske analize, jer je u ovom posljednjem slučaju $XX^T = I$. Kako je Q nesingularna transformacijska matrica, funkcije pogreške su nezavisne od transformacije inducirane matricom cilja T.

3. PROGRAM CORAMIN

Algoritam za višekriterijalno odabiranje je implementiran kao makro program u verziji 5.2/M SS jezika pod imenom CORAMIN. Dokumentirani program je podijeljen u slijedeće blokove:

(0) Učitavanje prediktorske i kriterijske matrice podataka, učitavanje matrice cilja, sortiranje i uskladjivanje ulaznih podataka;

(1) Standardizacija prediktorskih varijabli i računanje matrice interkorelacija prediktora;

(2) Standardizacija kriterijskih varijabli i računanje matrice interkorelacija kriterija;

(3) Računanje matrice kroskorelacija prediktora i kriterija;

(4) Analiza redundancije bez uvjeta ortogonalnosti i sa redukcijom podataka;

(5) Računanje matrice kovarijanci pogreške;

(6) Neuvjetovana PROMAX transformacija rješenja dobijenog analizom redundancije;

(7) Pohranjivanje transformacijske matrice koja predstavlja bazu postupka selekcije.

Ovaj program dozvoljava analizu do 10.000 entiteta i do 250 prediktorskih i do 250 kriterijskih varijabli, s ograničenjem da maksimalni broj faktora redundancije nije veći od 110.

Dodatni kratki program pisan takodjer u SS jeziku, vrši operacije selekcije. Ovaj se program sastoji od slijedećih blokova:

(0) Učitavanje podataka o entitetima koji će biti podvrgnuti proceduri selekcije, učitavanje transformacijske matrice i vektora kritičnih vrijednosti;

(1) Operacija selekcije i štampanje rezultata izabranih entiteta.

4. NUMERIČKI PRIMJER

Skup prediktorskih varijabli sastavljen je od rezultata 200 muških ispitanika, aktivno uključenih u razne kineziološke aktivnosti, u slijedećim kognitivnim i konativnim dimenzijama:

(1) efikasnost perceptivnog procesiranja
(2) efikasnost serijalnog procesiranja
(3) efikasnost paralelnog procesiranja
(4) efikasnost regulacije i kontrole reakcija obrane
(5) efikasnost regulacije i kontrole reakcija napada
(6) efikasnost regulacije organskih funkcija
(7) efikasnost homeostatičke regulacije
(8) efikasnost integracije i koordinacije regulativnih funkcija
(9) efikasnost regulacije nivoa aktiviteta.

Na istom uzorku entiteta definiran je skup kriterijskih varijabli i izneseni su rezultati u slijedećim latentnim psihomotornim dimenzijama:

(1) koordinaciji
(2) ritmu
(3) ravnoteži
(4) frekvenciji pokreta
(5) brzini jednostavnih pokreta
(6) preciznosti
(7) fleksibilnosti
(8) sili
(9) eksplozivnoj snazi
(10) repetitivnoj snazi
(11) izdržljivosti.

Tabela 1 prikazuje sklop faktora redundancije skupa kognitivnih i konativnih varijabli, a tabela 2 korelacije dobivenih latentnih dimenzija. Algoritam je proizveo tri lako interpretabilna faktora. Prvi od njih je generalni faktor kognitivnih sposobnosti, drugi je generalni faktor regulacije i kontrole konativnih funkcija a treći je faktor mjera efikasnosti motivacijskih procesa.

Tabela 3 prikazuje korelacije motoričkih sposobnosti i faktora redundancije izvedenih iz skupa kognitivnih i konativnih faktora. Ritam, pa zatim koordinacija i frekvencija pokreta imaju najveće korelacije sa generalnim faktorom kognitivnih sposobnosti. Generalni faktor regulacije i kontrole konativnih funkcija ima signifikantne korelacije sa koordinacijom, ravnotežom i preciznošću. Treći faktor redundancije nema značajnih korelacija sa bilo kojom motoričkom sposobnošću.

Prediktorski sistem, kojim dominira generalni kognitivni faktor, omogućuje efikasnu detekciju entiteta s iznad prosječnim ritmom, koordinacijom, ravnotežom, frekvencijom pokreta i preciznošću, kao što se vidi iz tabele 3. To je, očito, posljedica činjenice da u varijanci kompleksnih motoričkih sposobnosti varijanca kognitivnih i regulatornih funkcija igra značajnu ulogu.

Tabela 1 - SKLOP PRVOG SKUPA FAKTORA REDUNDANCIJE

	FAC 1	FAC 2	FAC 3
ALPHA	.13	(.71)	.28
SIGMA	-.05	(.79)	-.01
HI	.14	(.94)	-.33
DELTA	.01	(.66)	.37
ETA	.12	(.78)	.08
E	-.45	-.08	(.45)
INPUT	(.88)	.02	-.41
PARAL	(.87)	-.10	.26
SERIJ	(.81)	-.22	.36

Tabela 2 - KORELACIJE PRVOG SKUPA FAKTORA REDUNDANCIJE

	FAC 1	FAC 2	FAC 3
FAC 1	1.00	.07	.24
FAC 2	.07	1.00	.48
FAC 3	.24	.48	1.00

Tabela 3 - KROSFAKTORI DRUGOG SKUPA

	FAC 1	FAC 2	FAC 3
KOORDI	(.32)	(.29)	-.09
RITAM	(.74)	-.08	.11
BALANS	.12	(.25)	.00
BRZFRQ	(.24)	.03	-.14
BRZJEP	.14	-.16	-.14
PRECIZ	-.05	(.26)	.02
FLEKSI	.18	.01	.00
DIMSIL	-.04	.03	.00
EKSPLO	.04	-.04	-.07
REPSTA	-.14	.16	-.05
IZDRZL	-.08	.19	.08
ρ	.64	.24	.06

LITERATURA

1. Anderson, T.W.:
An introduction to multivariate statistical analysis. Wiley, New York, 1958.

2. Gauss, K.F.:
Theoria motus corporum coelestium in sectimibus conicis solem ambientum. Perthes und Besser, Hamburg, 1809.

3. Hotelling, H.:
The most predictable criterion. Journal of Educational Psychology (1935), 26, 139-142.

4. Hotelling, H.:
Relations between two sets of variates. Biometrika (1936), 28, 321-377.

5. Johansson, I.K.:
An extension of Wollenberg´s redundancy analysis. Psychometrika (1981), 46, 93-103.

6. Markov, A.A.:
Wahrscheinlichkeitsneckung (Deutsche Übersetzung des Russisches Original; Erste Auflage, 1900). Teubner, Leipzig, 1912.

7. Momirović, K., V. Dobrić and Ž. Karaman:
Canonical covariance analysis.Proceedings of 5° International Symposium "Computer at the University" (1983), 5, 1, 1-11.

8. Prot, F., K. Bosnar and K. Momirović:
An algorithm for asymetric redundancy analysis without orthogonality constraints. Proceedings of 5° International symposium "Computer at the University" (1983), 5, 2, 1-9.

9. Van den Wollenberg, A.L.:
Redundancy analysis - An alternative model for canonical correlation analysis. Psychometrika (1977), 42, 207-219.

10. Wilks, S.S.:
Mathematical statistics. Wiley, New York, 1962.

2.4 ALGORITAM I PROGRAM CAOS

Jednadžbe specifikacije su osnov ne samo operacija izbora, već i kontrole sportske forme tokom trenažnog procesa, jer definiraju sposobnosti i osobine koje su od značaja za postizanje sportskih rezultata. Ako se rezultat u sportu može definirati samo jednom varijablom, ili se na razuman način može svesti na samo jednu varijablu, jednadžbu specifikacije je moguće izvesti u okviru regresijskog modela.

Uzorci kvalitetnih takmičara pojedinih sportova su skromnih efektiva, pa formiranje jednadžbi specifikacije pod standardnim LSR modelom (najmanjih kvadrata) ne može dati dovoljno stabilne rezultate za donošenje pouzdanih sudova.

Model SRA (model maksimizacije kovarijanci) Momirovića i Štaleca (1983) znatno je manje osjetljiv na efektive uzoraka.

Program CAOS omogućuje analizu pod oba modela (LSR, SRA), a usporedbom dobivenih rezultata znatno doprinosi sigurnosti zaključivanja o parametrima traženih jednadžbi specifikacije.

```
*
*   NAZIV PROGRAMA
*
*                  *** CAOS ***
*
*
*   AUTORI
*
*
*          V. DOBRIC
*          J. STALEC
*          K. MOMIROVIC
*
*
*   FUNKCIJA
*
*
*     CAOS ANALIZIRA SKUP MULTIVARIJATNO NORMALNO DISTRIBUIRANIH KRITERIJSKIH
*     VARIJABLI REGRESIRANIH NA SKUP MULTIVARIJATNO NORMALNO
*     DISTRIBUIRANIH PREDIKTORSKIH VARIJABLI UOBICAJENOM
*     REGRESIJSKOM ANALIZOM NAJMANJIM KVADRATIMA, UZ MINIMIZACIJU
*     VARIJANCE, I STUPIDNOM REGRESIJSKOM ANALIZOM (STALEC I MOMIROVIC,
      1983) I USPOREDJUJE REZULTATE TIH DVIJU METODA
*
*
*   METODA I ALGORITAM OPISANI SU U
*
*          DOBRIC, V., J. STALEC AND K. MOMIROVIC:
*          NOTE ON SOME RELATIONSHIPS BETWEEN LEAST SQUARES AND
*          STUPID REGRESSION ANALYSIS, 6TH INTERNATIONAL SYMPOSIUM ON
*          COMPUTATIONAL STATISTICS, PRAGUE, 1984.
*
*
*   PRIPREMA PODATAKA
*
*
*     PRETPOSTAVJA SE DA JE PREDIKTORSKI SKUP U DATOTECI  D A T A,  A
*     KRITERIJSKI SKUP U DATOTECI  A T A
*
*
*   IZVRSNI DIO PROGRAMA

*     BLOK 0.
*     UCITAVANJE I USKLADJIVANJE PODATAKA
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=C A O S.T)
HEADING(TEXT=KOMPARATIVNA ANALIZA STUPIDNE I REGRESIJE NAJMANJIM KVADRATIMA)
HEADING(TEXT=VERZIJA 1.D)
TEXT(TEXT=            CAOS)
INPUT(DATA=DATA,SCORE=PRED)
INPUT(DATA=ATA,SCORE=CRIT)
CONFORM(IN1=PRED,IN2=CRIT,OUT1=P,OUT2=C)
      DELETE(MATRIX=PRED)
      DELETE(MATRIX=CRIT)
*
*     BLOK 1.
*     DISTRIBUCIJE, PARAMETRI I KORELACIJE PREDIKTORSKIH VARIJABLI
*
HEADING(TEXT=DISTRIBUCIJE PARAMETRI I KORELACIJE PREDIKTORA,D)
STATISTICS(SCORE=P,CLASS=9,S,Z=ZP)
CORRELATION(SCORE=ZP,R=RPP)
PRINT(MATRIX=RPP,TEXT=INTERKORELACIJE PREDIKTORSKIH VARIJABLI)
      DELETE(MATRIX=P)
```

```
*
*       BLOK 2.
*       DISTRIBUCIJE, PARAMETRI I KORELACIJE KRITERIJSKIH VARIJABLI
*
HEADING(TEXT=DISTRIBUCIJE PARAMETRI I KORELACIJE KRITERIJA,D)
STATISTICS(SCORE=C,CLASS=9,S,Z=ZC)
CORRELATION(SCORE=ZC,R=RCC)
PRINT(MATRIX=RCC,TEXT=INTERKORELACIJE KRITERIJSKIH VARIJABLI)
        DELETE(MATRIX=C)
*
*       BLOK 3.
*       KROSKORELACIJE PREDIKTORSKIH I KRITERIJSKIH VARIJABLI
*
HEADING(TEXT=KROSKORELACIJE PREDIKTORSKIH I KRITERIJSKIH VARIJABLI,D)
CROSSCORRELATION(P1=ZP,P2=ZC,R12=RPC)
PRINT(MATRIX=RPC,TEXT=KROSKORELACIJE PREDIKTORA I KRITERIJA)
*
*       BLOK 4.
*       REGRESIJSKA ANALIZA NAJMANJIM KVADRATIMA
*
HEADING(TEXT=REGRESIJSKA ANALIZA NAJMANJIM KVADRATIMA,D)
TRANSPOSE(OLD=RPC,NEW=RCP)
INVERSION(R=RPP)
REGRESSION(SCORE=ZP,R12=RCP,BETA=BETA,FB=F)
MULT(A=BETA,B=RPC,M=G)
LINEAR(A=RCC,B=G,CB=-1.0,M=ERRORL)
PRINT(MATRIX=ERRORL,TEXT=KOVARIJANCE POGRESKE - NAJMANJI KVADRATI)
*
*      BLOK 5.
*      STUPIDNA REGRESIJSKA ANALIZA
*
HEADING(TEXT=STUPIDNA REGRESIJSKA ANALIZA,D)
MULT(A=RPC,TA,B=RPC,M=Q)
DIAGMULT(A=RPC,D=Q,C=-0.5,R,M=GAMA)
MULT(A=RPP,B=GAMA,M=ITA)
MULT(A=GAMA,TA,B=ITA,M=C)
DIAG(A=C,C=-0.5,D=D)
DIAGMULT(A=GAMA,D=D,R,M=ALPHA)
SCALE(C=C,R=M)
MULT(A=RPC,TA,B=ALPHA,M=RO)
PRINT(MATRIX=GAMA,TEXT=STUPIDNI REGRESIJSKI KOEFICIJENTI)
PRINT(MATRIX=C,TEXT=KOVARIJANCE VARIJABLI STUPIDNE REGRESIJE)
PRINT(MATRIX=ALPHA,TEXT=STANDARDIZIRANI STUPIDNI REGRESIJSKI KOEFICIJENTI)
PRINT(MATRIX=M,TEXT=KORELACIJE VARIJABLI STUPIDNE REGRESIJE)
MULT(A=RPP,B=ALPHA,M=S)
PRINT(MATRIX=S,TEXT=STRUKTURA STUPIDNIH REGRESIJSKIH FAKTORA)
PRINT(MATRIX=RO,TEXT=STUPIDNE MULTIPLE KORELACIJE I KROSKORELACIJE)
LINEAR(A=RCC,B=M,M=VESNA1)
LINEAR(A=VESNA1,B=RO,CB=-1.0,M=VESNA2)
TRANSPOSE(OLD=RO,NEW=ROT)
LINEAR(A=VESNA2,B=ROT,CB=-1.0,M=E)
PRINT(MATRIX=E,TEXT=STUPIDNE KOVARIJANCE POGRESKE,D)
       DELETE(MATRIX=ITA)
       DELETE(MATRIX=VESNA1)
       DELETE(MATRIX=VESNA2)
*
*      BLOK 6.
*      USPOREDBA STUPIDNE REGRESIJSKE ANALIZE I REGRESIJSKE ANALIZE NAJMANJIM
*      KVADRATIMA
*
HEADING(TEXT=USPOREDBA STUPIDNE I REGRESIJE NAJMANJIM KVAADRATIMA,D)
MULT(A=RPC,TA,B=GAMA,M=COVLSS)
DIAGMULT(A=RO,D=G,C=-0.5,L,M=CORLSS)
DIAGMULT (A=G,D=G,C=-0.5,R,M=CONBGB)
```

```
DIAGMULT(A=CONBGB,D=D,R,M=ITA)
DIAGMULT(A=ITA,D=G,C=-0.5,L,M=CONBG)
MULT(A=BETA,B=BETA,TB,M=XENIA)
DIAGMULT(A=CONBGB,D=XENIA,C=-0.5,L,M=COSBGB)
DIAGMULT(A=XENIA,D=G,C=-0.5,L,R,M=VESNA)
MULT(A=ALPHA,TA,B=ALPHA,M=MAJA)
DIAGMULT(A=CONBG,D=VESNA,C=-0.5,L,M=JASNA)
DIAGMULT(A=JASNA,D=MAJA,C=-0.5,R,M=COSBG)
PRINT(MATRIX=COVLSS,TEXT=KOVARIJANCE NAJMANJIH KVADRATA I STUPIDNIH REGVARIABLI)
PRINT(MATRIX=CORLSS,TEXT=CORELACIJE NAJMANJIH KVADRATA I STUPIDNIH REGVARIJABLI)
PRINT (MATRIX=COSBGB,TEXT=KONGRUENCIJE LS AND SRA REGRESIJSKIH TEZINA)
PRINT(MATRIX=COSBG,TEXT=KONGRUENCIJE STANDARDIZIRANIH REGRESIJSKIH TEZINA)
      DELETE(MATRIX=ITA)
      DELETE(MATRIX=XENIA)
      DELETE(MATRIX=VESNA)
      DELETE(MATRIX=MAJA)
      DELETE(MATRIX=JASNA)
TRANSPOSE(OLD=S,NEW=ST)
CONGRUENCE(F1=F,F2=ST,CONGRU=COSFS)
PRINT(MATRIX=COSFS,TEXT=KONGRUENCIJE LS AND SRA REGRESIJSKIH FAKTORA)
*
HEADING(TEXT=KRAJ PROGRAMA CAOS)
*
```

NEKE RELACIJE IZMEDJU LINEARNE REGRESIJSKE ANALIZE POD MODELOM NAJMANJIH KVADRATA I STUPIDNE REGRESIJSKE ANALIZE*

Dobrić, V., J. Štalec i K. Momirović

SAŽETAK

Uobičajeni postupak linearne regresijske analize, pod modelom najmanjih kvadrata (LSR), baziran je na maksimizaciji korelacije linearne kombinacije prediktorskih varijabli i kriterija i može se promatrati kao posebni slučaj kanoničke korelacijske analize (Hotelling, 1936; Wilks, 1962; Guanadesikan, 1977). Nasuprot tome, nedavno predloženi SRA (Stupid Regression Analysis) model multivarijatne robustne regresijske analize Štaleca i Momirovića (1983) baziran je na maksimizaciji kovarijance linearne kombinacije standardiziranih prediktorskih varijabli i kriterijske varijable i predstavlja specijalni slučaj kanoničke analize kovarijance (Momirović, Dobrić i Karaman, 1983).

Usporedba procijenjenih parametara dobivenih pod LSR i SRA modelom pokazuje slijedeće:

(1) kovarijanca izmedju LSR i SRA procjene kriterijske varijable jednaka je kovarijanci kriterijske varijable i SRA procjene kriterijske varijable;

(2) korelacija izmedju LSR i SRA procjene kriterijske varijable jednaka je η/ρ, gdje je ρ multipla korelacija dobijena pod LSR modelom, a η multipla korelacije pod SRA modelom;

(3) skalarni produkt vektora regresijskih koeficijenata dobijen pod LSR i SRA modelom jednak je ρ^2/c gdje je c kovarijanca izmedju kriterijske varijable i SRA procjene kriterijske varijable;

(4) kovarijanca izmedju reziduala dobijenih pod LSR i SRA modelom jednaka je varijanci reziduala dobijenih pod LSR modelom;

(5) korelacija izmedju reziduala dobijenih pod LSR i SRA modelom jednaka je omjeru standardne devijacije reziduala dobijenih pod LSR i SRA modelom.

---

* *Ovaj rad izvorno je napisan na engleskom jeziku i bit će izložen na 6. simpoziju računarske statistike, COMPSTAT '84, Prag, 1984.*

1. UVOD

Neka je $E = \{e_i;\ i=1,\dots,n\}$ slučajni uzorak iz populacije P, i neka je $V = \{v_j;\ j=1,\dots,m\}$ skup prediktorskih varijabli izabran iz univerzuma U svih mogućih prediktora za kriterijsku varijablu k. Neka je $Z = (z_{ij});\ i=1,\dots,n;\ j=1,\dots,m$ matrica podataka dobijena kao rezultat nad uredjenim parom $E \times V$ i neka je $K = (k_i),\ i=1,\dots,n$ vektor podataka dobijen nad uredjenim parom $E \times k$. Bez gubljenja općenitosti, dozvoljeno je odrediti da je $Z^T 1 = 0$, $K^T 1 = 0$, $\mathrm{diag}\,(Z^T Z \frac{1}{n}) = I$ i $K^T K \frac{1}{n} = 1$, tj. da su standardizirane varijable iz V i k.

Označimo li sa $R = Z^T Z \frac{1}{n}$ matricu interkorelacija prediktorskih varijabli i sa $Q = Z^T K \frac{1}{n}$ vektor korelacija prediktorskih varijabli i kriterija, rješenje generalnog linearnog problema

$$K = Z\beta + E \quad \Big| \quad E^T E = \min$$

je $\beta = R^{-1} Q$. Varijanca kriterijskog vektora $K^* = Z\beta$, procijenjenog pod modelom najmanjih kvadrata, je $\rho^2 = K^{*T} K \frac{1}{n} = Q^T R^{-1} Q$; varijanca rezidualnog vektora E jednaka je $\varepsilon^2 = (K - K^*)^T (K - K^*) \frac{1}{n} = 1 - \rho^2$ a korelacija izmedju kriterija i najmanjim kvadratima procijenjenog kriterija, poznata kao multipla korelacija, je $K^T K^* \rho^{-1} \frac{1}{n} = \rho$.

Isti rezultat se može dobiti u okviru kanoničke korelacijske analize. Rješenje problema

$$\tilde{K} = Z\gamma \quad \Big| \quad \tilde{K}^T \tilde{K} \frac{1}{n} = \gamma^T R \gamma = 1, \quad K^T \tilde{K} \frac{1}{n} = \rho = \max$$

je $Q^T R^{-1} Q = \rho^2$ sa $\gamma = R^{-1} Q \rho^{-1} = \beta\rho^{-1}$. U oba slučaja $F = Z^T \tilde{K} \frac{1}{n} = Z^T K^* \rho^{-1} \frac{1}{n} = Q\rho^{-1}$ je struktura regresijskog faktora.

Regresijska analiza pod vidom najmanjih kvadrata se, dakle, može shvatiti kao problem pronalaženja maksimalne korelacije izmedju linearne kombinacije prediktorskih varijabli i (standardizirane) kriterijske varijable. Neki njeni nedostaci su dobro poznati npr. da je $E(\rho^2 | \tilde{\rho}^2 = 0) = (m-1)/(n-1)$ pri čemu je $\tilde{\rho}^2$ populacijska vrijednost koeficijenta determinacije; nadalje, β koeficijenti su vrlo nestabilni za mali broj stupnjeva slobode $df = n - m - 1$ kao i u slučaju kad je R skoro singularna matrica, jer matrica kovarijanci regresijskih koeficijenata $V(\beta) = (1 - \delta^2) R^{-1} (n - m - 1)^{-1}$ ovisi i o broju stupnjeva slobode i o R^{-1}, i posebno o $U^{-2} = \mathrm{diag} R^{-1}$ (Wilks, 1962; Anderson, 1958; Kendall i Stuart, 1973; Guanadesikan, 1977).

Model SRA (Momirović i Štalec, 1983) je jedan od modela robustne regresijske analize, definiran maksimiziranjem kovarijanci, a ne korelacije, linearne kombinacije standardiziranih prediktorskih varijabli i standardizirane kriterijske varijable, pod uvjetom da je norma vektora regresijskih koeficijenata jednaka 1. Rješenje problema

$$Zx = \psi \quad \Bigg| \quad \begin{aligned} &K^T \psi \frac{1}{n} = c = \max \\ &x^T x = 1 \end{aligned}$$

je, očito, $x = Q(Q^T Q)^{-1/2}$ sa $c = (Q^T Q)^{1/2}$. Varijanca od ψ je $\sigma^2 = x^T R x = (Q^T Q)^{-1} (Q^T R Q)$ tako da je SRA multipla korelacija jednaka $\eta = K^T \psi\, \sigma^{-1} \frac{1}{n} = c\, \sigma^{-1} = (Q^T Q)(Q^T R Q)^{-1/2}$. SRA, naravno, ne minimizira varijancu reziduala jer je $r^2 = (K - \psi)^T (K - \psi) \frac{1}{n} = 1 + \sigma^2 - 2c$. Ako sa $\phi = \psi\sigma^{-1}$ označimo standardizirane SRA procjene od K, $\tilde{r}^2 = (K - \phi)^T (K - \phi) \frac{1}{n} = 2(1 - \eta)$, pa je struktura SRA regresijskih faktora jednaka $S = Z^T \phi \frac{1}{n} = R x \sigma^{-1}$.

Test hipoteze $\tilde{\rho} = 0 \leftrightarrow \tilde{Q} = 0$ u LSR modelu je $f = \rho^2 (n - m - 1)/(1 - \rho^2) m$ sa Snedecorovom distribucijom sa m i n - m - 1 stupnjeva slobode. Test hipoteze $\tilde{\eta} = 0 \leftrightarrow \tilde{Q} = 0$ u SRA je jednostavno $f^* = q^2(n-2)(1-q^2)^{-1}$, sa Snedecorovom distribucijom uz 1 i n - 2 stupnjeva slobode, pri čemu je $q = \max_j(q_j)$, gdje je $Q^T = (q_1, \ldots, q_j, \ldots, q_m)$.

Očito, ρ^2 nije samo varijanca od K^*, već u isto vrijeme i kovarijanca izmedju K i K^*; nadalje, ρ nije samo korelacija izmedju K i K^* već u isto vrijeme i standardna devijacija od K^*. Ovaj tip jednostavnosti ne postoji u SRA modelu, no, c je jedina nenulta svojstvena vrijednost od Q pa je SRA, u stvari, bazična dekompozicija matrice Q.

2. ODNOS LRS I SRA MODELA

Razmotrimo prvo odnos LSR i SRA procjene kriterijske varijable. Kovarijanca izmedju $K^* = Z\beta$ i $\Psi = Z\chi$ je

$$K^{*T}\psi\frac{1}{n} = Q^T\chi = (Q^TQ)^{1/2}$$

što znači da je kovarijanca izmedju LSR i SRA procjena kriterija jednaka kovarijanci kriterijske varijable i SRA procjene kriterijske varijable, tj. upravo vrijednosti koje se dobiju maksimiziranjem u SRA modelu. Korelacija izmedju $\tilde{K} = Z\gamma = Z\beta\rho^{-1}$ i $\phi = Z\chi\sigma^{-1}$ je

$$\tilde{K}^T\phi\frac{1}{n} = \rho^{-1}c\,\sigma^{-1} = \rho^{-1}\eta$$

pa je korelacija izmedju LSR i SRA procjene kriterija jednaka omjeru multiple korelacije, dobijene SRA modelom i multiple korelacije dobijene LSR modelom. Odatle, $\eta \leq \rho$; jednakost se postiže samo za $Q = 0$ ili $R = I$. Naravno, $Q = 0$ je trivijalni slučaj ortogonalnosti kriterijskog vektora na prostor razapet prediktorskim vektorima u slučaju da su varijable iz skupa $\{V, k\}$ multivarijatno normalno distribuirane, odnosno ako se mogu pretpostaviti linearne veze varijabli iz $\{V, k\}$. U slučaju $R = I$,

$\beta = Q$ i LSR i SRA model su jednaki.

S interpretativne točke gledišta, iz korelacije izmedju LSR i SRA procjena mogu se izvesti vrlo korisni zaključci. Najvažniji medju njima je da se LSR procjena kriterijske varijable može dobiti iz njene SRA procjene i obratno, premda ovaj drugi smjer nije od posebnog značaja u većini realnih situacija. Procjena vrijednosti $\tilde{K}$ iz ϕ je

$$\phi^* = \phi\,\eta\,\rho^{-1}$$

sa varijancom pogreške $1 - \eta^2\rho^{-2}$.

Razmotrimo odnos izmedju LSR i SRA vektora regresije. Skalarni produkt izmedju $\beta = R^{-1}Q$ i $X = Q(Q^TQ)^{-1/2}$ je

$${}^TX = Q^TR^{-1}Q(Q^TQ)^{-1/2} = \rho^2c^{-1}$$

što odgovara omjeru kovarijanci izmedju kriterijske varijable i SRA procjene kriterijske varijable. Kosinus kuta vektora β i X je

$$\cos\alpha_{(\beta,X)} = (Q^TR^{-2}Q)^{-1/2}\,\rho^2c^{-1}$$

jer je $\beta^T\beta = Q^TR^{-2}Q$ i $X^TX = 1$.

Interesantno je pogledati i odnos reziduala dobijenih pod LSR i SRA modelom. Kovarijanca izmedju $E = (K - Z\beta)$ i $N = (K - ZX)$ je

$$E^TN\frac{1}{n} = 1 - Q^TR^{-1}Q = 1 - \rho^2 = \varepsilon^2$$

pa je kovarijanca izmedju reziduala dobijenih pod LSR i SRA modelom jednaka varijanci reziduala dobijenih pod LSR modelom. Korelacija izmedju E i N

$$(1-\rho^2)^{-1/2}E^TN(1+\sigma^2-2c)^{-1/2}\,\frac{1}{n} = (1-\rho^2)^{1/2}(1+\sigma^2-2c)^{-1/2},$$

jednaka je omjeru standardnih devijacija reziduala dobijenih pod LSR i SRA modelom.

Kao i većina metoda robustne regresijske analize i algoritam SRA je izveden u namjeri da se omogući i obrada takvih podataka, sakupljenih od neiskusnih istraživača, koji se ne mogu podvrći uobičajenom statističkom tretmanu. Jednostavnost veza rezultata LSR i SRA modela je možda dokaz da SRA model nije promašaj.

3. PROGRAM CAOS

Da bi se omogućila komparativna analiza LSR i SRA modela, napisan je, u verziji 5.2/M SS jezika (Zakrajšek, Štalec i Momirović, 1974) kratki program pod imenom CAOS. CAOS analizira do 250 multivarijatno normalno distribuiranih kriterijskih varijabli regresiranih na skup do 250 multivarijatno normalno distribuiranih prediktorskih varijabli, pod LSR i SRA modelom i usporedjuje rezultate tih dviju metoda. Program je podijeljen u slijedeće blokove:

(0) učitavanje, kontrola i sredjivanje podataka;
(1) distribucije, parametri i korelacije prediktorskih varijabli;
(2) distribucije, parametri i korelacije kriterijskih varijabli;
(3) kroskorelacije prediktorskih i kriterijskih varijabli;
(4) regresijska analiza pod vidom najmanjih kvadrata;
(5) SRA regresijska analiza;
(6) usporedba LSR i SRA rezultata.

Program je dokumentiran i sastoji se od 79 linija izvršnog koda i 39 linija komentara, a pohranjen je u SRCE*SS-MAKRO biblioteci programa Sveučilišnog računskog centra.

LITERATURA

1. Anderson, T.W.:
An introduction to multivariate statistical analysis. Wiley, New York, 1958.

2. Guanadesikan, R.:
Methods for statistical data analysis of multivariate observations. Wiley, New York, 1977.

3. Hotelling, H.:
Relations between two sets of variates. Biometrika (1936), 28, 321-377.

4. Kendall, M.G. and A. Stuart. The advanced theory of statistics. 2. Inference and relationship (3rd ed.). Griffin, London, 1973.

5. Momirović, K., V. Dobrić and Ž. Karaman:
Canonical covariance analysis. Proceedings of 5th International Symposium "Computer at the University", 1983.

6. Štalec, J. and K. Momirović:
Some properties of a very simple model for robust regression analysis. Proceedings of 5th International Symposium "Computer at the University", 1983.

7. Wilks, S.S.:
Mathematical statistics (2nd ed.). Wiley, New York, 1962.

8. Zakrajšek, E., J. Štalec and K. Momirović:
SS - Programming system for multivariate data analysis (in Croatian; English summary). Proceedings of 1th International Symposium "Computer at the University", 1974, c8, 1-16.

2.5 ALGORITAM I PROGRAM DIANA

U suvremenoj sportskoj praksi, pravovremena klasifikacija i selekcija osnovni su preduvjeti za postizanje vrhunskih sportskih dostignuća. Stoga je utvrdjivanje objektivnih kvantitativnih pokazatelja razlika u osobinama takmičara u različitim sportskim disciplinama ili pojedinim uzrasnim kategorijama odredjene sportske discipline od presudnog značaja za mogućnost formiranja racionalnih postupaka za izbor, usmjeravanje i praćenje potencijalnih kandidata.

Procedure zasnovane na kanoničkom diskriminativnom modelu osnov su za efikasno utvrdjivanje broja i strukture funkcija razlika.

Na osnovu diskriminacijskih koeficijenata moguća je objektivizacija postupaka izbora, usmjeravanja i praćenja u sportu. Niz pogodnih matematičkih svojstava algoritma i programa DIANA omogućuje efikasno rješavanje problema ovog tipa.

```
*
*   NAZIV PROGRAMA
*
*
*                    ***************
*                    *  D I A N A  *
*                    ***************
*
*
*   AUTORI
*
*
*             K. BOSNAR
*             K. MOMIROVIC
*             F. PROT
*
*
*   FUNKCIJA
*
*
*   DIANA (DI(SKRIMINATIVNA) ANA(LIZA)) JE PROGRAM ZA KANONICKU DISKRIMINATIVNU
*   ANALIZU SKUPA OBJEKATA OPISANIH NAD SKUPOM MULTIVARIJATNO NORMALNO
*   DISTRIBUIRANIH VARIJABLI S NESINGULARNOM MATRICOM KOVARIJANCI,
*   TRANSFORMIRANIH U MAHALANOBISOV OBLIK.
*
*
*   METODA I ALGORITAM OPISANI SU U
*
*             BOSNAR, K., K. MOMIROVIC I F. PROT:
*             ALGORITAM ZA DISKRIMINATIVNU ANALIZU U MAHALANOBISOVOM
*             PROSTORU, KINEZIOLOGIJA,16,1(1984), U TISKU
*
*
*   PRIPREMA PODATAKA
*
*   U OVOJ VERZIJI SS JEZIKA, DOZVOLJEN JE EFEKTIV UZORKA DO
*   10 000, BROJ VARIJABLI DO 250 I BROJ GRUPA DO 20, BROJ
*   GRUPA MOZE DOSECI I 250 AKO SE PRVA I POSLJEDNJA NAREDBA STATISTICS
*   MODIFICIRA UKLANJANJEM SIGNALA V=6, NO U TOM SLUCAJU DIANA NECE
*   IZVRSITI ANALIZU VARIJANCE ORIGINALNIH VARIJABLI I NEROTIRANIH I
*   ROTIRANIH DISKRIMINATIVNIH VARIJABLI.
*
*   PRIJE UPOTREBE PROGRAMA POTREBNO JE PRIREDITI:
*
*   (1) DATOTEKU DATA SA POTPUNIM ORIGINALNIM PODACIMA. SESTI ZNAK
*       IMENA ENTITETA TREBA BITI INDIKATOR PRIPADNOSTI GRUPI.
*
*   (2) DATOTEKU ATA SA SELEKTORSKOM MATRICOM U KOJOJ JE OZNACENA
*       PRIPADNOST ENTITETA GRUPI
*
*   (3) KONTROLNE ZAPISE PRIDRUZENE PODACIMA I SELEKTORSKOJ MATRICI
*
*   (4) U NAREDBI LINEAR U SEKCIJI 3, 1/N TREBA ZAMIJENITI REALNIM
*       BROJEM 1/N, GDJE JE N = BROJ ENTITETA U UZORKU
*
*   (5) U NAREDBI LINEAR U SEKCIJI 4, N-K TREBA ZAMIJENITI REALNIM
*       BROJEM N-K, GDJE JE N = BROJ ENTITETA I K = BROJ GRUPA.
*
*
*   IZVRSNI DIO PROGRAMA
*
*
*   BLOK 0.
*   INPUT, CONTROL, SORT AND CONFORM ULAZNIH PODATAKA.
```

```
*
OUTPUT(DEVICE=PR)
HEADING(TEXT= D  I  A  N  A ,T)
HEADING(TEXT=DISKRIMINATIVNA ANALIZA U MAHALANOBISOVOM PROSTORU)
HEADING(TEXT=S ROTACIJOM DISKRIMINATIVNIH FAKTORA,D)
TEXT(TEXT=                DIANA)
INPUT(SCORE=SCR,DATA=DATA)
INPUT(SCORE=SEL,DATA=ATA)
CONFORM(IN1=SCR,IN2=SEL,OUT1=B,OUT2=S)
          DELETE(MATRIX=SCR)
          DELETE(MATRIX=SEL)
*
*    BLOK 1.
*    DISTRIBUCIJE, PARAMETRI I UNIVARIJATNA ANALIZA VARIJANCE
*    ORIGINALNIH, NESTANDARDIZIRANIH VARIJABLI.
*
HEADING(TEXT=DISTRIBUCIJE I PARAMETRI ORIGINALNIH VARIJABLI)
HEADING(TEXT=I REZULTATI UNIVARIJATNE ANALIZE VARIJANCE,D)
STATISTICS(SCORE=B,S,V=6)
*
*    BLOK 2.
*    INTERKORELACIJE VARIJABLI I TRANSFORMACIJA PODATAKA U
*    MAHALANOBISOVU FORMU.
*
HEADING(TEXT=INTERKORELACIJE VARIJABLI)
HEADING(TEXT=I TRANSFORMACIJA PODATAKA U MAHALANOBISOV OBLIK,D)
CORRELATION(C=C,R=R)
PRINT(MATRIX=R,TEXT=INTERKORELACIJE VARIJABLI)
DIAGONALISATION.
ORTHOSCORES.
DIAGMULT(A=X,T,D=LAMBDA,C=-0.5,R,M=XLM)
MULT(A=XLM,B=X,    M=RMP)
DIAGMULT(A=X,T,D=LAMBDA,C=0.5,M=XL,R)
MULT(A=XL,B=X,    M=RP)
          DELETE(MATRIX=LAMBDA)
          DELETE(MATRIX=X)
          DELETE(MATRIX=XLM)
          DELETE(MATRIX=XL)
PRINT(MATRIX=RP,TEXT=KORELACIJE ORIGINALNIH I MAHALANOBISOVIH VARIABLI)
MULT(A=S,TA,B=S,M=STS)
DIAGMULT(A=S,T,D=STS,C=-1.0,L,M=SS)
MULT(A=SS,B=Z,M=MZ)
PRINT(MATRIX=MZ,TEXT=ARITMETICKE SREDINE GRUPA U STANDARDIZIRANOM PROSTORU)
MULT(A=SS,B=V,M=M)
PRINT(MATRIX=M,TEXT=ARITMETICKE SREDINE GRUPA U MAHALANOBISOVOM PROSTORU)
          DELETE(MATRIX=B)
*
*    BLOK 3.
*    DISKRIMINITIVNA ANALIZA U MAHALANOBISOVOM PROSTORU.
*
HEADING(TEXT=DISKRIMINATIVNA ANALIZA U MAHALANOBISOVOM PROSTORU)
HEADING(TEXT=I REDUKCIJA NA ZNACAJNE DISKRIMINATIVNE FUNKCIJE,D)
MULT(A=STS,B=M,M=STSM)
MULT(A=STSM,TA,B=M,M=AA)
LINEAR(A=AA,CA=0.0014472,M=A)
          DELETE(MATRIX=AA)
          DELETE(MATRIX=STSM)
          DELETE(MATRIX=S)
DIAGONALISATION(R=A)
HOTELLING(F=H,MIN=0.05)
          DELETE(MATRIX=X)
          DELETE(MATRIX=LAMBDA)
MULT(A=H,B=H,TB,M=L)
DIAGMULT(A=H,T,D=L,C=-0.5,R,M=X)
```

```
DIAG(A=L,C=0.5,D=RO)
PRINT(MATRIX=RO,TEXT=KOEFICIJENTI KANONICKE DISKRIMINACIJE)
PRINT(MATRIX=X,TEXT=DISKRIMINATIVNI FAKTORI U MAHALANOBISOVOM PROSTORU)
MULT(A=M,B=X,M=G)
PRINT(MATRIX=G,TEXT=CENTROIDI GRUPA U DISKRIMINATIVNOM PROSTORU)
*
*    BLOK 4.
*    TESTOVI STATISTICKE ZNACAJNOSTI.
*
HEADING(TEXT=TESTOVI STATISTICKE ZNACAJNOSTI,D)
SCALE(C=L,R=I)
LINEAR(A=L,CA= 687.0 ,M=LL)
LINEAR(A=I,B=L,CB=-1.0,M=LLL)
DIAGMULT(A=LL,D=LLL,C=-1.0,R,M=CHISQ)
         DELETE(MATRIX=I)
         DELETE(MATRIX=LL)
         DELETE(MATRIX=LLL)
PRINT(MATRIX=CHISQ,D,TEXT=RAOV TEST ZNACAJNOSTI DISKRIMINATIVNIH FUNKCIJA)
MULT(A=V,B=X,M=D)
STATISTICS(SCORE=D,V=6,Z=ZDZ)
*
*    BLOK 5.
*    DISKRIMINATIVNE FUNKCIJE U METRICI STANDARDIZIRANIH VARIJABLI.
*
HEADING(TEXT=DISKRIMINATIVNE FUNKCIJE STANDARDIZIRANIH VARIJABLI,D)
MULT(A=RMP,B=X,M=Y)
MULT(A=RP,B=X,M=F)
PRINT(MATRIX=Y,TEXT=DISKRIMINATIVNI KOEFICIJENTI STANDARDIZIRANIH VARIABLI)
PRINT(MATRIX=F,TEXT=DISKRIMINATIVNI FAKTORI STANDARDIZIRANIH VARIABLI)
*
*    BLOK 6.
*    ROTACIJA DISKRIMINATIVNIH FAKTORA.
*
HEADING(TEXT=ROTACIJA DISKRIMINATIVNIH FAKTORA,D)
TRANSPOSE(OLD=X,NEW=XT)
QUARTIMAX(F=XT,TAU=Q,FN=P,N)
PRINT(MATRIX=P,T,TEXT=ROTIRANI DISKRIMIN FAKTORI U MAHALANOBISOVOM PROSTORU)
MULT(A=M,B=P,TB,M=GQ)
PRINT(MATRIX=GQ,TEXT=CENTROIDI NA ROTIRANIM DISKRIMINATIVNIM FAKTORIMA)
MULT(A=V,B=P,TB,M=DQ)
STATISTICS(SCORE=DQ,V=6,Z=ZQZ)
PRINT(MATRIX=Q,T,TEXT=KORELACIJE NEROTIRANIH I ROTIRANIH FAKTORA)
*
*    BLOK 7.
*    ROTIRANI FAKTORI U METRICI STANDARDIZIRANIH VARIJABLI.
*
HEADING(TEXT=ROTIRANI FAKTORI U METRICI STANDARDIZIRANIH VARIJABLI,D)
MULT(A=Y,B=Q,TB,M=YQ)
MULT(A=F,B=Q,TB,M=FQ)
PRINT(MATRIX=YQ,TEXT=ROTIRANI DISKRIMIN KOEFICIJENTI STANDARDIZIRANIH VARIABLI)
PRINT(MATRIX=FQ,TEXT=ROTIRANI DISKRIMIN FAKTORI STANDARDIZIRANIH VARIABLI)
*
*    BLOK 8.
*    UOBICAJENE DISKRIMINATIVNE FUNKCIJE ORIGINALNIH, NESTANDARDIZIRANIH
*    VARIJABLI.
*
HEADING(TEXT=DISKRIMINATIVNE FUNKCIJE ORIGINALNIH VARIJABLI,D)
DIAGMULT(A=Y,D=C,C=-0.5,L,M=W)
PRINT(MATRIX=W,TEXT=DISKRIMINATIVNI KOEFICIJENTI ORIGINALNIH VARIABLI)
*
*
*
HEADING(TEXT=KRAJ LADY DIANE)
*
```

ALGORITAM ZA DISKRIMINATIVNU ANALIZU U MAHALANOBISOVOM PROSTORU*

Ksenija Bosnar, Konstantin Momirović i Franjo Prot

Institut za kineziologiju Fakulteta za fizičku kulturu sveučilišta u Zagrebu
Odjel za informatiku i statistiku

SAŽETAK

Predložen je algoritam i napisan program za kanoničku diskriminativnu analizu skupa entiteta opisanih nad skupom multivarijatno normalno distribuiranih varijabli s nesingularnom matricom kovarijanci, transformiranih u Mahalanobisov oblik.

Algoritam se razlikuje od standardnih algoritama za kanoničku diskriminativnu analizu po tome što diskriminativne koeficijente i diskriminativne faktore odredjuje i u metrici standardiziranih i u metrici varijabli transformiranih u Mahalanobisov oblik; što rotira značajne diskriminativne faktore u položaj koji maksimizira normalizirani quartimax kriterij, pri čemu broj zadržanih faktora odredjuje prema informacijskom doprinosu faktora objašnjenju ukupne intergrupne varijance; što odredjuje centroide grupa i za nerotirane i za rotirane diskriminativne varijable na koordinatnom sistemu s ishodištem u centroidu ukupnog uzorka; što statistički provjerava značajnost diskriminativnih funkcija zadržanih informacijskim kriterijem putem χ^2 testa (Rao, 1951); što provjerava rezultate χ^2 testa jednofaktorskom, univarijatnom analizom varijance za svaku zadržanu rotiranu i nerotiranu diskriminativnu varijablu.

1. UVOD

Analiza homogenosti nekog skupa entiteta opisanog nad skupom varijabli dozvoljava više strategija empirijskih istraživanja koje uglavnom ovise o preciznosti hipoteze o postojanju podskupova unutar osnovnog skupa. Kada je moguća apriorna podjela entiteta u mutualno ekskluzivne podskupove, na osnovu nekog obilježja koje ne pripada analiziranim varijablama, i kad postoji hipoteza o kvantitativnim razlikama medju tako stvorenim grupama, značajnost razlika i strukturu činilaca koji utiču na te razlike najpodesnije je ispitati kanoničkom diskriminativnom analizom.

Kanonička diskriminativna analiza je, zapravo, poseban slučaj kanoničke korelacijske analize ako se nominalna varijabla koja odredjuje pripadanje entiteta pojedinim grupama definira kao binarna selektorska matrica.

---

* *Ovaj rad će biti objavljen u časopisu Kineziologija, 16, (1984), 2, u tisku*

Zbog toga je ova metoda invarijantna na metriku varijabli, i može se definirati na više različitih načina. Neki od njih su, u nekim slučajevima, pogodniji za interpretaciju i/ili izračunavanje rezultata.

Predloženi algoritam i njemu pridruženi program DIANA razlikuje se od standardnih algoritama za kanoničku diskriminativnu analizu u više svojstava:

(1) Diskriminativna analiza izvodi se u Mahalanobisovom prostoru, a diskriminativni koeficijenti i diskriminativni faktori odredjuju se i u metrici standardiziranih i u metrici varijabli transformiranih u Mahalanobisov oblik;

(2) Broj diskriminativnih funkcija zadržanih za interpretaciju odredjuje se na temelju informacijskih svojstava diskriminativnih varijabli. Diskriminativna varijabla se smatra informacijski značajnom ako je odgovorna za neki zadovoljavajući postotak od ukupne intergrupne varijance. Ukoliko se ne odredi drugačije, program će prihvatiti kao značajnu svaku diskriminativnu varijablu koja objašnjava 5 ili više posto od ukupne intergrupne varijance;

(3) Značajnost zadržanih diskriminativnih funkcija provjerava se i statistički, χ^2 testom (Rao, 1951);

(4) Diskriminativni faktori rotiraju se u poziciju koja maksimizira normalizirani quartimax kriterij (Ferguson, 1954) i odredjuju se matrice koeficijenata i strukture za rotirane faktore i u metrici standardiziranih i u metrici varijabli transformiranih u Mahalanobisov oblik;

(5) Centroidi grupa i za rotirane i za nerotirane diskriminativne varijable odredjuju se na koordinatnom sustavu s ishodištem u centroidu ukupnog uzorka;

(6) Značajnost diskriminativnih varijabli još se jednom provjerava jednofaktorskom univarijatnom analizom varijance za svaku zadržanu rotiranu i nerotiranu diskriminativnu varijablu.

2. ALGORITAM

2.1 Uvodne operacije

Neka je $B = (b_{ij})$; $i=1,\ldots,n$; $j=1,\ldots,m$; matrica izvornih podataka entiteta skupa $E = \{e_i, i=1,\ldots,n\}$ opisanih nas skupom multivarijatno normalno distribuiranih varijabli $V = \{v_j; j=1,\ldots,m\}$. Neka je skup entiteta podijeljen u podskupove G_k, $k=1,\ldots,g$, i neka matrica $S = (s_{ik})$ apriorno odredjuje pripadnost entiteta ovim podskupovima tako da je $s_{ik}=1$ ako je entitet e_i element podskupa G_k a $s_{ik}=0$ ako e_i nije element podskupa G_k.

Odredimo matricu kovarijanci varijabli iz V na cijelom skupu entiteta E

$$C = (B^T B - B^T 1 1^T \tfrac{1}{n} B) \tfrac{1}{n}$$

gdje je 1 vektor jedinice S n elemenata i dijagonalnu matricu varijanci

$$V^2 = \text{diag } C.$$

Standardizacija varijabli na nultu aritmetičku sredinu i jediničnu varijancu rezultat je operacije

$$Z = (B - 1 1^T \tfrac{1}{n} B) V^{-1}.$$

2.2 Interkorelacije varijabli i transformacija rezultata u Mahalanobisov oblik

Interkorelacije varijabli na cijelom skupu entiteta definirane su oepracijom

$$R = V^{-1} C V^{-1} = Z^T Z \tfrac{1}{n}.$$

Eliminacija kovarijabiliteta medju varijablama izvedena je Mahalanobisovom transformacijom

$$\Psi = Z R^{-1/2}.$$

Korelacije originalnih varijabli i varijabli transformiranih u Mahalanobisov oblik su

$$Z^T \Psi \tfrac{1}{n} = R^{1/2} .$$

Očito, varijable iz Ψ su, pod kriterijem najmanjih kvadrata, ortogonalne varijable najsličnije varijablama iz Z, jer je

$$tr(Z - \Psi)^T (Z - \Psi) = \text{minimum}.$$

Aritmetičke sredine rezultata u grupama entiteta su elementi vektora redaka matrice

$$M_Z = (\text{diag}(S^T S))^{-1} S^T Z$$

za standardizirane rezultate i elementi vektora redaka matrice

$$M_M = (\text{diag}(S^T S))^{-1} S^T \Psi$$

za rezultate transformirane u Mahalanobisov oblik.

2.3 Diskriminativna analiza u Mahalanobisovom prostoru i redukcija broja diskriminativnih funkcija

Problem diskriminativne analize u pravilu se definira na osnovu dekompozicije matrice kovarijabiliteta na ukupnom uzorku na matricu kovarijabiliteta unutar grupa i matricu kovarijabiliteta izmedju grupa. Neka je T matrica kovarijabiliteta u cijelom uzorku, W matrica kovarijabiliteta unutar grupa, a A matrica kovarijabiliteta izmedju grupa. Tada je

$$T = W + A.$$

Na osnovu ove definicije problem utvrdjivanja diskriminativnih funkcija moguće je formulirati na više ekvivalentnih načina, kao što je npr. pokazao Romeder (1973), jedan od njih je maksimizacija omjera

$$\frac{xAx^T}{xTx^T} = \rho^2$$

uz neki pogodan uvjet, najčešće $x^T T x = 1$,

gdje je x vektor diskriminativnih koeficijenata koji tražimo.

Ako su varijable transformirane u Mahalanobisov oblik,

$$T = \Psi^T \Psi \frac{1}{n} = I,$$

maksimizira se

$$x^T A x \Big|_{x^T x = 1},$$

što problem svodi na rješavanje karakteristične jednadžbe

$$(A - \rho^2) x = 0 \;\; ; \;\; x^T x = 1$$

gdje je Lagrangeov množitelj ρ^2 traženi maksimum i kvadrat koeficijenta kanoničke diskriminacije.

Aritmetičke sredine varijabli transformiranih u Mahalanobisov oblik za ukupni uzorak jednake su nuli, pa je matrica A u ovom slučaju jednostavno

$$A = ((S^T S) M_M)^T M_M \, \frac{1}{n} \, .$$

U diskriminativnoj analizi moguće je odrediti do $k - 1$ ortogonalnih diskriminativnih funkcija ako je $k \leq m$. Neka ih u daljoj analizi bude zadržano samo q, tj.

$$(A - \rho_p^2) x_p = 0 \; ; \qquad p = 1, \ldots, q;$$

$$q = \text{num}\,(\rho_p^2 / (\sum_{p=1}^{v} \rho^2 \div v) \geq 0.05) \qquad v = \min\,((k-1), m)$$

samo onih funkcija čiji je doprinos ukupnoj diskriminaciji jednak ili veći od 5 posto.

Definirajmo matricu

$$X = (x_p), \text{ za koju vrijedi } X^T X = I,$$

formiranu od zadržanih diskriminativnih vektora i matricu

$$\rho = (\rho_p)$$

u čijoj su dijagonali koeficijenti kanoničke diskriminacije za zadržane diskriminativne funkcije.

Značajnost zadržanih diskriminativnih funkcija može se testirati i statistički, χ^2 testom, Raovom aproksimacijom Wilksovog kriterija

(Rao, 1951)

$$\chi_p^2 = (\eta * \rho_p^2)(1 - \rho_p^2)^{-1} \;; \quad \eta = n - k$$

sa stupnjevima slobode $df_p = m + k - 2p$, što odgovara vrijednosti na F-distribuciji

$$f_p = \chi_p^2 \div df_p$$

sa stupnjevima slobode $df_1 = m + k - 2p$, $df_2 = +\infty$.

Vrijednosti entiteta u ovom diskriminativnom prostoru su

$$D = \Psi X = ZR^{-1/2} X$$

a centroidi grupa su odredjeni sa

$$G_D = (S^T S)^{-1} S^T D = (S^T S)^{-1} S^T \Psi X = (S^T S)^{-1} S^T Z R^{-1/2} X = M_M X .$$

2.4 Diskriminativne funkcije u metrici standardiziranih varijabli

Diskriminativne funkcije su standardizirane

$$D^T D \frac{1}{n} = X^T \Psi^T \Psi X \frac{1}{n} = I$$

pa su korelacije standardiziranih varijabli i diskriminativnih funkcija, tj. diskriminativni faktori standardiziranih varijabli jednaki

$$F = Z^T D \frac{1}{n} = Z^T \Psi X \frac{1}{n} = R^{1/2} X.$$

Diskriminativni ponderi za standardizirane varijable su, naravno,

$$Y = R^{-1/2} X$$

jer

$$D = Z R^{-1/2} X = Z Y .$$

2.5 Rotacija diskriminativnih faktora

Diskriminativne funkcije nije lako interpretirati. Ovaj se problem može ponekad pojednostavniti ortogonalnom rotacijom diskriminativnih faktora u neku parsimoničnu poziciju.

U ovoj se analizi izvodi ortogonalna transformacija matrice X tako da rezultat zadovoljava Fergusonov quartimax kriterij (Ferguson, 1954), tj.

$$P = XQ = (p_{jp})$$

uz uvjet

$$Q^T Q = QQ^T = I$$

tako da je

$$\sum_{j=1}^{m} \cdot \sum_{p=1}^{q} p_{jp}^4 = \max.$$

Vrijednosti entiteta na rotiranim diskriminativnim faktorima, procijenjene regresijskim postupkom, su

$$\Phi = \Psi P (P^T P)^{-1}$$

no kako je

$$P^T P = Q^T X^T X Q = I$$

to je

$$\Phi = \Psi P = \Psi XQ = ZR^{-1/2} XQ = ZYQ = DQ.$$

Varijable iz Φ su i dalje ortogonalne, jer

$$\Phi^T \Phi \frac{1}{n} = P^T \Psi^T \Psi P \frac{1}{n} = P^T P = I.$$

Centroidi grupa na rotiranim diskriminativnim faktorima su

$$G_\Phi = (S^T S)^{-1} S^T \Phi = (S^T S)^{-1} S^T DQ = (S^T S)^{-1} S^T ZYQ =$$
$$= (S^T S)^{-1} S^T Z R^{-1/2} XQ = M_M XQ = M_M P.$$

Transformacijska matrica Q je istovremeno i matrica korelacija rotiranih i nerotiranih faktora:

$$D^T \Phi \frac{1}{n} = X^T \Psi^T \Psi P \frac{1}{n} = X^T P = X^T X Q = Q.$$

Testovi značajnosti učinjeni za nerotirane diskriminativne funkcije, naravno, ne vrijede za rotirane faktore, pa se značajnost pojedinog rotiranog diskriminativnog faktora može provjeriti jednofaktorskom, univarijatnom analizom varijance.

2.6 Rotirani faktori u metrici standardiziranih varijabli

Korelacije standardiziranih varijabli i rotiranih diskriminativnih varijabli, tj. rotirani diskriminativni faktori standardiziranih varijabli su

$$H = Z^T \Phi \frac{1}{n} = Z^T Z R^{-1/2} X Q \frac{1}{n} = R^{1/2} X Q = F Q.$$

Rotirani diskriminativni ponderi za standardizirane varijable su

$$U = YQ$$

jer

$$\Phi = \Psi X Q = Z R^{-1/2} X Q = Z Y Q = Z U.$$

2.7 Diskriminativne funkcije originalnih varijabli

Na kraju, algoritmom se odredjuju i diskriminativni ponderi za originalne, nestandardizirane varijable.

Ako je

$$Z = (B - 1 1^T \tfrac{1}{n} B)\, V^{-1}$$

iz toga slijedi da je

$$ZV = (B - 1 1^T \tfrac{1}{n} B)$$

pa, jer je

$$D = Z R^{-1/2} X = Z V V^{-1} R^{-1/2} X$$

ponderi za nestandardizirane varijable su

$$K = V^{-1} R^{-1/2} X.$$

3. PROGRAM

Program DIANA napisan je u meta jeziku SS (Zakrajšek, Štalec i Momirović, 1974), verzija 5.2/M, i implementiran je na računalu UNIVAC Sveučilišnog računskog centra u Zagrebu gdje je pohranjen u programskoj biblioteci SRCE * SS - MAKRO.

U ovoj verziji program omogućava analizu do 10.000 entiteta svrstanih u do 20 grupa i opisanih s najviše 250 varijabli. Broj grupa je moguće povećati na 250, no u tom slučaju neće biti izvedene univarijatne analize varijance na originalnim varijablama i nerotiranim i rotiranim diskriminativnim varijablama.

Korisnik programa dužan je pripremiti matricu originalnih podataka tako da u njoj nema entiteta s nepotpunim podacima, a takodjer je potrebno da prethodno kreira selektorsku matricu, najbolje uz pomoć jednog od programa za binarizaciju kvalitativnih varijabli, npr. BINAR ili VESNA, takodjer dostupnih na Sveučilišnom računskom centru. Detaljnije upute o korištenju programa nalaze se unutar samog programa.

4. NUMERIČKI PRIMJER

Efikasnost algoritma i programa DIANA provjerena je na rezultatima 653 klinički zdrava muška ispitanika, stara 19-27 godina. Ispitanici su podijeljeni u četiri grupe prema slijedećim obilježjima mjesta u kome žive:

(1) Republičko središte (MJES 4)
(2) Mjesto u kom je sjedište okružnog suda (MJES 3)
(3) Mjesto u kom je sjedište općine (MJES 2)
(4) Mjesto koje nema ni jednu od navedenih karakteristika (MJES 1).

Ispitanici su izmjereni sa deset morfoloških mjera: (1) visina tijela (VISINA), (2) dužina ruke (DUZIRU), (3) dužina noge (DUZINO), (4) težina (TEZINA), (5) opseg nadlaktice (OPNAD), (6) opseg podlaktice (OPPODL), (7) opseg potkoljenice (OPPOTK), (8) nabor na ledjima (NANALE), (9) nabor na trbuhu (NATRBU) i (10) nabor na nadlaktici (NANADL) na način kako propisuje Internacionalni Biološki Program.

Primijenjenim informacijskim kriterijem značajnosti sve tri diskriminativne funkcije smatrale bi se značajnim (tabela 1). Raov χ^2 test je, medjutim, znatno stroži i po njemu je značajna samo prva diskriminativna funkcija. Sudeći po rezultatima χ^2 testa, a isto tako i prema slabo razmaknutim centroidima grupa na trećoj diskriminativnoj funkciji (tabela 3), treća funkcija je vjerojatno suvišna i trebalo bi, bar u ovoj analizi, pooštriti kriterij od 5% doprinosa ukupnoj diskriminaciji.

Morfološke varijable bile su odabrane tako da "pokriju" tri poznata morfološka faktora: longitudinalnu dimenzionalnost skeleta, volumen i masu tijela i faktor masnog tkiva. Kao što je vidljivo u tabeli 2, samo diskriminativni faktori standardiziranih varijabli donekle slijede taj model, što je potpuno razumljivo, jer transformacija u Mahalanobisov prostor uništava postojeću zajedničku varijancu varijabli pa treba napustiti interpretaciju u terminima zajedničkih činilaca i prihvatiti interpretaciju u terminima nezavisnih varijabli.

Prvi diskriminativni faktor u Mahalanobisovom prostoru definiran je najviše visinom, zatim naborom na ledjima (koji značajno ovisi od broja masnih stanica) te opsegom nadlaktice i opsegom podlaktice s negativnim predznakom, tj. slabijom i više podložnom specifičnim utjecajima mjerom mišićne mase i boljom mjerom ukupne mišićne mase sa suprotnim predznakom. Iz vrijednosti centroida grupa na prvoj funkciji (tabela 3) vidljiv je razmak izmedju stanovnika malih mjesta i republičkog središta, a po položaju na funkciji može se zaključiti da je vjerojatnije da će veću visinu, više masnog tkiva dijelom endogenog porijekla i više mišićne mase koja je pretežno uvjetovana egzogenim činiocima imati stanovnici republičkog središta nego stanovnici malog mjesta ili sela.

Zanimljiv je podatak negativna vrijednost za MJES 3 na drugoj diskriminativnoj funkciji. Stanovnici mjesta koje je sjedište okružnog suda, što su obično srednje veliki gradovi, razlikuju se od ostalih ispitanika po tome što imaju manju visinu, kraće udove, manju težinu i manji opseg podlaktice, tj. manje vrijednosti na varijablama koje definiraju drugu funkciju. Ovo može biti posljedica negativne selekcije uslijed migracionih kretanja i/ili lošijih uvjeta za rast i razvoj i od malih mjesta i sela i od velikih gradova.

Rotirani diskriminativni faktori pružaju nešto drugačiju sliku (vidi tabelu 6). Donekle je reproducirana druga nerotirana diskriminativna funkcija u Mahalanobisovom prostoru (tabela 4) na kojoj centroid grupe MJES 3 ima ponovo negativnu vrijednost (tabela 5).

Prva rotirana diskriminativna funkcija u Mahalanobisovom prostoru definirana je pretežno Mahalanobisovom reprezentacijom visine i, znatno manje, težine, i ponovo dobro diferencira stanovnike malih mjesta i sela na negativnom polu od stanovnika republičkih središta na pozitivnom polu.

Treći rotirani diskriminativni faktor u Mahalanobisovom prostoru pomalo je neobičan. Definiraju ga uglavnom nabor na ledjima i opseg nadlaktice i zanimljiv je po tome što ne razlikuje stanovnike MJES 3 i MJES 4, tj. stanovnike uglavnom srednjih i velikih gradova.

U prostoru standardiziranih varijabli, prvi rotirani diskriminativni faktor nije interpretabilan, dok se preostala dva lako prepoznaju kao faktor općeg rasta bez učešća masnog tkiva (Q_2) i kao faktor mekih tkiva (Q_3).

Kako se vidi, algoritam i program DIANA su upotrebljivo orudje za diskriminativnu analizu, koje može dati zanimljive informacije i kad se analizira dobro poznat problem kao što su morfološke karakteristike ispitanika različitog rezidencijalnog statusa.

Tabela 1 - SVOJSTVENE VRIJEDNOSTI DISKRIMINATIVNE JEDNADŽBE (ρ^2), KOEFICIJENTI KANONIČKE DISKRIMINACIJE (ρ), POSTOTAK DORPINOSA POJEDINE DISKRIMINATIVNE FUNKCIJE UKUPNOJ DISKRIMINACIJI $(\rho\%)$, REZULTAT RAOVOG χ^2 TESTA (χ^2), STUPNJEVI SLOBODE ZA χ^2 TEST (df) I ZNAČAJNOST χ^2 VRIJEDNOSTI (p). SA RB JE OZNAČEN INDEKS DISKRIMINATIVNE FUNKCIJE

RB	ρ^2	ρ	$\rho\%$	χ^2	df	p
1	.039	.197	69.8	26.08	12	.0106
2	.012	.110	21.9	7.98	10	.6316
3	.005	.068	8.3	2.99	8	.9345

Tabela 2 - NEROTIRANI DISKRIMINATIVNI FAKTORI U MAHALANOBISOVOM PROSTORU (x), DISKRIMINATIVNI KOEFICIJENTI (y) I DISKRIMINATIVNI FAKTORI (F) STANDARDIZIRANIH VARIJABLI

	x_1	x_2	x_3	y_1	y_2	y_3	F_1	F_2	F_3
VISINA	(.67)	(.49)	-.25	-.29	(-.93)	-.50	-.07	(.56)	.18
DUZIRU	-.05	(.30)	(.32)	.49	-.30	.61	.10	(.57)	.35
DUZINO	-.13	(.50)	.11	-.10	(1.84)	-.48	-.03	(.82)	.15
TEZINA	-.01	(.42)	-.20	-.24	.21	(1.23)	.11	.32	(.52)
OPNADL	(.34)	-.23	.26	(-1.08)	-.20	.45	-.01	-.03	(.62)
OPPODL	(-.32)	(.34)	.16	(.81)	.17	.26	.21	.09	(.63)
OPPOTK	.21	.00	-.19	-.06	-.15	(-1.04)	.05	.05	.10
NANALE	(.48)	-.08	(.62)	(1.08)	-.42	-.14	(.56)	-.13	.24
NATRBU	.12	-.26	(-.47)	.50	.37	-.33	(.55)	.13	.04
NANADL	.18	.05	-.22	-.50	.00	.08	.21	-.02	.19

Tabela 3 - CENTROIDI GRUPA NA NEROTIRANIM DISKRIMINATIVNIM VARIJABLAMA

	x_1	x_2	x_3
MJES 1	-.220	.018	.059
MJES 2	-.013	.022	-.086
MJES 3	.161	-.276	.024
MJES 4	.340	.120	.059

Tabela 4 - ROTIRANI DKSIRMINATIVNI FAKTORI U MAHALANOBISOVOM PROSTORU (P), ROTIRANI DISKRIMINATIVNI KOEFICIJENTI (K) I DISKRIMINATIVNI FAKTORI (Q) STANDARDIZIRANIH VARIJABLI

	P_1	P_2	P_3	K_1	K_2	K_3	Q_1	Q_2	Q_3
VISINA	(.85)	.11	.13	-.46	(-.95)	-.27	.16	(.57)	-.09
DUZIRU	-.03	(.43)	.10	-.10	-.02	(.84)	.21	(.63)	.15
DUZINO	.12	(.49)	-.15	(1.11)	(1.24)	(-.92)	(.34)	(.75)	-.15
TEZINA	(.31)	.23	-.26	-.65	(.90)	.62	.00	(.52)	(.34)
OPNADL	.00	-.12	(.47)	(-1.08)	.32	-.35	(-.32)	(.32)	(.42)
OPPODL	-.13	(.43)	-.20	.54	.10	.67	-.10	(.37)	(.54)
OPPOTK	.24	-.15	.01	.37	-.68	-.71	.01	.08	.09
NANALE	.00	.18	(.77)	.61	-.66	(.75)	.22	-.10	(.57)
NATRBU	.17	(-.49)	-.17	.70	.00	.00	(.44)	.00	(.36)
NANADL	.26	-.12	-.04	-.39	.16	-.28	.05	.05	(.28)

Tabela 5 - CENTROIDI GRUPA NA ROTIRANIM DISKRIMINATIVNIM VARIJABLAMA

	P_1	P_2	P_3
MJES 1	-.175	.096	-.111
MJES 2	.043	-.027	-.074
MJES 3	-.042	-.245	.203
MJES 4	.276	.053	.234

Tabela 6 - KORELACIJE NEROTIRANIH (X) I ROTIRANIH DISKRIMINATIVNIH FAKTORA (P)

	P_1	P_2	P_3
X_1	.71	-.22	.67
X_2	.52	.80	-.28
X_3	-.47	.55	.69

LITERATURA

1. Ferguson, G.A.: The concept of parsimony in factor analysis. Psychometrika, 19, (1954), 281-290.

2. Rao, C.R.: An asymptotic expansion of the distribution of Wilks' criterion. Bulletin of international statistics institute, 33, 2, (1951), 177-180.

3. Romeder, J.: Methodes et programmes d'analyse discriminante. DUNOD, Paris, 1973.

4. Zakrajšek, E., J. Štalec i K. Momirović: SS-programski sistem za multivarijatnu analizu podataka. Zbornik Simpozija "Kompjuter na Sveučilištu", Sveučilišni računski centar, Zagreb, 1974, C8-1 - C8-16.

2.6 ALGORITAM I PROGRAM ITA

Za postizanje vrhunskih sportskih dostignuća sve je manje mjesta za neuspjele pokušaje i pogreške. Proučavanje sportskih grupa značajno je za cjelokupnu aktivnost takmičara, koja se ostvaruje kroz djelovanje unutar grupe ili za ciljeve grupe.

Grupa i grupni odnosi mogu biti presudni za podnošenje maksimalnih psihofizičkih opterećenja, a izuzetni napori na velikim utakmicama ili takmičenjima traže maksimalnu mobilizaciju pojedinca i kolektiva i njihovu interakciju.

Primjenom algoritma i programa ITA moguće je na osnovu komunikacije unutar grupe utvrditi mikrosocijalnu strukturu sportskog kolektiva, koja je bazični podatak na osnovu kojeg treneri mogu racionalno organizirati vodjenje kolektiva i pojedinca kroz pripremni i takmičarski period uz maksimizaciju pozitivnih efekata.

NAZIV PROGRAMA

*** ITA ***

AUTORI

K. MOMIROVIC
A. HOSEK
K. BOSNAR
F. PROT

FUNKCIJA

ITA ANALIZIRA BINARNU MATRICU KOJOM JE OPISANA MREZA OD M CVOROVA. U KOLONAMA TE MATRICE VEKTORI, PRIDRUZENI SVAKOM CVORU, DEFINIRAJU POSTOJANJE DOPUSTIVO UNIPOLARNOG KOMUNIKACIJSKOG KANALA IZMEDJU TOG CVORA I SVAKOG DRUGOG CVORA U MREZI. U RETCIMA TE MATRICE VEKTORI, PRIDRUZENI SVAKOM CVORU, DEFINIRAJU POSTOJANJE DOPUSTIVO UNIPOLARNOG KOMUNIKACIJSKOG KANALA KOJI, PRIDRUZEN NEKOM OD OSTALIH CVOROVA MREZE, ZAVRSAVA NA TOM CVORU.

ITA IZVODI SPEKTRALNU DEKOMPOZICIJU MATRICE MREZE, REDUCIRA BROJ LIJEVIH I DESNIH SVOJSTVENIH VEKTORA NA BROJ IZNADPROSJECNIH SPEKTRALNIH VRIJEDNOSTI, ORTONORMALNO TRANSFORMIRA LIJEVE I DESNE SVOJSTVENE VEKTORE TAKO DA MAKSIMIZIRA NORMALIZIRANU VARIMAX FUNKCIJU, ODREDJUJE SKLOP I STRUKTURU IZLAZNIH I ULAZNIH VEKTORA U KOORDINATNIM SUSTAVIMA KOJI SU DOBIJENI OVIM TRANSFORMACIJAMA, I ODREDJUJE RELACIJE KOORDINATNIH SUSTAVA.

ITA JE NAMJENJENA ZA ANALIZU SOCIOMETRIJSKIH PODATAKA, ALI SE, NARAVNO, MOZE PRIMJENITI I ZA ANALIZU BILO KAKVIH MREZA KOJE SE MOGU REPREZENTIRATI BINARNIM MATRICAMA.

METODA I ALGORITAM OPISANI SU U

MOMIROVIC, K., A. HOSEK, K. BOSNAR I F. PROT:
ALGORITAM ZA DETEKCIJU KLIKA NA OSNOVU STRUKTURE
KOMUNIKACIJSKIH MREZA, KINEZIOLOGIJA,16,1(1984), U TISKU

PRIPREMA PODATAKA

ITA MOZE ANALIZIRATI MREZE OD DO 250 CVOROVA. ULAZNE OPERACIJE PRETPOSTAVLJAJU DA SU PODACI UNETI PO VEKTORIMA ULAZNIH PODATAKA, PRI CEMU SE CVOROVI RETCI MATRICE I TRETIRAJU SE KAO ENTITETI. VEKTORI ULAZNIH PODATAKA MORAJU BITI OPISANI IMENOM CVORA, KOJE MERA SADRZI DO 6 ZNAKOVA OD KOJIH JEDAN TREBA BITI IZABRAN TAKO DA OMOGUCI DA SE TAJ VEKTOR KASNIJE IDENTIFICIRA KAO ULAZNI VEKTOR. VEKTORI IZLAZA IZ CVOROVA TRETIRAJU SE KAO VARIJABLE. IMENA CVOROVA MORAJU BITI ZAPISANA U VARIABLE NAREDBAMA, A SVAKOM IMENU TREBA PRIDRUZITI ZNAK KOJI OMOGUCUJE DA SE TI VEKTORI KASNIJE IDENTIFICIRAJU KAO IZLAZNI VEKTORI.
KONTROLNI ZAPIS I PODACI NALAZE SE U DATOTECI DATA.

IZVRSNI DIO PROGRAMA

```
*              SEKCIJA 0.
*              ULAZNE OPERACIJE I ISPISIVANJE ZAGLAVLJA I NASLOVA.
*
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=I T A,T)
HEADING(TEXT=ANALIZA MREZA OPISANIH BINARNIM MATRICAMA)
HEADING(TEXT=EXPERIMENTALNA VERZIJA,D)
TEXT(TEXT=ITA              ITA               ITA)
INPUT(SCORE=B,DATA=DATA)
*
*              SEKCIJA 1.
*              ULAZNA MATRICA, MATRICA SKALARNIH PRODUKATA ULAZNIH VEKTORA I
*              MATRICA SKALARNIH PRODUKATA IZLAZNIH VEKTORA.
*
MULT(A=B,B=B.TB,M=G)
MULT(A=B,B=B.TA,M=H)
PRINT(MATRIX=B,TEXT=OPIS MREZE)
PRINT(MATRIX=G,TEXT=SKALARNI PRODUKTI ULAZNIH VEKTORA)
PRINT(MATRIX=H,TEXT=SKALARNI PRODUKTI IZLAZNIH VEKTORA)
*
*              SEKCIJA 2.
*              REDUKCIJA PROSTORA, LIJEVI I DESNI SVOJSTVENI VEKTORI
*              MATRICE PODATAKA.
*
DIAGONALISATION(R=G)
HOTELLING,
MULT(A=F,B=F.TB,M=L)
              DELETE(MATRIX=X)
              DELETE(MATRIX=LAMBDA)
DIAGMULT(A=F.T,D=L,C=-0.5,R,M=X)
MULT(A=B.TA,B=X,M=BTX)
DIAGMULT(A=BTX,D=L,C=-0.5,R,M=Y)
DIAG(A=L,C=0.5,D=LAMBDA)
PRINT(MATRIX=LAMBDA,TEXT=SPEKTRALNE VRIJEDNOSTI MATRICE PODATAKA)
PRINT(MATRIX=X,TEXT=LIJEVI VEKTORI MATRICE PODATAKA)
PRINT(MATRIX=Y,TEXT=DESNI VEKTORI MATRICE PODATAKA)
*
*              SEKCIJA 3.
*              REPRODUKCIJA MREZE U REDUCIRANOM PROSTORU.
*
MUL  A=X,B=LAMBDA,M=XL)
MULT(A=XL,B=Y.TB,M=BTEOR)
LINEAR(A=B,B=BTEOR,CB=-1.0,M=BRES)
PRINT(MATRIX=BTEOR,TEXT=REPRODUKCIJA MREZE U REDUCIRANOM PROSTORU,D)
PRINT(MATRIX=BRES,TEXT=REZIDUALNA MATRICA MREZE,D)
*
*              SEKCIJA 4.
*              FAKTORI MATRICA SKALARNIH PRODUKATA ULAZNIH VEKTORA I
*              REPRODUKCIJA OVE MATRICE U REDUCIRANOM PROSTORU.
*              KAKO SU DIJAGONALNI CLANOVI MATRICE G JEDNAKI BROJU KANALA
*              KOJI ZAVRSAVAJU SA ZNAKOM CVORU, A VANDIJAGONALNI CVOROVI
*              JEDNAKI BROJU KANALA ZAJEDNICKIH PAROVIMA CVOROVA, TO SU
*              CLANOVI REPRODUCIRANE MATRICE JEDNAKI JEDNAKI OCEKIVANOM
*              BROJU KANALA KOJI ZAVRSAVAJU NA SVAKOM CVORU ILI SU, POD
*              MODELOM KOJIM JE DEFINIRAN REDUCIRANI PROSTOR, ZAJEDNICKI
*              ZA PAROVE CVOROVA. NEGATIVNE VRIJEDNOSTI U RELIDUALNOJ MA-
*              TRICI OZNACAVAJU BROJ KANALA KOJI, AKO JE MODEL ISPRAVAN,
*              NEDOSTAJU SVAKOM PARU CVOROVA; POZITIVNE VRIJEDNOSTI U
*              REZIDUALNOJ MATRICI OZNACAVAJU BROJ KANALA KOJI JE SUVISAN
*              ZA SVAKI PAR CVOROVA.
*
MULT(A=F.TA,B=F,M=GTEOR)
LINEAR(A=G,B=GTEOR,CB=-1.0,M=GRES)
```

```
PRINT(MATRIX=F,T,TEXT=FAKTORI MATRICE SKALARNIH PRODUKATA ULAZNIH VEKTORA)
PRINT(MATRIX=GTEOR,TEXT=REPRODUKCIJA SKALARNIH PRODUKATA ULAZNIH VEKTORA,D)
PRINT(MATRIX=GRES,TEXT=REZIDUALI SKALARNIH PRODUKATA ULAZNIH VEKTORA,D)
*
*              SEKCIJA 5.
*              FAKTORI MATRICE SKALARNIH PRODUKATA IZLAZNIH VEKTORA I
*              REPRODUKCIJA OVE MATRICE U REDUCIRANOM PROSTORU.
*              KAKO SU DIJAGONALNI CLANOVI MATRICE H JEDNAKI BROJU KANALA
*              KOJI ZAPOCINJU OD SVAKOG CVORA, A VANDIJAGONALNI CLANOVI
*              JEDNAKI BROJU KANALA ZAJEDNICKIH PAROVIMA IZLAZNIH CVOROVA,
*              TO SU CLANOVI REPRODUCIRANE MATRICE JEDNAKI OCEKIVANOM BROJU
*              IZLAZNIH KANALA IZ SVAKOG CVORA, ILI SU, POD MODELOM POD KO-
*              JIM JE DEFINIRAN REDUCIRANI PROSTOR, JEDNAK, ZAJEDNICKOM
*              OCEKIVANOM BROJU IZLAZNIH KANALA ZA PAROVE CVOROVA. NEGATIVNE
*              VRIJEDNOSTI U RELIDUALNOJ MATRICI OZNACAVAJU TEORIJSKI NEDOSTA-
*              TAK, A POZITIVNE VRIJEDNOSTI U TOJ MATRICI TEORIJSKI SUVISNE
*              IZLAZNIH KANALA IZ CVOROVA ILI PAROVA CVOROVA.
*
MULT(A=BTX,B=BTX,TB,M=HTEOR)
LINEAR(A=H,B=HTEOR,CB=-1.0,M=HRES)
PRINT(MATRIX=BTX,TEXT=FAKTORI MATRICE SKALARNIH PRODUKATA IZLAZNIH VEKTORA)
PRINT(MATRIX=HTEOR,TEXT=REPRODUKCIJA SKALARNIH PRODUKATA IZLAZNIH VEKTORA,D)
PRI  (MATRIX=HRES,TEXT=REZIDUALI SKALARNIH PRODUKATA IZLAZNIH VEKTORA,D)
*
*              SEKCIJA 6.
*              GROZDOVI ULAZNIH CVOROVA.
*              FORMIRANJE GROZDOVA IZVEDENO JE EKSTREMIZIRANJEM NORMALIZIRANE
*              VARIMAX FUNKCIJE NA DESNIM SVOJSTVENIM VEKTORIMA MATRICE PODATAKA.
*
TRANSPOSE(OLD=Y,NEW=YT)
VARIMAX(F=YT,TAU=TY,FN=YQ,N)
PRINT(MATRIX=YQ,T,TEXT=GROZDOVI ULAZNIH CVOROVA)
*
*              SEKCIJA 7.
*              GROZDOVI IZLAZNIH CVOROVA.
*              FORMIRANJE GROZDOVA IZVEDENO JE EKSTREMIZIRANJEM NORMALIZIRANE
*              VARIMAY FUNKCIJE NA LIJEVIM SVOJSTVENIM VEKTORIMA MATRICE PODATAKA
*
TRANSPOSE(OLD=X,NEW=XT)
VARIMAX(F=XT,TAU=TX,FN=XQ,N)
PRINT(MATRIX=XQ,T,TEXT=GROZDOVI IZLAZNIH CVOROVA)
*
*              SEKCIJA 8.
*              RELACIJE GROZDOVA ULAZNIH CVOROVA.
*              RELACIJE SU DEFINIRANE KAO KOSINUSI PUTEVA VEKTORA KOJI DEFINIRAJU
*              GROZDOVE ULAZNIH CVOROVA U PROSTORU IZLAZNIH CVOROVA
*
MULT(A=TY,B=L,M=TYL)
MULT(A=TYL,B=TY,TB,M=CY)
SCALE(C=CY,R=MY)
PRINT(MATRIX=MY,TEXT=RELACIJE GROZDOVA ULAZNIH CVOROVA)
*
*              SEKCIJA 9.
*              RELACIJE GROZDOVA IZLAZNIH CVOROVA.
*              RELACIJE SU DEFINIRANE KAO KOSINUSI KUTEVA VEKTORA KOJI DEFINIRAJU
*              GROZDOVE IZLAZNIH CVOROVA U PROSTORU ULAZNIH CVOROVA.
*
MULT(A=TX,B=L,M=TXL)
MULT(A=TXL,B=TX,TB,M=CX)
SCALE(C=CX,R=MX)
PRINT(MATRIX=MX,TEXT=RELACIJE GROZDOVA IZLAZNIH CVOROVA)
*
*              SEKCIJA 10.
*              RELACIJE ULAZNIH I IZLAZNIH CVOROVA
```

```
*
MULT(A=YQ,B=XQ,TB,M=COSIO)
PRINT(MATRIX=COSIO,TEXT=RELACIJE ULAZNIH I IZLAZNIH CVOROVA)
*
*            KRAJ PROGRAMA.
*
```

ALGORITAM ZA DETEKCIJU KLIKA NA OSNOVU STRUKTURE KOMUNIKACIJSKIH MREŽA"

Konstantin Momirović, Ankica Hošek,*
Ksenija Bosnar i Franjo Prot

Institut za kineziologiju Fakulteta za fizičku kulturu u Zagrebu
Odjel za informatiku i statistiku
*Sveučilišni računski centar u Zagrebu
Odjel za informatičku matematiku i statistiku

Predložen je algoritam i napisan program za detekciju klika na osnovu strukture komunikacijske mreže definirane binarnom reprezentacijom komunikacijskih kanala izmedju čvorova. Algoritam odredjuje klike parsimonijskim ortonormalnim transformacijama informatički značajnih lijevih i desnih svojstvenih vektora pridruženih nadprosječnim vrijednostima spektra mreže tako da jedna solucija odredjuje klike na temelju ulaznih, a druga na temelju izlaznih kanala. Ponašanje algoritma prikazano je na podacima dobijenim sociometrijskim ispitivanjima jedne poluzatvorene grupe.

1. UVOD

Medju uistinu impresivnom kolekcijom heurističkih postupaka za analizu sociometrijskih podataka* predloženi su i neki čija je svrha detekcija klika (Katz, 1947; Festinger, Schackter i Beck, 1950; Coleman, 1964; Festinger, 1950; Luce i Perry, 1950; Nehnevajsa, 1973)** na temelju operacija nad sociometrijskim matricama. Najveći se broj tih postupaka svodi na potenciranje matrica, ili na prostu dekompoziciju tih matrica.

Detekcija klika je, naravno, taksonomski problem, i može biti rješavan efikasnim taksonomskim algoritmima. U ovom je radu predložen jedan takav algoritam koji odredjuje klike na temelju komunikacija izmedju entiteta parsimonijskom transformacijom lijevih i desnih svojstvenih vektora binarne matrice koja opisuje strukturu komunikacijskih kanala.

---

* *Izvrstan i vrlo sistematičan pregled tih postupaka vidi u I. Nehnevajsa, Soziometrie (In R. König, Grundlegende Methoden und Techniken der empirischen Sozialforschung, 1° Teil, Band 2, pp. 268-285, Enke, Stuttgart, 1973).*

** *Opis ovih postupaka i popis referenci vidi u König, 1973.*

" *Ovaj rad će biti objavljen u časopisu Kineziologija, 16 (1984), 2, u tisku*

2. ALGORITAM

Neka je $S = \{s_i, \ i = 1,\ldots,n\}$ grupa koju tvori n entiteta povezanih odredjenim brojem jednosmjernih ili dvosmjernih komunikacijskih kanala. Neka su intenzitet, frekvencija, opseg i sadržaj komunikacija izmedju entiteta irelevantne varijable, tako da se struktura mreže koju tvore entiteti povezani komunikacijskim kanalima može svesti na binarni graf mreže.

Neka je $M = (m_{ij}); \ i = 1,\ldots,n \ ; \ j = 1,\ldots,n$ matrica kojom je opisana struktura mreže i neka je ta matrica orijentirana tako da retci definiraju ulazne, a kolone izlazne kanale čvorova. Matrica M bit će općenito asimetrična, pozitivno semidefinitna matrica sa strukturom

$$M = X\Lambda X^{-1}$$

gdje je Λ dijagonalna matrica spektralnih vrijednosti, a X matrica njima pridruženih vektora.

Razmotrimo, medjutim, matrice

$$R = M^T M$$

i

$$G = MM^T.$$

Kako je M pozitivno definitna, ili pozitivno semidefinitna matrica, njena se struktura može napisati u obliku

$$M = Y\Lambda X^T$$

gdje su Y i X matrice ortogonalnih svojstvenih vektora, normiranih tako da je $Y^TY = I$ i $X^TX = I$. Otuda

$$R = X\Lambda^2 X^T$$

i

$$G = Y\Lambda^2 Y^T$$

(Green i Carrol, 1976; Fishera, 1978; Horst, 1965) pa se lijevi svojstveni vektori, pridruženi nenultim svojstvenim vrijednostima matrice M mogu definirati kao

$$Y = MX\Lambda^{-1}$$

ako su vektori x definirani na matrici R*.

Očito, u velikoj dijagonali matrice R je broj izlaznih kanala iz svakog čvora, a u vandijagonalnim članovima broj izlaznih kanala zajedničkih svakom paru čvorova. Naravno, u velikoj dijagonali matrice G je broj ulaznih kanala u svaki čvor, a u vandijagonalnim članovima broj ulaznih kanala zajedničkih parovima čvorova.

Koordinatni sistem definiran lijevim svojstvenim vektorom y matrice M odredjuje stoga položaj čvorova u prostoru ulaznih komunikacijskih kanala, a koordinatni sistem definiran desnim svojstvenim vektorima x matrice M odredjuje položaj čvorova u prostoru izlaznih komunikacijskih kanala.

Ovi su koordinatni sustavi, naravno, M - biortogonalni, jer

$$Y^T M X = X^T M^T Y = \Lambda.$$

Poredajmo svojstvene vrijednosti matrica R i G tako da je $\lambda_1^2 > \lambda_2^2 > \ldots > \lambda_q^2$, gdje je q posljednja nenulta svojstvena vrijednost, i organizirajmo tako i njima pridružene svojstvene vektore u matricama Y i X. Sada je matrice R i G moguće napisati kao zbroj matrica ranga jedan, dekomponiran u dvije komponente

$$R = \sum_{p=1}^{k} \lambda_p^2 x_p x_p^T + \sum_{p=k+1}^{q} \lambda_p^2 x_p x_p^T$$

odnosno

$$G = \sum_{p=1}^{k} \lambda_p^2 y_p y_p^T + \sum_{p=q+1}^{q} \lambda_p^2 y_p y_p^T .$$

Matrice

$$R^* = \sum_{p=1}^{k} \lambda_p^2 x_p x_p^T$$

---

* *Naravno, osnov solucije može biti i matrica* G, *ili, jednostavno, spektralna dekompozicija matrice* M

odnosno

$$G^* = \sum_{p=1}^{k} \lambda_p^2 Y_p Y_p^T$$

su, za neki fiksni k, optimalna aproksimacija matrica R i G singularnim matricama R^* i G^* ranga $k < q < n$, pa je stoga i matrica

$$M^* = \sum_{p=1}^{k} \lambda_p Y_p X_p^T$$

optimalna, za neki fiksni rang k, Eckart-Youngova aproksimacija matrice M(Eckart i Young, 1936).

Neka je $k < q$ izabran na neki pogodan način*, i neka je u matricama Y^* i X^* zadržano prvih k svojstvenih vektora, pridruženih zadržanim svojstvenim vrijedstima organiziranim u dijagonalnoj matrici Λ^*. Sada je

$$R^* = X^* \Lambda^{*2} X^{*T}$$

$$G^* = Y^* \Lambda^{*2} Y^{*T}$$

i

$$M^* = Y^* \Lambda^* X^{*T}.$$

Ako čvorovi tvore klike, tj. nakupine čvorova povezanih velikim brojem komunikacijskih kanala, sa slabim vezama sa ostalim čvorovima ili drugim klikama, te je klike moguće detektirati nekom parsimonijskom transformacijom matrica Y^* i X^*.

Neka su Q i T ortonormalne transformacijske matrice odredjene tako, da elementi matrica

$$P = Y^* Q = (p_{ip}) \qquad i = 1,\dots,n \quad p = 1,\dots,k$$

i

---

* *U programu ITA primijenjena je jednostavna Guttman-Kaiserova strategija:*

$$k = num(\lambda_p^2 > \sum_{p=1}^{q} \lambda_p^2 / q)$$

$$A = X^*T = (a_{ip}) \qquad i = 1,\ldots,n \quad p = 1,\ldots,k$$

ekstremiziraju neku parsimonijsku funkciju*.

Ako klike postoje, i ako je k odredjen sukladno broju klika, u vektorima p_p, $p = 1,\ldots,k$ matrice P imat će elemente značajno različite od nule čvorovi koji tvore klike na temelju ulaznih kanala, a u vektorima A_p, $p = 1,\ldots,k$ matrice A imat će elemente značajno različite od nule čvorovi koji tvore klike na temelju izlaznih kanala.

Vektori u matricama P i A i dalje su, naravno, ortogonalni. Medjutim, ta dva sistema više nisu biortogonalna, jer

$$P^TA = Q^T\Lambda T$$

što u općem slučaju nije dijagonalna matrica.

Ako projiciramo vektore ulaznih kanala u sistem definiran grozdovima izlaznih kanala, i ako projiciramo vektore izlaznih kanala u sustav definiran grozdovima ulaznih kanala, tj. ako izvedemo operacije

$$K = MX^*T$$

i

$$L = M^TY^*Q$$

možemo utvrditi relacije grozdova ulaznih kanala operacijom

$$L^TL = Q^T\Lambda^2Q = W$$

i relacije grozdova izlaznih kanala operacijom

$$K^TK = T^T\Lambda^2T = V.$$

Neka je $D_W = (\text{diag}W)^{-1/2}$ i neka je $D_V = (\text{diag}V)^{-1/2}$.

---

* *U programu ITA ekstremizirana je Kaiserova (1958) brutto varimax funkcija*

$$w = n\sum_{i=1}^{n}\sum_{p=1}^{k} e_{ip}^4 - \sum_{p=1}^{k}\left(\sum_{i=1}^{n} e_{ip}^2\right)^2$$

za $e_{ip} = p_{ip}$, a_{ip} .

Sada su kosinusi kuteva vektora grozdova ulaznih kanala u prostoru izlaznih kanala elementi matrice

$$C_L = D_W W D_W$$

a kosinusi kuteva vektora grozdova izlaznih kanala u prostoru ulaznih kanala elementi matrice

$$C_K = D_V V D_V \ .$$

Ova operacija omogućuje i odredjivanje kosinusa kuteva vektora grozdova ulaznih i izlaznih kanala jer je

$$L^T K = Q^T \Lambda^3 T$$

a kosinusi kuteva vektora grozdova definiranih na temelju ulaznih i izlaznih kanala elementi matrice

$$C_{LK} = D_W Q^T \Lambda^3 T D_V \ .$$

3. PROGRAM ITA

Program ITA napisan je u verziji 5 SS jezika i implementiran u programskim bibliotekama SRCE*SS-MAKRO* i FFK*LIB**. Program dopušta grupe do 250 entiteta i pretpostavlja da su u retcima matrice podataka vektori ulaznih, a u kolonama te matrice vektori izlaznih kanala. Imenima entiteta u zapisu koji sadrži podatke valja pridružiti oznaku da se radi o ulaznim kanalima. Izlazne kanale program tretira kao varijable, pa imena varijabli moraju biti identična imenima entiteta kojima je pridružena oznaka da se radi o izlaznim kanalima.

---

* *Javna biblioteka Sveučilišnog računskog centra u Zagrebu*

** *Zatvorena biblioteka Računskog centra Fakulteta za fizičku kulturu u Zagrebu*

4. REZULTATI ANALIZE JEDNE MALE POLUZATVORENE GRUPE

Analizirana je struktura jedne male grupe čija je djelatnost u velikoj mjeri ovisila o kooperaciji članova grupe*. Entiteti su tretirani kao čvorovi mreže, a komunikacijski kanal odredjenog smjera smatran je otvorenim ako je intenzitet komunikacija izmedju ma koja dva čvora prešao neku liminalnu vrijednost.

U tabeli 1 dati su rezultati dobijeni analizom strukture ulaznih, a u tabeli 2 rezultati dobijeni analizom strukture izlaznih kanala.

Tabela 1
FORMACIJA KLIKA NA OSNOVU ULAZNIH KANALA

	U VRX 1	U VRX 2	U VRX 3
U 07	-.0306	(.3699)	-.1245
U 08	.0687	(.5138)	-.0362
U 09	-.1498	(.3089)	.1957
U 10	.0142	-.0568	(.4707)
U 11	-.0646	.1734	(.5500)
U 12	.0125	-.0552	(.4623)
U 13	(.4968)	-.1731	.2527
U 14	(.5607)	.0589	-.2463
U 17	(.5976)	.0762	.0900
U 19	.0017	(.4440)	.1845
U 20	-.0306	(.3699)	-.1245
U 21	.0798	-.0895	(.1545)
U 02	.2045	(.2937)	-.0324

---

* *Priroda djelatnosti grupe bila je takva da su članovi grupe najveći dio ne samo radnog, već i slobodnog vremena morali provoditi zajedno.*

Tabela 2

FORMACIJA KLIKA NA OSNOVU IZLAZNIH KANALA

	K VRX 1	K VRX 2	K VRX 3
K2 07	-.0845	(.5171)	-.0438
K2 08	.0443	(.5980)	.1057
K2 09	(.3535)	.2950	-.1660
K2 10	(.3884)	.2941	-.1583
K2 11	(.4786)	-.1504	.0441
K2 12	(.3797)	-.0915	.2842
K2 13	(.5617)	-.0764	-.0036
K2 14	-.0472	.1483	(.3414)
K2 17	-.0279	.0196	(.4611)
K2 19	.0126	.1285	(.4070)
K2 20	-.0945	(.3222)	.1202
K2 21	.1168	-.1405	(.2801)
K2 02	-.0247	-.0467	(.5137)

Tri klike formirane na osnovu ulaznih kanala u nenegativnim su medjusobnim vezama (tabela 3), no te su veze niže od osrednjih. Prvu kliku tvore entiteti koji su u grupu došli relativno kasno i ne potiču iz sredine iz koje potiču ostali članovi grupe. Druga, po broju najveća klika sastavljena je od entiteta superiorne profesionalne efikasnosti. Treću kliku tvore tri entiteta sa vrlo visokim nivoom aspiracije i naglim trendom postizanja vrhunske profesionalne efikasnosti. Posljednja dva entiteta od kojih prvi pripada trećoj, a drugi drugoj kliki zapravo su bili autsajderi pod vidom njihove profesionalne efikasnosti u vrijeme kada je provedeno ovo istraživanje.

Klike formirane na temelju izlaznih kanala suštinski se razlikuju od klika formiranih na temelju ulaznih kanala. Relacije izmedju ovako formiranih klika su takodjer nenegativne, ali i dalje jedva osrednje (tabela 4). Kriterij na osnovu kojeg su se formirale klike bio je različit od kriterija koji je formirao klike na osnovu ulaznih kanala. Prvu kliku tvorila je relativno homogena grupa entiteta suprotstavljenih politici rukovodstva grupe. Drugu, slabu kliku, sačinjavalo je nekoliko entiteta sa izrazitom sklonošću ka kooperaciji. Treću kliku sačinjavao je ostatak grupe, koji se prema rukovodstvu odnosio neutralno.

Tabela 3

RELACIJE KLIKA ODREDJENIH NA OSNOVU ULAZNIH KANALA

	U VRX 1	*U VRX 2*	*U VRX 3*
U VRX 1	*1.0000*	*.5313*	*.5789*
U VRX 2	*.5313*	*1.0000*	*.3741*
U VRX 3	*.5789*	*.3741*	*1.0000*

Tabela 4

RELACIJE KLIKA ODREDJENIH NA OSNOVU IZLAZNIH KANALA

	K VRX 1	*K VRX 2*	*K VRX 3*
K VRX 1	*1.0000*	*.4387*	*.5923*
K VRX 2	*.4387*	*1.0000*	*.4822*
K VRX 3	*.5923*	*.4822*	*1.0000*

Profesionalna efikasnost grupe u vrijeme kada je provedena ova analiza bila je vrlo visoka, a standardne sociometrijske procedure govorile su u prilog hipotezi da je kolektiv vrlo homogen. Medjutim, samo godinu dana kasnije došlo je do otvorenog sukoba izmedju prve klike formirane na osnovu izlaznih kanala i rukovodstva grupe. Dio članova grupe napustio je kolektiv, a rukovodstvo grupe bilo je uklonjeno. Heterogenost kolektiva postala je sasvim očita, a profesionalna efikasnost je drastično opala.

Čini se prema tome, da ovaj algoritam može biti od neke koristi za detekciju klika, posebno u slučajevima kada postojanje tih klika ne može biti utvrdjeno jednostavnom inspekcijom sociometrijskih matrica, ili primjenom uobičajenih procedura za detekciju klika koje se temelje na potenciranju sociometrijskih matrica.

LITERATURA

1. Eckart, C. and G. Young. *The approximation of one matrix by another of lower rank.* Psychometrika, 1(1936),211-218.
2. Fishera, G. *Numerical and quantitative analysis.* Pitman, London, 1978.
3. Green, P.E. and J. D. Carroll. *Mathematical tools for applied multivariate analysis.* Academic Press, New York, 1976.
4. Horst, P. *Factor analysis of data matrices.* Holt, Rinehart and Winston, New York, 1965.
5. Kaiser, H. F. *The varimax criterion for analytic rotation in factor analysis.* Psychometrika, 23(1958), 187-200.
6. König, R. *Grundlgende Methoden und Techniken der empirischen Sozialforschung.* 1° Teil, Band 2, 3° Auflage. Enke Verlag, Stuttgart, 1973.

2.7 ALGORITAM I PROGRAM COMPAT-ZZ

Tokom trenažnog procesa javljaju se faze, poznate pod imenom platoi krivulje učenja ili vježbe, gdje kvantitativne promjene u dimenzijama psihosomatskog statusa i promjene u broju motoričkih informacija ne dovode do zamjetnog povećanja uspjeha takmičara. Poboljšanje rezultata može se očekivati tek nakon prestrukturiranja dimenzija i/ili informacija na kvalitativno višem nivou. Za praćenje procesa treninga su, stoga, podjednako važni podaci o strukturalnim promjenama nakon nekog perioda treninga kao i kvantitativni pokazatelji. Stoga, predložen je algoritam i program COMPAT-ZZ koji usporedjuje strukture latentnih dimenzija varijabli izmjerenih na početku i na kraju nekog trenažnog ciklusa i analizira eventualne strukturalne promjene.

```
*
*    NAZIV PROGRAMA
*
*
*                *** COMPAT-ZZ ***
*
*
*    AUTORI
*
*              K. MOMIROVIC
*              V. DOBRIC
*              Z. KARAMAN
*
*
*    FUNKCIJA
*
*         PROGRAM COMPAT-ZZ ANALIZIRA, POD KOMPONENTNIM MODELOM, KOMPATIBILNOST
*         FAKTORSKIH STRUKTURA DVA ZAVISNA UZORKA ENTITETA OPISANA NAD IDEN-
*         TICNIM SKUPOM KVANTITATIVNIH VARIJABLI U METRICI KOJA JE DEFINIRANA
*         NA POPULACIJI KOJOJ PRIPADA PRVI UZORAK. STRUKTURE KOJE
*         COMPAT-ZZ IDENTIFICIRA I USPOREDJUJE DEFINIRANE SU KAO GLAVNE OSOVINE
*         MATRICA KOVARIJANCI I IZ NJIH IZVEDENI ORTHOBLIQUE FAKTORI. ANALI-
*         ZU KOMPATIBILNOSTI COMPAT IZVODI ODREDJIVANJEM KONGRUENCIJE
*         SKLOPOVA I KORELACIJA LATENTNIH DIMENZIJA.
*
*
*    METODA I ALGORITAM OPISANI SU U
*
*         MULAIK (1972), FULGOSI (1979), MOMIROVIC (1979,1980),
*
*              I U RADU
*
*         DOBRIC,V., Z.KARAMAN I K.MOMIROVIC:
*         JEDNOSTAVNI ALGORITAM ZA ANALIZU STRUKTURALNIH PROMJENA,
*         KINEZIOLOGIJA,16,1(1984), U TISKU
*
*
*    PRIPREMA PODATAKA
*
*          KONTROLNI ZAPIS I PODACI NALAZE SE U DATOTECI  DATA.
*
*    IZVRSNI DIO PROGRAMA
*
*          BLOK 0. ULAZNE OPERACIJE I REPARAMETRIZACIJA VARIJABLI.
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=C O M P A T - ZZ,T)
HEADING(TEXT=ANALIZA KOMPATIBILNOSTI FAKTORSKIH STRUKTURA)
HEADING(TEXT=MODEL 2 - ZAVISNI UZORCI U PROSTORU SA ISTOM METRIKOM,D)
TEXT(TEXT=                 COMPAT      ZZ)
*
HEADING(TEXT=REPARAMETRIZACIJA VARIJABLI)
INPUT(DATA=DATA,SCORE=S1)
REWIND(FILE=DATA)
INPUT(DATA=DATA,SCORE=B2)
*CONFORM(IN1=SS1,IN2=SS2,OUT1=S1,OUT2=B2)
CORRELATION(SCORE=S1,C=C1,R=R1)
CORRELATION(SCORE=B2,C=C2,R=R2)
          DELETE(MATRIX=R2)
STATISTICS(SCORE=B2,Z=S2,S)
DIAGMULT(A=S2,D=C2,C=0.5,R,M=ZZ2)
DIAG(A=C1,D=DC1)
DIAG(A=C2,D=DC2)
DIAGMULT(A=DC1,D=DC2,C=-1.0,R,M=PND)
```

```
*
*          BLOK 1.DISTRIBUCIJE I PARAMETRI VARIJABLI U PRVOM UZORKU.
*
HEADING(TEXT=DISTRIBUCIJE I PARAMETRI VARIJABLI)
HEADING(TEXT=PRVI UZORAK,D)
STATISTICS(SCORE=S1,CLASS=9,Z=Z1,S)
            DELETE(MATRIX=S1)
PRINT(MATRIX=R1,TEXT=KOVARIJANCE VARIJABLI U PRVOM UZORKU)
INVERSION(R=R1,RINV=RINV1,SMC=SMC1)
*
*          BLOK 2. DISTRIBUCIJE I PARAMETRI VARIJABLI U DRUGOM UZORKU.
*
HEADING(TEXT=DISTRIBUCIJE I PARAMETRI VARIJABLI)
HEADING(TEXT=DRUGI UZORAK,D)
DIAGMULT(A=ZZ2,D=C1,C=-0.5,R,M=Z2)
            DELETE(MATRIX=ZZ2)
STATISTICS(SCORE=Z2,CLASS=9,Z=ZZ2,S)
            DELETE(MATRIX=ZZ2)
CORRELATION(SCORE=Z2,R=RR2,C=R2)
            DELETE(MATRIX=RR2)
PRINT(MATRIX=R2,TEXT=KOVARIJANCE VARIJABLI U DRUGOM UZORKU)
INVERSION(R=R2,RINV=CINV2,SMC=SMC2)

*          BLOK 3. GLAVNE KOMPONENTE U PRVOM UZORKU.
*
HEADING(TEXT=GLAVNE KOMPONENTE)
HEADING(TEXT=PRVI UZORAK,D)
DIAGONALISATION(R=R1,X=X1,LAMBDA=L1)
HOTELLING(LAMBDA=L1,X=X1,F=H1,PBC,SMC=SMC1)
            DELETE(MATRIX=L1)
            DELETE(MATRIX=X1)
MULT(A=H1,B=H1,TB,M=L1)
DIAGMULT(A=H1,D=L1,C=-0.5,L,M=X1)
PRINT(MATRIX=H1,T,TEXT=GLAVNE OSOVINE U PRVOM UZORKU)
MULT(A=Z1,B=X1,TB,M=K1)
STATISTICS(SCORE=K1,CLASS=9,Z=ZK1,S)
            DELETE(MATRIX=ZK1)
*
*          BLOK 4. GLAVNE KOMPONENTE U DRUGOM UZORKU.
*
HEADING(TEXT=GLAVNE KOMPONENTE)
HEADING(TEXT=DRUGI UZORAK,D)
DIAGONALISATION(R=R2,X=X2,LAMBDA=L2)
HOTELLING(LAMBDA=L2,X=X2,F=H2,PBC,SMC=SMC2)
            DELETE(MATRIX=L2)
            DELETE(MATRIX=X2)
MULT(A=H2,B=H2,TB,M=L2)
DIAGMULT(A=H2,D=L2,C=-0.5,L,M=X2)
PRINT(MATRIX=H2,T,TEXT=GLAVNE OSOVINE U DRUGOM UZORKU)
DIAGMULT(A=H2,T,D=PND,L,M=SH2)
PRINT(MATRIX=SH2,TEXT= STANDARDIZIRANE OSOVINE DRUGOG UZORKA)
MULT(A=Z2,B=X2,TB,M=K2)
STATISTICS(SCORE=K2,CLASS=9,Z=ZK2,S)
            DELETE(MATRIX=ZK2)
*
*          BLOK 5. KONGRUENCIJA GLAVNIH OSOVINA.
*
HEADING(TEXT=KOMPATIBILNOST GLAVNIH OSOVINA)
HEADING(TEXT=KONGRUENCIJA GLAVNIH OSOVINA,D)
MULT(A=X1,B=X2,TB,M=CONG12)
PRINT(MATRIX=CONG12,TEXT=KOEFICIJENTI KONGRUENCIJE GLAVNIH OSOVINA)
*
*          BLOK 6. KORELACIJE GLAVNIH KOMPONENATA.
*
```

```
HEADING(TEXT=KORELACIJE GLAVNIH KOMPONENATA,D)
CROSSCORRELATION(P1=K1,P2=K2,R12=R12,C12=W12)
PRINT(MATRIX=W12,TEXT=KOVARIJANCE GLAVNIH KOMPONENATA PRVOG I DRUGOG UZORKA)
PRINT(MATRIX=R12,TEXT=KORELACIJE GLAVNIH KOMPONENATA PRVOG I DRUGOG UZORKA)
*
*          BLOK 7. ORTHOBLIQUE FAKTORI U PRVOM UZORKU.
*
HEADING(TEXT=ORTHOBLIQUE FAKTORI)
HEADING(TEXT=PRVI UZORAK,D)
VARIMAX(F=X1,TAU=Q1,FN=AA1)
MULT(A=Q1,B=L1,M=QL1)
MULT(A=QL1,B=Q1,TB,M=MM1)
DIAGMULT(A=AA1,D=MM1,L,C=0.5,M=A1)
SCALE(C=MM1,R=M1)
MULT(A=M1,B=A1,M=F1)
PRINT(MATRIX=Q1,T,TEXT=ORTHOBLIQUE TRANSFORMACIJSKA MATRICA U PRVOM UZORKU)
PRINT(MATRIX=A1,T,TEXT=SKLOP ORTHOBLIQUE FAKTORA U PRVOM UZORKU)
PRINT(MATRIX=MM1,TEXT=KOVARIJANCE ORTHOBLIQUE FAKTORA U PRVOM UZORKU)
PRINT(MATRIX=M1,TEXT=KORELACIJE ORTHOBLIQUE FAKTORA U PRVOM UZORKU,D)
PRINT(MATRIX=F1,T,TEXT=STRUKTURA ORTHOBLIQUE FAKTORA U PRVOM UZORKU,D)
MULT(A=Z1,B=AA1,TB,M=PSI1)
STATISTICS(SCORE=PSI1,CLASS=9,Z=ZPSI1,S)
            DELETE(MATRIX=ZPSI1)
            DELETE(MATRIX=QL1)
*
*          BLOK 8. ORTHOBLIQUE FAKTORI U DRUGOM UZORKU.
*
HEADING(TEXT=ORTHOBLIQUE FAKTORI)
HEADING(TEXT=DRUGI UZORAK,D)
VARIMAX(F=X2,TAU=Q2,FN=AA2)
MULT(A=Q2,B=L2,M=QL2)
MULT(A=QL2,B=Q2,TB,M=MM2)
DIAGMULT(A=AA2,D=MM2,L,C=0.5,M=A2)
SCALE(C=MM2,R=M2)
MULT(A=M2,B=A2,M=F2)
PRINT(MATRIX=Q2,T,TEXT=ORTHOBLIQUE TRANSFORMACIJSKA MATRICA U DRUGOM UZORKU)
PRINT(MATRIX=A2,T,TEXT=SKLOP ORTHOBLIQUE FAKTORA U DRUGOM UZORKU)
PRINT(MATRIX=MM2,TEXT=KOVARIJANCE ORTHOBLIQUE FAKTORA U DRUGOM UZORKU)
PRINT(MATRIX=M2,TEXT=KORELACIJE ORTHOBLIQUE FAKTORA U DRUGOM UZORKU,D)
PRINT(MATRIX=F2,T,TEXT=STRUKTURA ORTHOBLIQUE FAKTORA U DRUGOM UZORKU,D)
DIAGMULT(A=A2,T,D=PND,L,M=SA2)
PRINT(MATRIX=SA2,TEXT= STANDARDIZIRANI SKLOP DRUGOG UZORKA)
DIAGMULT(A=F2,T,D=PND,L,M=SF2)
PRINT(MATRIX=SF2,TEXT= STANDARDIZIRANA STRUKTURA DRUGOG UZORKA)
MULT(A=Z2,B=AA2,TB,M=PSI2)
STATISTICS(SCORE=PSI2,CLASS=9,Z=ZPSI2,S)
            DELETE(MATRIX=ZPSI2)
            DELETE(MATRIX=QL2)
*
*          BLOK 9. KONGRUENCIJA ORTHOBLIQUE SKLOPOVA.
*
HEADING(TEXT=KOMPATIBILNOST ORTHOBLIQUE FAKTORA)
HEADING(TEXT=KONGRUENCIJA ORTHOBLIQUE SKLOPOVA,D)
MULT(A=Q1,B=CONG12,M=BIK)
MULT(A=BIK,B=Q2,TB,M=COMP12)
PRINT(MATRIX=COMP12,TEXT=KOEFICIJENTI KONGRUENCIJE ORTHOBLIQUE FAKTORA)
            DELETE(MATRIX=BIK)
*
*          BLOK 10. KORELACIJE ORTHOBLIQUE FAKTORA.
*
HEADING(TEXT=KORELACIJE ORTHOBLIQUE FAKTORA,D)
MULT(A=Q1,B=W12,M=SVINJA)
MULT(A=SVINJA,B=Q2,TB,M=C12)
            DELETE(MATRIX=SVINJA)
```

```
DIAGMULT(A=C12,D=MM1,C=-0.5,L,M=KRAVA)
DIAGMULT(A=KRAVA,D=MM2,C=-0.5,R,M=M12)
PRINT(MATRIX=C12,TEXT=KOVARIJANCE ORTHOBLIQUE FAKTORA PRVOG I DRUGOG UZORKA)
PRINT(MATRIX=M12,TEXT=KORELACIJE ORTHOBLIQUE FAKTORA PRVOG I DRUGOG UZORKA)
*
*         KRAJ PROGRAMA, NAPOMENE:
*      (1) IMENA VARIJABLI ZA OBA UZORKA MORAJU BITI POSVE IDENTICNA.
*          REDOSLIJED VARIJABLI U OBA SKUPA MORA BITI IDENTICAN, ILI
*          SE TE MATRICE, NAKON TRANSPOZICIJE, MORAJU SORTIRATI, PONOVO
*          TRANSPONIRATI I TEK  ZATIM PODVRCI EGZEKUTIVNIM NAREDBAMA U
*          BLOKU 0. OVO KORISNIK MOZE UCINITI DOPISAVSI TE NAREDBE U BLOK
*          0., NO RAZUMNIJE JE DA OD POCETKA KOREKTNO PRIPREMI PODATKE.
*      (2) COMPAT-ZZ MOZE, ZA SADA, ANALIZIRATI UZORKE OD PO 10000 ENTITETA,
*          OPISANE SA DO 250 VARIJABLI KOJE NE SMIJU BITI IZVORNO STANDAR-
*          DIZIRANE.
*      (3) BROJ FAKTORA ODREDJEN JE, U SVAKOM UZORKU POSEBNO I NEZAVISNO,
*          NA TEMELJU PB KRITERIJA (STALEC I MOMIROVIC,1971) PA NE MORA BITI
*          JEDNAK. AKO JE OVO ZBOG NECEGA KORISNIKU NEUGODNO, A IMA NEKU PA-
*          METNU HIPOTEZU O BROJU FAKTORA, NEKA U NAREDBAMA HOTELLING U
*          BLOKOVIMA 3. I 4. IZBACI SIGNAL PBC, I OZNAKE SMC=SMC1/2, I
*          BROJ FAKTORA DEFINIRA OPCIJOM NUM=K, GDJE JE K HIPOTETSKI BROJ
*          FAKTORA. U TOM SLUCAJU NISU POTREBNE NAREDBE INVERSION U BLO-
           KOVIMA 1. I 2.
*      (4) COMPAT-ZZ, ZA PROBLEME NETRIVIJALNIH DIMENZIJA, TRAJE RELATIVNO
*          DUGO, A IMA I PRILICNO VELIKI ISPIS, ZBOG TOGA JE RAZUMNO AKTI-
*          VIRATI OVAJ PROGRAM SAMO ZA ANALIZU KOREKTNO DOBIJENIH REZULTATA.
*
HEADING(TEXT=KRAJ PROGRAMA)
*
```

JEDNOSTAVAN ALGORITAM ZA ANALIZU STRUKTURALNIH PROMJENA"

*Vesna Dobrić, Živan Karaman i Konstantin Momirović**
SVEUČILIŠNI RAČUNSKI CENTAR, ZAGREB
Odjel za informatičku matematiku i statistiku
*INSTITUT ZA KINEZIOLOGIJU FAKULTETA ZA FIZIČKU KULTURU U ZAGREBU
Odjel za informatiku i statistiku

Definiran je algoritam i napisan program koji analizira, pod komponentnim modelom, kompatibilnost faktorskih struktura dva zavisna uzorka entiteta opisana nad identičnim skupom kvantitativnih varijabli u metrici koja je definirana na populaciji kojoj pripada prvi uzorak. Strukture koje algoritam identificira i usporedjuje definirane su kao glavne osovine matrica kovarijanci i iz njih izvedeni orthoblique faktori. Analizu kompatibilnosti algoritam izvodi odredjivanjem kongruencije sklopova i kroskorelacija latentnih dimenzija.

1. UVOD

Primjenu faktorske analize za odredjivanje strukturalnih promjena prvi je predložio C. Harris prije više od 20 godina (Harris, 1963). Kasnije su u ovu svrhu bile predložene osjetno složenije metode od jednostavnog Harrisovog kanoničkog modela (longitudinalna faktorska analiza, Corbalis i Traub, 1970; trimodalna faktorska analiza, Tucker, 1963; 1964 i 1966). Meredith (1964) je pokazao da usporedba faktorskih struktura, pa stoga i struktura u raznim vremenskim točkama ima smisla samo ako su varijable, u ma kojoj od analiziranih vremenskih točaka, definirane u istoj metrici;* no ovaj se uvjet ne poštuje uvijek, dijelom stoga što ga neki istraživači ne poznaju, a dijelom stoga što ga ne razumiju.

---

* *vidi o tom opširnije u Mulaik, 1972; Fulgosi, 1979*

" *ovaj rad će biti objavljen u časopisu Kineziologija, 16, (1984), 2, u tisku*

U ovom je radu definiran algoritam i opisan program, napisan u SS jeziku, koji analizira pod komponentnim, a ne pod uobičajenijim faktorskim modelom, kompatibilnost faktorskih struktura dva zavisna uzorka entiteta** opisana nad identičnim skupom kvantitativnih varijabli u metrici koja je definirana na populaciji kojoj pripada prvi uzorak. Strukture koje algoritam identificira i usporedjuje definirane su kao glavne osovine matrica kovarijanci, čiji je broj odredjen jednom generalizacijom PB kriterija (Štalec i Momirović, 1971), i kao iz njih izvedeni orthoblique faktori (Harris i Kaiser, 1964). Analizu kompatibilnosti algoritam izvodi odredjivanjem kongruencije sklopova i izračunavanjem kroskorelacija latentnih dimenzija.

2. ALGORITAM

2.1. *Reparametrizacija varijabli*

Neka su $U_1 = \{e_i^{(1)}\ ;\ i = 1,\ldots,n\}$ i $U_2 = \{e_i^{(2)}\ ;\ i=1,\ldots,n\}$ dva zavisna uzorka s jednakim brojem entiteta (ili jedan uzorak entiteta mjeren u dvije vremenske točke) opisana nad identičnim uzorkom $V = \{v_j;\ j = 1,\ldots,m\}$, $m<n$, kvantitativnih, multivarijatno normalno distribuiranih varijabli. Označimo li sa $B_1 = (b_{ij}^{(1)})$ $i = 1,\ldots,n;\ j = 1,\ldots,m$ i $B_2 = (b_{ij}^{(2)})$, $i = 1,\ldots,n;\ j = 1,\ldots,m$ rezultate entiteta iz uzoraka U_1 i U_2 u varijablama iz skupa V, centriranje rezultata iz matrice B_1, odnosno B_2, postiže se operacijama

$B_1 - 11^T \frac{1}{n} B_1 = S_1$ i $B_2 - 11^T \frac{1}{n} B_2 = S_2$ gdje je

1 n-elementni sumacioni vektor jedinica; maksimalnom vjerojatnošću procijenjene matrice varijanci - kovarijanci

$C_1 = S_1^T S_1 \frac{1}{n}$ i $C_2 = S_2^T S_2 \frac{1}{n}$ su nosioci informacija o medjusobnoj povezanosti varijabli iz skupa V na uzorcima U_1 odnosno U_2.

** *pa stoga i istog uzorka u dvije različite vremenske točke*

U narednom koraku algoritam zahtijeva standardizaciju rezultata $S_1(\text{diag } C_1)^{-\frac{1}{2}} = Z_1$ prvog skupa i restandardizaciju rezultata $S_2(\text{diag } C_1)^{-\frac{1}{2}} = Z_2$ drugog skupa na metriku prvog skupa. Ishodišne matrice za daljnju analizu su matrica korelacija $R_1 = Z_1^T Z_1 \frac{1}{n}$ standardiziranog prvog skupa varijabli i matrica varijanci-kovarijanci $R_2 = Z_2^T Z_2 \frac{1}{n}$ nestandardiziranog drugog skupa rezultata.

2.2. Glavne komponente

Nestandardizirane ortogonalne komponente K_i, $i = 1,2$ restandardiziranih varijabli prvog odnosno drugog uzorka dobiju se uobičajenom procedurom izračunavanja matrica svojstvenih vrijednosti Λ_i, $i=1,2$ i pripadnih svojstvenih vektora X_i, $i = 1,2$ matrica $R_1 = X_1 \Lambda_1 X_1^T$ i $R_2 = X_2 \Lambda_2 X_2^T$ operacijama

$$K_1 = Z_1 X_1$$
$$K_2 = Z_2 X_2 \, .$$

Broj q_i, $i = 1,2$ zadržanih glavnih komponenata prvog, odnosno drugog skupa, odredjuje se jednom generalizacijom PB kriterija:

$$q_i = (\text{núm } \lambda_j \, ; \sum_{j=1}^{q_i - 1} \lambda_j < b_i < \sum_{j=1}^{q_i} \lambda_j)$$

pri čemu je

$$b_i = \text{tr}(R_i + (\text{diag} R_i^{-1})^{-1} R_i^{-1} (\text{diag} R_i^{-1})^{-1} - 2(\text{diag} R_i^{-1})^{-1})$$

mjera zajedničkog varijabiliteta varijabli iz Z_i , $i = 1,2$.

U matricama glavnih osi

$$H_i = X_i \Lambda_i^{\frac{1}{2}} \qquad i = 1,2$$

sadržane su kovarijance varijabli i zadržanih glavnih komponenata.

Kompatibilnost ovako dobijenih faktorskih struktura ispituje se koeficijentima kongruencije glavnih osi

$$\rho_{12} = X_1^T X_2 \; .$$

Povezanost glavnih komponenti prvog i drugog skupa daju njihove kroskovarijance

$$W_{12} = K_1^T K_2 \frac{1}{n}$$

odnosno kroskorelacije

$$R_{12} = \Lambda_1^{-\frac{1}{2}} W_{12} \Lambda_2^{-\frac{1}{2}} \; .$$

2.3. *Orthoblique faktori*

Kako ortogonalni faktori ne definiraju latentnu strukturu nekog skupa varijabli na optimalan način, algoritam definira jedno od mogućih kosih rješenja, postižući parsimonijsku poziciju orthoblique rotacijama zadržanih glavnih komponenti K_1 odnosno K_2. Ortogonalnim rotacijama zadržanih svojstvenih vektora

$$X_i T_i = A_i = (a_{jk})_i \; , \quad i = 1,2 \; ; \; j = 1,\dots,m \; ; \; k = 1,\dots,q$$

gdje su T_i , $i = 1,2$, transformacijske matrice sa svojstvom $T_i T_i^T = T_i^T T_i = I$, maksimizira se dobro poznata varimax funkcija

$$m \sum_{j=1}^{m} \sum_{k=1}^{q} a_{jk}^4 - \sum_{k=1}^{q} \left(\sum_{j=1}^{m} a_{jk}^2 \right)^2$$

što rezultira kosim latentnim dimenzijama

$$\Psi_i = Z_i X_i T_i \qquad i = 1,2$$

prvog, odnosno drugog skupa. Matrice

$$\Psi_i^T \Psi_i \frac{1}{n} = T_i^T \Lambda_i T_i \qquad i = 1,2$$

varijanci - kovarijanci, te matrice

$$M_i = (\text{diag} T_i^T \Lambda_i T_i)^{-\frac{1}{2}} (T_i^T \Lambda_i T_i)(\text{diag}\ T_i^T \Lambda_i T_i)^{-\frac{1}{2}} \qquad i = 1,2$$

korelacija orthoblique latentnih dimenzija prvog, odnosno drugog skupa pokazuju medjusobnu povezanost ovih dimenzija. Matrice sklopa

$$A_i = X_i T_i (\text{diag} T_i^T \Lambda_i T_i)^{-\frac{1}{2}} \qquad i = 1,2$$

sadrže koordinate originalnih varijabli u prostoru kosih dimenzija za prvi, odnosno drugi skup, a matrice strukture

$$F_i = X_i \Lambda_i T_i (\text{diag}\ T_i^T \Lambda_i T_i)^{-\frac{1}{2}} \qquad i = 1,2$$

sadrže kovarijance originalnih varijabli i orthoblique faktora za prvi, odnosno drugi skup.

U posljednjem koraku algoritam ispituje kompatibilnost dobivenih orthoblique faktorskih sklopova prvog i drugog skupa preko Tuckerovih koeficijenata kongruencije

$$\tau_{12} = T_1^T X_1^T X_2 T_2$$

te povezanost orthoblique latentnih dimenzija prvog i drugog skupa preko matrice kroskovarijanci

$$C_{12} = T_1^T W_{12} T_2$$

odnosno matrice kroskorelacija

$$M_{12} = (\text{diag} T_1^T \Lambda_1 T_1)^{-\frac{1}{2}}\ C_{12} (\text{diag} T_2^T \Lambda_2 T_2)^{-\frac{1}{2}} .$$

3. PROGRAM COMPAT-ZZ

Program COMPAT-ZZ napisan je u verziji 5.2/M programskog jezika SS i pohranjen u javnoj biblioteci SRCE*SS-MAKRO. Dopušta da u svakom od zavisnih uzoraka bude do 10 000 entiteta, opisanih sa do 250 kvantitativnih varijabli koje moraju biti zapisane u izvornom obliku, bez prethodnog centriranja i/ili normiranja. Program sam sredjuje zapise u uzorcima; ti zapisi, naravno, moraju imati ista imena entiteta u oba uzorka. Imena varijabli moraju takodjer biti posve identična, i sortirana na isti način u oba uzorka.

COMPAT-ZZ ima 112 linija izvršnog koda i 45 linija komentara; ovi komentari daju sve potrebne informacije o načinu korištenja, opcijama, uvjetima i ograničenjima, tako da nije potrebna nikakva daljnja dokumentacija, osim poznavanja procedura za aktiviranje makroprograma iz biblioteke SRCE*SS-MAKRO*.

4. PONAŠANJE ALGORITMA

Ponašanje algoritma i njemu pridruženog programa COMPAT-Z prikazano je na jednoj analizi društveno-ekonomskih promjena.

Za sve općine u jednoj našoj republici registrirani su, za 1971. i 1981. godinu, ovi indikatori razvijenosti:**

1. Društveni proizvod po stanovniku (DRUPRO)
2. Akumulacija po stanovniku (AKUMUL)
3. Broj zaposlenih na 1000 stanovnika (ZAPOSL)
4. Broj zaposlenih sa višim i visokim obrazovanjem (ZAPVSS)
5. Broj liječnika na 1000 stanovnika (LJEKAR)
6. Potrošnja električne energije po stanovniku (ELENER)
7. Dužina suvremenih puteva na 100 km^2 površine općine (GUSTPU)
8. Vrijednost osnovnih aktivnih sredstava privrede po stanovniku (VRIJED)

---

* *COMPAT-Z se ne smije zamijeniti s SS programom COMPAT koji odredjuje kompatibilnost faktorskih struktura u dva nezavisna uzorka.*

** *U zagradi su mnemoničke šifre varijabli upotrebljavane pri izračunavanju i izradi tabela. Podaci su uzeti iz izvora Zavoda za statistiku. Broj entiteta iznosio je 103.*

9. Broj zaposlenih koji rade van općine u kojoj stanuju (ZAPVOP)
10. Stopa zaposlenosti (ZAPOSS)
11. Broj stanovnika na 1 km^2 (BRSTAN)
12. Urbanizirano stanovništvo (URBANS)
13. Deagrarizirano stanovništvo (DEORGS)

U tabeli 1 su prvi (μ) i drugi moment (σ^2) distribucija varijabli u 1971. i 1981. godini, koeficijent translacije (μ_{1981}/μ_{1971}) i koeficijent dilatacije ($\sigma_{1981}/\sigma_{1971}$).

Kako se vidi iz koeficijenata translacije, stopa rasta pojedinih indikatora bila je sasvim nejednaka. Osnovni indikatori ekonomskog razvoja, kao što su društveni proizvod, akumulacija i potrošnja energije porasli su u ovom periodu 6 do 9 puta, i više se nego udvostručila dužina suvremenih puteva; medjutim, vrijednost osnovnih aktivnih sredstava porasla je samo za 69%. Promjene u demografskim indikatorima i strukturi zaposlenih bile su mnogo manje, jedino se više od tri puta povećao broj radnika koji rade van općine u kojoj žive.

Ove promjene bile su popraćene i znatnim porastom varijabiliteta pojedinih indikatora; no kako se vidi iz koeficijenata dilatacije, promjene varijabiliteta bile su vrlo različite i nisu uvijek bile u skladu sa promjenama lokacionih parametara. Zbog toga su se s razlogom mogle očekivati znatne strukturalne promjene.

Algoritam je najprije ispitao sve promjene na nivou glavnih komponenata. U tabeli 2 je struktura glavnih osovina u 1971. i 1981. godini, zajedno sa podacima o relativnoj varijanci svake komponente. Analize su, naravno, izvršene na metrici definiranoj na podacima iz 1971. godine; no struktura glavnih osovina indikatora socioekonomskog razvoja u 1981. godini reparametrizirana je u tabeli 2, radi lakšeg poredjenja, dijeljenjem sa koeficijentima dilatacije.

Tabela 1

ARITMETIČKE SREDINE (μ) I VARIJANCE (σ^2) INDIKATORA DRUŠTVENO-EKONOMSKOG RAZVOJA U 1971. i 1981. GODINI, TE KOEFICIJENTI TRANSLACIJE I DILATACIJE

VARIJABLA	1971		1981		TRANSLACIJA	DILATACIJA
	μ	σ^2	μ	σ^2		
1. DRUPRO	5.34	6.50	48.61	538.00	9.10	9.10
2. AKUMUL	2.64	2.85	19.43	97.61	7.36	5.85
3. ZAPOSL	113.21	4102.30	178.04	5472.89	1.57	1.16
4. ZAPVSS	486.79	2793228.66	799.25	5611222.13	1.64	1.42
5. LJEKAR	3.80	12.38	6.74	24.33	1.77	1.40
6. ELENER	97.40	8150.43	591.17	205050.92	6.07	5.02
7. GUSTPU	6.27	29.60	15.30	95.31	2.44	1.79
8. VRIJED	9.16	129.47	15.47	240.44	1.69	1.36
9. ZAPVOP	823.07	1604406.58	2873.72	80696158.00	3.49	7.09
10. ZAPOSS	5.28	130.97	8.71	360.42	1.65	1.66
11. BRSTAN	84.25	5192.14	94.37	10254.46	1.12	1.41
12. URBANS	10.84	741.80	14.67	1322.01	1.35	1.34
13. DEORGS	22.47	1095.05	30.74	1703.79	1.37	1.25

Tabela 2

STRUKTURA GLAVNIH OSOVINA INDIKATORA DRUŠTVENO-EKONOMSKOG RAZVOJA U 1971. i 1981. GODINI. SVI SU KOEFICIJENTI STANDARDIZIRANI

VARIJABLA	1971. H_1	1971. H_2	1971. H_3	1971. h^2	1981. H_1	1981. H_2	1981. H_3	1981. h^2
1. DRUPRO	.81	.51	.05	.91	.96	-.24	.11	.99
2. AKUMUL	.85	.34	.05	.85	.96	-.18	-.10	.96
3. ZAPOSL	.77	.50	.04	.85	.77	-.11	.26	.67
4. ZAPVSS	.91	-.32	-.16	.95	.62	.74	.04	.94
5. LJEKAR	.84	.19	-.12	.76	.70	.19	.26	.59
6. ELENER	.89	.25	.05	.85	.71	.00	.70	.99
7. GUSTPU	.34	-.38	.83	.96	.07	.17	.01	.03
8. VRIJED	.55	.59	.30	.74	.56	-.13	.16	.35
9. ZAPVOP	.87	-.52	.04	.87	.47	.88	-.08	.99
10. ZAPOSS	.95	-.20	-.17	.98	.66	.68	.07	.90
11. BRSTAN	.79	-.43	.12	.82	.49	.72	.07	.77
12. URBANS	.94	-.26	-.18	.97	.65	.66	.08	.86
13. DEORGS	.93	-.21	-.16	.94	.65	.64	.07	.83
λ_j/trC	.6596	.1487	.0721		.6538	.2357	.0680	
$\sum_{j=1}^{3} \lambda_j/trC$	0.8805				0.9575			

Evidentne su vrlo znatne strukturalne promjene. Prva glavna komponenta, determinirana u 1971. većinom indikatora, no osobito ipak brojem i strukturom zaposlenih i procesom urbanizacije, u 1981. postaje dominantno odredjena akumulacijom i vrijednošću društvenog proizvoda. Druga komponenta u 1981. udvostručuje svoju varijancu i postaje mjera demografskih karakteristika i zaposlenosti. Treća je komponenta u 1971. definirana gustinom suvremenih puteva; u 1981. ova varijabla leži gotovo sasvim izvan zajedničkog prostora, a treća komponenta u 1981. postaje dominantno determinirana potrošnjom električne energije. Iz tabela 3 (kongruencije glavnih osovina) i 4 (korelacije glavnih komponenata) izvrsno se vidi da je od 1971. do 1981. došlo do tako velikih strukturalnih promjena, da je prostor indikatora društveno-ekonomskog razvoja gotovo neusporediv u te dvije vremenske točke.

Tabela 3
KONGRUENCIJA GLAVNIH OSOVINA

			1981.	
		H_1	H_2	H_3
1971.	H_1	.68	.32	.22
	H_2	.33	-.65	.13
	H_3	.07	-.04	.01

Tabela 4
KORELACIJE GLAVNIH KOMPONENATA

			1981.	
		H_1	H_2	H_3
1971.	H_1	.79	.49	.14
	H_2	.30	-.73	.14
	H_3	-.01	-.12	.09

Poseban uvid u strukturu i strukturalne promjene daju podaci dobijeni orthoblique transformacijama bazične solucije. Rezultati, definirani matricama sklopa za 1971. i 1981., dati su u tabeli 5. U tabelama 6 i 7 su korelacije orthoblique faktora u 1971. i 1981. godini, u tabeli 8 kongruencija sklopova, a u tabeli 9 kroskorelacije orthoblique faktora. Rezultati za 1981. ponovno su, radi lakše usporedbe restandardizirani dijeljenjem sa koeficijentima dilatacije.

Tabela 5

SKLOP ORTHOBLIQUE FAKTORA INDIKATORA DRUŠTVENO-EKONOMSKOG RAZVOJA U 1971. I 1981. GODINI. SVI SU KOEFICIJENTI STANDARDIZIRANI.

VARIJABLA	1971.			1981.		
	A_1	A_2	A_3	A_1	A_2	A_3
1. DRUPRO	.02	(.94)	-.04	(1.00)	-.02	-.01
2. AKUMUL	.21	(.77)	.01	(.97)	.04	.00
3. ZAPOSL	.02	(.91)	-.05	(.46)	-.02	(.45)
4. ZAPVSS	(1.03)	-.07	-.05	.08	(.88)	.15
5. LJEKAR	(.48)	(.52)	-.14	.25	.27	(.44)
6. ELENER	.33	(.67)	.03	-.01	-.03	(1.02)
7. GUSTPU	-.01	.04	(.98)	-.03	.18	.02
8. VRIJED	-.40	(1.03)	.19	(.39)	.06	.29
9. ZAPVOP	(.98)	-.27	.20	.01	(1.01)	-.02
10. ZAPOSS	(.96)	.09	-.08	.12	(.81)	.20
11. BRSTAN	(.85)	-.12	.27	-.03	(.82)	.17
12. URBANS	(1.00)	.01	-.09	.12	(.79)	.20
13. DEORGS	(.94)	.08	-.07	.13	(.76)	.20

Tabela 6

KORELACIJE ORTHOBLIQUE FAKTORA U 1971. GODINI

	A_1	A_2	A_3
A_1	1.00	.63	.34
A_2	.63	1.00	.13
A_3	.34	.13	1.00

Tabela 7

KORELACIJE ORTHOBLIQUE FAKTORA U 1981. GODINI

	A_1	A_2	A_3
A_1	1.00	.29	.63
A_2	.29	1.00	.34
A_3	.63	.34	1.00

Tabela 8

KONGRUENCIJA ORTHOBLIQUE FAKTORA U 1971. I 1981. GODINI

			1981.	
		A_1	A_2	A_3
	A_1	.07	.68	.23
1971.	A_2	.64	-.12	.43
	A_3	-.04	.20	.01

Tabela 9

KROSKORELACIJE ORTHOBLIQUE FAKTORA U 1971. I 1981. GODINI

		1981.		
		A_1	A_2	A_3
1971.	A_1	.47	.92	.55
	A_2	.78	.40	.74
	A_3	.03	.36	.14

Kako se vidi, prvi i drugi faktor su zamijenili pozicije, što je za ovaj tip solucije irelevantna promjena jer je redoslijed faktora nebitan. Mnogo je bitnije to što su solucije slabo kongruentne i pored toga što ovaj tip parsimonijske transformacije ima jaku sklonost da sklopove drži invarijantnim ako je to ikako moguće.

Faktor demografskih karakteristika (prvi faktor u 1971., drugi faktor u 1981.) proizvode faktorske vrijednosti koje su u znatnoj korelacije, no i pored toga njihova je kongruentnost slaba, ponajviše zbog toga što broj liječnika na 1000 stanovnika u 1981. godini sudjeluje u formiranju sasvim nove latentne strukture. Faktor ekonomskog razvoja (drugi faktor u 1971., prvi faktor 1981. godine) bitno je promijenio svoju strukturu, ostavši u 1981. godini praktički ograničen na društveni prihod i akumulaciju; uslijed toga je faktor ekonomskog razvoja u vrlo niskoj korelaciji sa faktorom demografskih kretanja. Treći faktor u 1971. je, u stvari, specifični faktor gustine puteva; u 1981. to je sasvim druga dimenzija definirana potrošnjom električne energije, brojem zaposlenih na 1000 stanovnika i brojem liječnika na 1000 stanovnika. Očito, u 1981. došlo je do disocijativnih promjena u društveno-ekonomskom razvoju, tako da je društveno-ekonomska situacija ne samo kvantitativno, već i kvalitativno drugačija od one u 1971. godini.

Čini se, prema tome, da je COMPAT-ZZ algoritam sa prilično pristojnim ponašanjem i da unatoč svoje jednostavnosti, ili upravo zbog toga, može biti od velike koristi za analizu strukturalnih promjena.

LITERATURA

1. Corbalis, M.C. and R.E. Traub. *Longitudinal factor analysis.* Psychometrika, 35(1979), 79-98
2. Fulgosi, A. *Faktorska analiza.* Školska knjiga, Zagreb, 1979
3. Harris, C.W. *Problems in measuring change.* University of Wisconsin Press, Madison, 1963
4. Harris, C.W. and H.F. Kaiser. *Oblique factor analytic solutions by orthogonal transformations.* Psychometrika, 29 (1964),347-362
5. Mulaik, S.A. *The foundations of factor analysis.* McGraw-Hill, New York, 1972
6. Tucker, L.R. *Implications of factor analysis of three-way matrices for measurement of change.* (In: C.W. Harris, *Problems in measuring change*, University of Wisconsin Press, Madison, 1963)
7. Tucker, L.R. *The extension of factor analysis to three-dimensional matrices* (In: N. Frederiksen and H. Gulliksen, *Contributions to mathematical psychology*, Holt, Rinehart and Winston, New York, 1964)
8. Tucker, L.R. *Some mathematical notes on three-mode factor analysis.* Psychometrika, 31(1966),279-311
9. Štalec, J. i K. Momirović. *Ukupna količina valjane varijance kao osnov kriterija za odredjivanje broja značajnih glavnih komponenata.* Kineziologija, 1,1(1971),79-81

2.8 ALGORITAM I PROGRAM LIMAX

Neprecizne a često i kontradiktorne definicije i opisi karakterističnih stanja sportske forme izvor su mnogih zabluda i pogrešaka u procesu treninga.

Sportska forma nije ništa drugo do skup kvantitativnih i kvalitativnih karakteristika psihosomatskog statusa, gdje se pod kvalitativnim karakteristikama podrazumijeva odnos izmedju psihosomatskih dimenzija.

Veoma često se pred trenere postavlja zadatak da optimalnu sportsku formu takmičara zadrže tokom cijelog natjecateljskog perioda, ili da je usmjereno mijenjaju sukladno programu natjecanja.

Algoritmom i programom LIMAX analiziraju se kvalitativne promjene dimenzija psihosomatskog statusa, što omogućuje kontrolu procesa razvoja sportske forme natjecatelja.

NAZIV PROGRAMA

*** L I M A X ***

AUTORI

F. PROT
A. HOSEK
K. BOSNAR
K. MOMIROVIC

FUNKCIJA

L I M A X JE PROGRAM ZA PSEUDOKANONICKU ANALIZU STRUKTURALNIH PROMJENA KOJI SE TEMLJI NA ALGORITMU ZA PSEUDOKANONICKU FAKTORSKU ANALIZU (BOSNAR, PROT, MOMIROVIC, LUZAR, DOBRIC; 1981) TE ALGORITMU ZA KANONICKU ANALIZU PROMJENA PSIHOSOMATSKOG STATUSA IZAZVANIH NEKIM KINEZIOLOSKIM TRETMANOM (MOMIROVIC; 1969,1975).

MOGUCE GA JE PRIMJENITI:

(1) KADA JE PROCES KINEZIOLOSKIH TRANSFORMACIJA TAKO PROGRAMIRAN, DA SE, OSIM KVANTITATIVNIH, MOGU OCEKIVATI I STRUKTURALNE PROMJENE, DAKLE PROMJENE MEDJUSOBNIH ODNOSA PSIHOSOMATSKIH DIMENZIJA, ILI SE SAMO TAKOVE PROMJENE MOGU OCEKIVATI
(2) KADA JE IZ BILO KOGA RAZLOGA, PODTREBNO UTVRDITI STRUKTURALNE PROMJENE U DVA SUKCESIVNA TRANZITIVNA STANJA

L I M A X IZVODI OVE OPERACIJE:

(1) FORMIRA SUPER MATRICU OPISA ENTITETA U TRANZITIVNIM STANJIMA (T) I (T+1)
(2) ANALIZIRA LATENTNE DIMENZIJE POD KOMPONENTNIM MODELOM I FIKSIRA BROJ DIMENZIJA U SKLADU SA GK KRITERIJEM (GUTTMAN; KAISER, 1960)
(3) ZA TAKO FIKSIRANI BROJ FAKTORA ODREDJUJE KOMUNALITETE ITERATIVNIM POSTUPKOM
(4) DEFINIRA UNIKVITETE NA TEMELJU TAKO PROCJENJENIH KOMUNALITETA I FORMIRA REDUCIRANU MATRICU KOVARIJANCI VARIJABLI RESKALIRANIH NA INVERZNU METRIKU UNIKNIH KOMPONENATA
(5) ODREDJUJE GLAVNE OSOVINE RESKALIRANE REDUCIRANE MATRICE KOVARIJANCI, FIKSIRA BROJ FAKTORA U SKLADU SA DMEAN KRITERIJEM (MOMIROVIC I STALEC, 1973) I RESKALIRA FAKTORE NA METRIKU STANDARDIZIRANIH VARIJABLI
(6) MATRICU PSEUDOKANONICKIH FAKTORA PARTICIONIRA NA MATRICE KOORDINATA VEKTORA VARIJABLI TRANZITIVNIH STANJA (T) I (T+1)
(7) RAZLIKOM KOORDINATA VARIJABLI TRANZITIVNIH STANJA (T) - (T+1) ODREDJUJE STRUKTURALNE PROMJENE U SISTEMU PSEUDOKANONICKIH FAKTORA
(8) ODREDJUJE KOSINUSE KUTEVA PARTICIJA PSEUDOKANONICKIH FAKTORA TRANZITIVNIH STANJA (T) I (T+1)
(9) ROTIRA PSEUDOKANONICKE FAKTORE U NORMAL VARIMAX POZICIJU (KAISER, 1958)
(10) MATRICU VARIMAX PSEUDOKANONICKIH FAKTORA PARTICIONIRA NA MATRICE KOORDINATA VEKTORA VARIJABLI TRANZITIVNIH STANJA (T) I (T+1)
(11) RAZLIKOM KOORDINATA VARIJABLI TRANZITIVNIH STANJA (T) -(T+1) ODREDJUJE STRUKTURALNE PROMJENE U SISTEMU VARIMAX PSEUDOKANONICKIH FAKTORA
(12) ODREDJUJE KOSINUSE KUTEVA PARTICIJA VARIMAX PSEUDOKANONICKIH

```
*         FAKTORA TRANZITIVNIH STANJA (T) I (T+1)
*          (13) TRANSFORMIRA PSEUDOKANONICKE FAKTORE U PROMAX  POZICIJU,
*         (HENDRICKSON I WHITE, 1964)
*          (14) MATRICU SKLOPA PROMAX  PSEUDOKANONICKIH FAKTORA
*         PARTICIONIRA NA MATRICE KOORDINATA VEKTORA VARIJABLI TRANZITIVNIH
*         STANJA (T) I (T+1)
*          (15) RAZLIKOM KOORDINATA VARIJABLI TRANZITIVNIH STANJA (T) - (T+1)
*         ODREDJUJE STRUKTURALNE PROMJENE U SISTEMU PROMAX  PSEUDO-
*         KANONICKIH FAKTORA
*          (16) ODREDJUJE KOSINUSE KUTEVA PARTICIJA PROMAX
*         PSEUDOKANONICKIH FAKTORA TRANZITIVNIH STANJA (T) I (T+1)
*
*
*
*    OSNOVNA METODA I ALGORITAM OPISANI SU U
*
*             PROF. F.. A. HOSEK, K. BOSNAR I K. MOMIROVIC:
*             ALGORITAM I PROGRAM ZA ANALIZU STRUKTURALNIH PROMJENA,
*             KINEZIOLOGIJA,16,1(1984), U TISKU
*
*    PRIPREMA PODATAKA
*
*             KONTROLNI ZAPIS I PODACI ZA PRVI SKUP U FILEU
*             DATA1, KONTROLNI ZAPIS I PODACI ZA DRUGI SKUP
*             U FILEU DATA2.
*
*    UPOZORENJE KORISNIKU:
*
*          (1) PROGRAM PODRAZUMJEVA DA SE PRVI SKUP PODATAKA
*        ODNOSI NA TRANZITIVNO STANJE (T), A DA SE DRUGI SKUP
*        PODATAKA ODNOSI NA TRANZITIVNO STANJE (T+1)
*
*          (2) VARIJABLE KOJIMA SU OPISANI ENTITETI U OBA
*        TRANZITIVNA STANJA (T) I (T+1) TREBAJU BITI IDENTICNE
*
*          (3) PREPORUCIJE SE DA KORISNIK U IMENIMA VARIJABLI
*        PRVU KOLONU REZERVIRA ZA OZNAKU TRANZITIVNOG STANJA (1.....2.....)
*
*          (4) NEOPHODNO JE DA IMENA VARIJABLI U PREOSTALIH PET KOLONA
*        BUDU IDENTICNA ZA TRANZITIVNA STANJA  (T)  I (T+1)
*
*
*    IZVRSNI DIO PROGRAMA
*
*
*
*              1.  ULAZ , SREDJIVANJE  PODATAKA , PARAMETRI I
*                  DISTRIBUCJE VARIJABLE TRANZITIVNIH STANJA (T) I (T+1)
*
*
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=L I M A X - PROGRAM ZA PSEUDOKANONICKU ANALIZU KVALITATIVNIH,T)
HEADING(TEXT= PROMJENA PSIHOSOMATSKOG STATUSA)
HEADING(TEXT= U DVIJE VREMENSKE TOCKE  T  I  T+1,D)
TEXT(TEXT=          PARAMETRI VARIJABLI)
INPUT(SCORE=ISSC1,DATA=DATA1)
INPUT(SCORE=ISSC2,DATA=DATA2)
CONFORM(IN1=ISSC1,IN2=ISSC2,OUT1=SSC1,OUT2=SSC2)
          DELETE(MATRIX=ISSC1)
          DELETE(MATRIX=ISSC2)
TRANSPOSE(OLD=SSC1,NEW=TSSC1)
TRANSPOSE(OLD=SSC2,NEW=TSSC2)
          DELETE(MATRIX=SSC1)
```

```
            DELETE(MATRIX=SSC2)
SORT(IN=TSSC1,OUT=TSC1)
SORT(IN=TSSC2,OUT=TSC2)
TRANSPOSE(OLD=TSC1,NEW=SC1)
TRANSPOSE(OLD=TSC2,NEW=SC2)
CROSSCORRELATION(P1=SC1,P2=SC2,R12=NAMT1)
TRANSPOSE(OLD=NAMT1,NEW=NAMT2)
            DELETE(MATRIX=TSC1)
            DELETE(MATRIX=TSC2)
HEADING(TEXT=PARAMETRI TRANZITIVNOG STANJA  = T,D)
STATISTICS(SCORE=SC1,S,CLASS=9,Z=ZSC1)
HEADING(TEXT=PARAMETRI TRANZITIVNOG STANJA  = T+1,D)
STATISTICS(SCORE=SC2,S,CLASS=9,Z=ZSC2)
            DELETE(MATRIX=SC1)
            DELETE(MATRIX=SC2)
MERGE(IN1=ZSC1,IN2=ZSC2,OUT=Z)
            DELETE(MATRIX=ZSC1)
            DELETE(MATRIX=ZSC2)
HEADING(TEXT= RELACIJE VARIJABLI TRANZITIVNIH STANJA  T  I T+1,D)
CORRELATION.
PRINT(MATRIX=R,TEXT= SUPERMATRICA KORELACIJE VRIJABLI TRANZITIVNIH STANJA)
            DELETE(MATRIX=Z)
 VERSION(P)
            DELETE(MATRIX=RINV)
*
*
*         2. INICIJALNA SOLUCIJA
*
*
TEXT(TEXT=              INICIJALNA SOLUCIJA)
HEADING(TEXT= INICIJALNA SOLUCIJA,D)
DIAGONALISATION.
HOTELLING(F=HT)
            DELETE(MATRIX=LAMBDA)
ITERCOMMUNALITIES(F=HT,FN=FCT,LAMBDA=LAMBDA,ITER=60,NAMEF=NMF)
            DELETE(MATRIX=LAMBDA)
            DELETE(MATRIX=X)
            DELETE(MATRIX=HT)
TRANSPOSE(OLD=FCT,NEW=FC)
MULT(A=FC,B=FCT,M=RTEOR)
DIAG(A=RTEOR,D=COM)
 AG(A=R,D=I)
LINEAR(A=I,B=COM,CB=-1.0,M=U2)
DIAG(A=U2,C=-0.5,D=UMIN)
            DELETE(MATRIX=FCT)
            DELETE(MATRIX=FC)
            DELETE(MATRIX=RTEOR)
*
*
*         3. ORTOGONALNI PSEUDOKANONICKI FAKTORI
*
*
TEXT(TEXT=PSEUDO    KANONICKI FAKTORI)
HEADING(TEXT= ORTOGONALNI PSEUDOKANONICKI FAKTORI,D)
MULT(A=UMIN,B=R,M=UR)
MULT(A=UR,B=UMIN,M=V)
LINEAR(A=V,B=I,CB=-1.0,M=W)
            DELETE(MATRIX=V)
DIAGONALISATION(R=W)
HOTELLING(F=HT)
DIAGMULT(A=HT,D=UMIN,C=-1.0,P,M=FT)
TRANSPOSE(OLD=FT,NEW=F)
RESIDUAL(F=FT)
PRINT(MATRIX=F,TEXT= PSEUDOKANONICKI FAKTORI)
```

```
PRINT(MATRIX=RES,TEXT= REZIDUALI SUPERMATRICE  R,D)
              DELETE(MATRIX=UMIN)
              DELETE(MATRIX=U2)
              DELETE(MATRIX=W)
              DELETE(MATRIX=I)
              DELETE(MATRIX=X)
              DELETE(MATRIX=LAMBDA)
              DELETE(MATRIX=HT)
HEADING(TEXT= STRUKTURALNE PROMJENE  PSEUDAOKANONICKI PROSTOR,D)
CONFORM(IN1=NAMT1,IN2=F,OUT1=ENA1,OUT2=F1)
CONFORM(IN1=NAMT2,IN2=F,OUT1=ENA2,OUT2=F2)
              DELETE(MATRIX=ENA1)
              DELETE(MATRIX=ENA2)
LINEAR(A=F1,B=F2,CB=-1.0,M=DF)
PRINT(MATRIX=DF,TEXT= RAZLIKE KOORDINATA TRANZITIVNIH STANJA PSEUDOKANF)
MULT(A=F1,TA,B=F1,M=1FF1)
MULT(A=F2,TA,B=F2,M=2FF2)
MULT(A=F1,TA,B=F2,M=1FF2)
DIAGMULT(A=1FF2,D=1FF1,C=-0.5,L,M=CCONF)
DIAGMULT(A=CCONF,D=2FF2,C=-0.5,R,M=CONF)
PRINT(MATRIX=CONF,TEXT= KOSINUSI KUTEVA PARTICIJA PSEUDOKANONICKIH FAC,D)
*
*
*          4. VARIMAX TRANSFORMACIJA PSEUDOKANONICKIH FAKTORA
*
*
TEXT(TEXT=VARIMAX   PSEUDOKAN FAKTORA)
HEADING(TEXT= VARIMAX TRANSFORMACIJA PSEUDOKANONICKIH FAKTORA,D)
VARIMAX(F=FT,TAU=TT,FN=VT,N)
TRANSPOSE(OLD=VT,NEW=V)
PRINT(MATRIX=TT,TEXT=VARIMAX TRANSFORMACIJSKA MATRICA,D)
PRINT(MATRIX=V,TEXT= STRUKTURA VARIMAX FAKTORA)
HEADING(TEXT= STRUKTURALNE PROMJENE VARIMAXPSEUDOKAN PROSTOR,D)
CONFORM(IN1=NAMT1,IN2=V,OUT1=ENA1,OUT2=V1)
CONFORM(IN1=NAMT2,IN2=V,OUT1=ENA2,OUT2=V2)
              DELETE(MATRIX=ENA1)
              DELETE(MATRIX=ENA2)
LINEAR(A=V1,B=V2,CB=-1.0,M=DV)
PRINT(MATRIX=DV,TEXT=RAZLIKE KOORDINATA TRANZITIVNIH STANJA VARIMAXF,D)
MULT(A=V1,TA,B=V1,M=1VV1)
MULT(A=V2,TA,B=V2,M=2VV2)
MULT(A=V1,TA,B=V2,M=1VV2)
DIAGMULT(A=1VV2,D=1VV1,C=-0.5,L,M=CCONV)
DIAGMULT(A=CCONV,D=2VV2,C=-0.5,R,M=CONV)
PRINT(MATRIX=CONV,TEXT=KOSINUSI KUTEVA PARTICIJA VARIMAX FAKTORA,D)
*
*
*          5. PROMAX  TRANSFORMACIJA PSEUDOKANONICKIH FAKTORA
*
*
TEXT(TEXT=PROMAX    PSEUDOKAN FAKTORI)
HEADING(TEXT= PROMAX  TRANSFORMACIJA PSEUDOKANONICKIH FAKTORA,D)
HADMULT(A=VT,C=5.0,M=VVT)
MULT(A=FT,B=VVT,TB,M=FVTF)
MULT(A=FT,B=FT,TB,M=FTF)
GAUSSJORDAN(M=FTF,MINV=IFTF)
MULT(A=IFTF,B=FVTF,M=TZ)
MULT(A=TZ,B=TZ,TA,M=TZTZ)
GAUSSJORDAN(M=TZTZ,MINV=ITZTZ)
DIAGMULT(A=TZ,T,D=ITZTZ,C=0.5,L,M=QT)
MULT(A=QT,B=FT,M=AT)
TRANSPOSE(OLD=AT,NEW=A)
MULT(A=QT,B=QT,TB,M=TTT)
GAUSSJORDAN(M=TTT,MINV=M)
```

```
MULT(A=M,B=AT,M=FPST)
PRINT(MATRIX=QT,T,TEXT=  PROMAX  TRANSFORMACIJSKA MATRICA,D)
PRINT(MATRIX=A,TEXT=SKLOP PROMAX  FAKTORA)
PRINT(MATRIX=FPST,T,TEXT=STRUKTURA PROMAX  FAKTORA,D)
PRINT(MATRIX=M,TEXT=KORELACIJE PROMAX  FAKTORA,D)
          DELETE(MATRIX=AT)
HEADING(TEXT=STRUKTURALNE PROMJENE PROMAX PSEUDOKAN PROSTOR,D)
CONFORM(IN1=NAMT1,IN2=A,OUT1=ENA1,OUT2=A1)
CONFORM(IN1=NAMT2,IN2=A,OUT1=ENA2,OUT2=A2)
          DELETE(MATRIX=ENA1)
          DELETE(MATRIX=ENA2)
LINEAR(A=A1,B=A2,CB=-1.0,M=DA)
PRINT(MATRIX=DA,TEXT= RAZLIKE KOORDINATA TRANZITIVNIH STANJA PROMAX F,D)
MULT(A=A1,TA,B=A1,M=1AA1)
MULT(A=A2,TA,B=A2,M=2AA2)
MULT(A=A1,TA,B=A2,M=1AA2)
DIAGMULT(A=1AA2,D=1AA1,C=-0.5,L,M=CCONA)
DIAGMULT(A=CCONA,D=2AA2,C=-0.5,R,M=CONO)
PRINT(MATRIX=CONO,TEXT= KOSINUSI KUTEVA PARTICIJA PROMAX  FAKTORA,D)
*
*
*           KRAJ  PROGRAMA      L I M A X
```

ALGORITAM I PROGRAM ZA ANALIZU STRUKTURALNIH PROMJENA *

Franjo Prot, Ankica Hošek, Ksenija Bosnar i Konstantin Momirović
INSTITUT ZA KINEZIOLOGIJU FAKULTETA ZA FIZIČKU KULTURU U ZAGREBU
Odjel za informatiku i statistiku

Predložen je algoritam i napisan program za analizu strukturalnih promjena u dvije konsekutivne vremenske točke. Rad se bazira na postupku za analizu strukturalnih promjena (Harris, 1963; Momirović, 1972) te algoritmu za pseudokanoničku faktorsku analizu (Bosnar, Prot, Momirović, Lužar i Dobrić, 1982). Iz super matrice korelacija kvantitativnih varijabli obadvije vremenske točke algoritam odredjuje ortogonalnu soluciju definiranu pseudokanoničkim faktorima koju zatim transformira u varimax i oblimin poziciju. Za svaku od dobivene solucije algoritam definira promjene struktura kao razlike koordinata vektora varijabli u prvoj i drugoj točki registracije na dobivenim faktorima, i kao koeficijente kongruencije tih koordinata. Ponašanje algoritma je prikazano na podacima jednog sociološkog istraživanja.

1. UVOD

Procesi koje istražuje sociologija ali i kineziologija, psihologija i druge antropološke znanosti u odredjenom su broju slučajeva takve naravi da se osim kvantitativnih promjena mogu očekivati i kvalitativne, strukturalne promjene, dakle promjene medjusobnih odnosa obilježja. Ponekad su neki transformacijski postupci usmjereni prema strukturalnim promjenama izmedju dva sukcesivna stanja, ili održanju konstantne strukture bez obzira na eventualne kvantitativne promjene.

Mogućnost analize strukturalnih promjena stoga može biti od presudnog značenja kako za daljnju analizu prikupljenih podataka, tako i za donošenje relevantnih odluka.

---

* *Ovaj rad će biti objavljen u časopisu Kineziologija, 16 (1984), 2, u tisku.*

Na osnovu algoritma za pseudokanoničku faktorsku analizu (Bosnar, Prot, Momirović, Lužar i Dobrić, 1981) koji se temelji na relacijama faktorskog modela (Rao, 1955), image modela (Gutman, 1953) i modela sa univerzalnom metrikom (Harris, 1962) te algoritma za kanoničku analizu promjena (Harris, 1963; Momirović, 1972), moguće je konstruirati jednostavan algoritam za analizu strukturalnih promjena u toku nekog stohastičkog procesa. U ovom je radu definiran jedan takav algoritam i opisan program kojim je taj algoritam implementiran.

2. ALGORITAM

2.1. *Preliminarne operacije*

Neka su $V_t = \{v_j;\ j = 1,\ldots,m\}$ i $V_{t+1} = \{v_l;\ l = 1,\ldots,r\}$; $m = r$, dva skupa istih multivarijatno normalno distribuiranih manifestnih ili latentnih varijabli registriranih na skupu entiteta $E = \{e_i;\ i = 1,\ldots,n\}$ u vremenskim točkama t i $t+1$ u nekom vremenskom intervalu $\{t_0,\ t_f\}$. Podrazumijeva se da su skupovi varijabli V_t i V_{t+1} odabrani tako da zadovoljavajuće dobro opisuju tranzitivna stanja u vremenskim točkama t i $t+1$.

Neka je $B_1 = (b_{1ij});\ i\ 0\ 1,\ldots,n;\ j = 1,\ldots,m$ nesingularna matrica podataka dobivena opisom entiteta iz E nad skupom varijabli V_t i neka je $B_2 = (b_{2il});\ i = 1,\ldots,n;\ l = 1,\ldots,r;$ $m = r$, nesingularna matrica podataka dobivena opisom istih entiteta iz E nad skupom varijabli V_{t+1}.

Organizirajmo matrice B_1 i B_2 u super matricu $B = (b_{is});\ i = 1,\ldots,n;\ s = 1,\ldots,m+r$, koja ima oblik

$$B = |B_1 \vdots B_2| \ .$$

Pretpostavimo da ne postoji značajan gubitak informacija ako super matricu B podvrgnemo postupku standardizacije. Takovu novu super matricu standardiziranih varijabli označimo sa Z; naravno, ta matrica ima oblik

$$Z = (z_{is}) = [Z_1 \vdots Z_2] \ .$$

Operacijom

$$Z^T Z \frac{1}{n} = R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

formiramo super matricu R gdje je R_{11} matrica interkorelacija varijabli u tranzitivnom stanju u t, R_{22} matrica interkorelacija varijabli u tranzitivnom stanju $t+1$ a matrica $R_{12} = R_{21}^T$ matrica kroskorelacija varijabli registriranih u točkama t i $t+1$.

Guttmanove procjene unikviteta varijabli iz skupa $\{V_t, V_{t+1}\}$ odredjene su operacijom

$$U^2 = (\text{diag}R^{-1})^{-1}.$$

2.2. *Inicijalna ortogonalna solucija*

Dekomponirajmo super matricu na na ove aditivne matrice

$$R = X\Lambda X^T + X^*\Lambda^* X^{*T}$$

gdje je $\Lambda = (\lambda_g)$; $g = 1,\ldots,k$, matrica prvih k svojstvenih vrijednosti koje zadovoljavaju kriterij $\lambda_g > 1$, a $X = (x_g)$; $X^T X = I$, matrica njima pridruženih svojstvenih vektora.

U matrici

$$H = X\Lambda^{1/2}$$

bit će prvih k glavnih osovina super matrice R.

Inicijalnu procjenu komunaliteta varijabli iz V_t i V_{t+1} u zajedničkom prostoru definiramo operacijom

$$h_0^2 = \mathrm{diag}(HH^T).$$

Konačna procjena komunaliteta odredjena je iterativnim procesom:

$$R_\alpha = R - I + h_{\alpha-1}^2$$

$$(R_\alpha - \lambda_{p\alpha} I) x_{p\alpha} = 0 \qquad p = 1,\ldots,\nu$$

$$H_\alpha = X_\alpha \Lambda_\alpha^{1/2}$$

$$h_\alpha^2 = \mathrm{diag}(H_\alpha H_\alpha^T)$$

gdje je α oznaka iteracije.

Iterativni proces se zaustavlja kada se zadovolji uvjet

$$|h_\alpha^2| - |h_{\alpha+1}^2| < \varepsilon$$

gdje je ε proizvoljno mali realni broj (u programu LIMA $\varepsilon = 0.005$), ili jednu iteraciju prije nego što nastupi generalizirani Heywoodov slučaj.

U matrici $H_\alpha = X_\alpha \Lambda_\alpha^{1/2}$, gdje su Λ_α i X_α matrice svojstvenih vrijednosti i svojstvenih vektora matrice $R - I + h_\alpha^2$ dobivene u posljednjoj iteraciji, su glavne osovine reducirane matrice korelacija procijenjene pod standardnim faktorskim modelom.

Unikviteti varijabli na kraju iterativnog procesa odredjeni su operacijom

$$S^2 = I - h_f^2$$

gdje je h_f^2 matrica komunaliteta varijabli dobivena na kraju iterativnog procesa.

Kovarijance varijabli u skupu $\{V_t, V_{t+1}\}$ reskaliranih na Harrisovu metriku procijenjene su operacijom

$$C = S^{-1}(R - S^2)S^{-1} = S^{-1}RS^{-1} - I.$$

Definira li se matrica $Y = (y_p)$; $p = 1,\ldots,q$ kao matrica prvih q svojstvenih vektora matrice C, u matrici Y su ujedno i prvih q svojstvenih vektora matrice $S^{-1}RS^{-1}$.

Ako su u matrici

$$N = (n_p) \qquad p = 1,\ldots,q$$

prvih q svojstvenih vrijednosti matrice $S^{-1}RS^{-1}$, onda su u matrici

$$\eta = (n_p - 1); \qquad p = 1,\ldots,q$$

prvih q svojstvenih vrijednosti matrice C.

Ortogonalna faktorska matrica, odnosno matrica pseudokanoničkih faktora

$$L = SY(\eta - I)^{1/2}$$

procjena je faktora matrice $R - S^2$ pod modelom analognim modelu najveće vjerodostojnosti (Harris, 1962; Momirović, 1964; 1982) jer su u $L^* = Y(\eta - I)$ prvih q** glavnih osovina matrice C.

Super matricu pseudokanoničkih faktora možemo posmatrati kao

$$L = \left[\begin{array}{c} L_1 \\ \hdashline L_2 \end{array}\right]$$

gdje su particije L_1 i L_2 koordinate vektora varijabli iz V_t i V_{t+1} koje opisuju tranzitivna stanja u točkama t i $t+1$.

** *Broj q je odredjen sukladno DMEAN kriteriju (Momirović i Štalec, 1973)*

Reprodukcija super matrice $R - S^2$ odredjena je operacijom

$$LL^T = R - S^2$$

što je

$$LL^T = \begin{bmatrix} L_1L_1^T & L_1L_2^T \\ L_2L_1^T & L_2L_2^T \end{bmatrix}$$

odnosno

$$LL^T = \begin{bmatrix} R_1 - S_1^2 & R_{12} \\ R_{21} & R_2 - S_2^2 \end{bmatrix}$$

Usporedbom koeficijenata particija super matrice pseudokanoničkih faktora L_1 i L_2 može se steći uvid u promjene strukturalnih odnosa varijabli iz V_t i V_{t+1} u tranzitivnim stanjima t i $t+1$.

Promjene u svakoj od odgovarajućih varijabli za svaki od q pseudokanoničkih faktora odredjene su razlikom

$$\Delta_L = L_1 - L_2 .$$

Odnos koordinata varijabli u tranzitivnim stanjima u t i $t+1$ u prostoru pseudokanoničkih faktora odredjen je i kosinusima kuteva vektora particija L_1 i L_2

$$\Psi_L = \mathrm{diag}(L_1^TL_1)^{-1/2}L_1^TL_2\ \mathrm{diag}\ (L_2^TL_2)^{-1/2} .$$

2.3. Varimax transformacija pseudokanoničkih faktora

Neka je $T = (t_{pc})$; $p,c = 1,\dots,q$; $T^TT = TT^T = I$ i neka se operacijom

$$\mathsf{LT} = \mathsf{V} = (v_{sp})\ ;\ s = 1,\ldots,m+r$$

maksimizira Kaiserova varimax funkcija (Kaiser, 1958)

$$w = (m+r)\sum_{s=1}^{m+r}\sum_{p=1}^{q}(v_{jp}/h_s)^4 - \sum_{p=1}^{q}\left(\sum_{s=1}^{m+r}(v_{jp}/h_s)^2\right)^2$$

gdje su h_s^2 komunaliteti varijabli procijenjeni operacijom $\mathsf{H}^2 = \mathrm{diag}(\mathsf{L}\mathsf{L}^T)$. Tada super matrica V odredjuje strukturu varijabli iz V_t i V_{t+1} sukladno parsimonijskoj soluciji definiranoj normal varimax funkcijom.

Particije V_1 i V_2 u

$$\mathsf{V} = \begin{bmatrix} \mathsf{V}_1 \\ ---- \\ \mathsf{V}_2 \end{bmatrix}$$

koordinate su vektora varijabli tranzitivnih stanja opisanih na skupinama varijabli V_t i V_{t+1} u prostoru pseudokanoničkih faktora transformiranih u varimax poziciju, pa se njihovom usporedbom može steći uvid u eventualne kvalitativne promjene. Promjene za svaku od varijabli V_t u odnosu na odgovarajuću varijablu u V_{t+1} na svakom od q pseudokanoničkih faktora transformiranih u varimax poziciju eksplicitno su odredjene razlikom particija

$$\Delta_V = \mathsf{V}_1 - \mathsf{V}_2\ .$$

Odnos koordinata varijabli stanja u V_t i V_{t+1} u prostoru varimax faktora odredjen je kosinusima kuteva particija V_1 i V_2

$$\Psi_V = \mathrm{diag}(\mathsf{V}_1^T\mathsf{V}_1)^{-1/2}\mathsf{V}_1^T\mathsf{V}_2\mathrm{diag}(\mathsf{V}_2^T\mathsf{V}_2)^{-1/2}\ .$$

2.4 *Promax transformacija pseudokanoničkih faktora*

Neortogonalno parsimonijsko rješenje je, očito, bliže stvarnim procesima promjene odnosa latentnih dimenzija. Stoga su pseudokanonički faktori podvrgnuti i promax transformaciji (Hendrickson i White, 1964).

Ako se definira matrica

$$W = (|v_{jp}|^{\delta+1}/\, v_{jp})$$

gdje je δ neka proizvoljna konstanta (u programu LIMAX $\delta = 4$). Operacijom

$$LQ = A$$

gdje je

$$Q = Q^* \, (\mathrm{diag}(Q^{*T}Q^*)^{-1})^{1/2}$$

a

$$Q^* = (L^TL)^{-1}L^TW\,,$$

dobit će se aproksimacija neortogonalne parsimonijske solucije definirane super matricom A; ova je solucija, naravno, promax solucija Hendricksona i Whitea (1964), odnosno, parsimonijski koordinirani sustav za vektore iz $\{V_t, V_{t+1}\}$. Procjena vjerojatnih korelacija promax faktora definirana je operacijom

$$M = (Q^TQ)^{-1},$$

a korelacije varijabli iz $\{V_t, V'_{t+1}\}$ i promax faktora tada su

$$F = AM.$$

Particije A_1 i A_2 matrice

$$A = \begin{bmatrix} A_1 \\ \hdashline A_2 \end{bmatrix}$$

koordinate su vektora varijabli tranzitivnog stanja opisanih nad varijablama V_t i V_{t+1} u prostoru promax faktora.

3. PROGRAM LIMAX

Na osnovu algoritma za analizu strukturalnih promjena (Harris, 1963; Momirović, 1972) i algoritma za pseudokanoničku faktorsku analizu (Bosnar, Prot, Momirović, Lužar i Dobrić, 1982) i ovdje opisanog algoritma za pseudokanoničku analizu strukturalnih promjena,napisan je program LIMAX.

Ovaj je program moguće primijeniti:

(1) Kada je neki stohastički proces takav da se, osim kvantitativnih, mogu očekivati i strukturalne promjene, dakle promjene medjusobnih odnosa obilježja, ili se samo takove promjene mogu očekivati;

(2) kada je iz bilo koga razloga potrebno utvrditi strukturalne promjene u dva sukcesivna tranzitivna stanja.

Program LIMAX izveden je u pet segmenata u okviru kojih izvodi ove operacije:

(1) formira super matricu opisa entiteta u tranzitivnim stanjima t i $t+1$;

(2) analizira latentne dimenzije pod komponentnim modelom i fiksira broj dimenzija u skladu sa Guttman-Kaiserovim kriterijem;

(3) za tako formiran broj faktora odredjuje komunalitete iterativnim postupkom;

(4) definira unikvitete na temelju tako procijenjenih komunaliteta i formira reduciranu matricu kovarijanci varijabli reskaliranih na inverznu metriku uniknih komponenata;

(5) odredjuje glavne osovine reskalirane reducirane matrice kovarijanci, fiksira broj faktora u skladu sa DMEAN kriterijem (Momirović i Štalec, 1973) i reskalira faktore na metriku standardiziranih varijabli;

(6) matricu pseudokanoničkih faktora particionira na matrice koordinata vektora varijabli tranzitivnih stanja *t* i *t+1*;

(7) razlikom koordinata varijabli u tranzitivnim stanjima *t* i *t+1* odredjuje strukturalne promjene u sistemu pseudokanoničkih faktora;

(8) odredjuje kosinuse kuteva particija pseudokanoničkih faktora u tranzitivnim stanjima *t* i *t+1*;

(9) rotira pseudokanoničke faktore u normal varimax poziciju (Kaiser, 1958);

(10) matricu pseudokanoničkih faktora u varimax poziciji particionira na matrice koordinata varijabli tranzitivnih stanja *t* i *t+1*;

(11) razlikom koordinata varijabli tranzitivnih stanja *t* i *t+1* odredjuje strukturalne promjene u sistemu pseudokanoničkih faktora u varimax poziciji;

(12) odredjuje kosinuse kuteva particija pseudokanoničkih faktora u varimax poziciji tranzitivnih stanja *t* i *t+1*;

(13) rotira pseudokanoničke faktore u promax poziciju (Hendrickson i White, 1964);

(14) matricu sklopa pseudokanoničkih faktora u promax poziciji particionira na matrice koordinata vektora varijabli tranzitivnih stanja *t* i *t+1*;

(15) razlikom koordinata varijabli u tranzitivnim stanjima *t* i *t+1* odredjuje strukturalne promjene u sistemu pseudokanoničkih faktora u promax poziciji;

(16) odredjuje kosinuse kuteva particija pseudokanoničkih faktora u promax poziciji u tranzitivnim stanjima *t* i *t+1*.

Program LIMAX napisan je u meta jeziku SS, verzija 5.2/M (Zakrajšek, Štalec i Momirović, 1974; Momirović, Štalec i Zakrajšek, 1982). Korisnicima je dostupan u datoteci SRCE*SS-MAKRO Sveučilišnog računskog centra u Zagrebu.

4. PONAŠANJE ALGORITMA

Za ilustraciju efikasnosti algoritma predočeni su rezultati oblimin transformacije pseudokanoničkih faktora* dobivenih u sklopu jednog kineziološkog istraživanja.

---

* *Ostali rezultati koji nisu mogli biti objavljeni u ovom radu pohranjeni su u arhivi Instituta za kineziologiju Fakulteta za fizičku kulturu Sveučilišta u Zagrebu*

U sklopu obuke rvanja za studente I godine fakulteta za fizičku kulturu, školske godine 1981/82 na uzorku od 115 studenata postavljen je cilj kvantitativne promjene jednog skupa varijabli energetske regulacije uz uvjet da kvalitativni (strukturalni) odnos varijabli ostane nepromijenjen.

Uzorak primijenjenih varijabli:

PRVO MJERENJE

1 CUC - čučnjevi sa partnerom na ledjima
1 DSM - skok udalj s mjesta
1 DTK - dizanje trupa na klupici
1 SKL - sklekovi na tlu
1 ZGIB - zgibovi na preči
1 NABAC - prebacivanje partnera hvatom oko pojasa.

DRUGO MJERENJE

2 CUC - čučnjevi sa partnerom na ledjima
2 DSM - skok udalj s mjesta
2 DTK - dizanje trupa na klupici
2 SKL - sklekovi na tlu
2 ZGIB - zgibovi na preči
2 NABAC - prebacivanje partnera hvatom oko pojasa.

U prostoru motoričkih varijabli registriranih prije i nakon obuke rvanja algoritam ekstrahira i transformira u promax poziciju četiri pseudokanoničkih faktora (rezultati su predočeni u tabelama 1, 2, 3, 4 i 5).

U matrici sklopa, tabela 1, prvi faktor (PRX 1) jasno je definiran podjednako visokom projekcijom varijabli dizanje trupa na klupici prvog i drugog mjerenja (1 DTK, 2 DTK). Projekcije ostalih varijabli daleko su niže, ali podjednake visine i u prvom i u drugom mjerenju. Drugi faktor (PRX 2) definiran je podjednako visokom projekcijom varijabli skok udalj s mjesta prvog i drugog mjerenja (1 DSM, 2 DSM). Preostale projekcije varijabli na ovaj faktor su niske, ali odgovaraju visinom projekcija korespondentnim varijablama prvog i drugog mjerenja. Treći faktor (PRX 3) podjednakim, visokim projekcijama u prvom i drugom mjerenju opisuju varijable zgibovi na preči i sklekovi na tlu (1 ZGIB, 2 ZGIB, 1 SKL, 2 SKL) a niže projekcije ostalih varijabli ponašaju se slično kao i kod prethodnih faktora. Ista forma se ponavlja i na četvrtom faktoru (PRX 4) s tom razlikom što ga visokim i podjednakim projekcijama za prvo i drugo mjerenje opisuje više varijabli: sklekovi na tlu, čučnjevi sa partnerom na ledjima i prebacivanje partnera hvatom oko pojasa (1 SKL, 2 SKL, 1 CUC, 2 CUC, 1 NABAC, 2 NABAC).

Da su projekcije pojedinih varijabli podjednake na ekstrahiranim faktorima za prvo i drugo mjerenje jasno je prikazano većinom nultim ili minimalnim razlikama koordinata tranzitivnih stanja u t i t+1 u prostoru oblimin faktora (tabela 2).

Isto tako podaci o visokim kosinusima kuteva particija korespondentnih oblimin faktora (tabela 3) potvrdjuju da je u ovom slučaju algoritam potvrdio da je ispunjen zadani uvjet pokusa, te da primjenom kineziološkog tretmana nije došlo do strukturalnih promjena u skupovima izmjerenih varijabli.

Tabela 1 - SKLOP PROMAX FAKTORA

	PRX 1	PRX 2	PRX 3	PRX 4
1 CUC	-.0186	.1924	-.1312	.5910
1 DSM	.0266	.8987	.1469	-.0587
1 DTK	.8332	.0043	-.0823	.0728
1 SKL	.1134	-.1780	.2738	.7195
1 ZGIB	-.1506	.0486	.8770	.0725
1 NABAC	-.1534	-.0167	-.1366	.6223
2 CUC	-.0703	.0971	-.1383	.6839
2 DSM	.0471	.9019	.0184	.0627
2 DTK	.9604	.0587	-.0387	-.0481
2 SKL	.1300	-.1455	.2988	.6136
2 ZGIB	.0336	.1220	.8908	-.0648
2 NABAC	-.0308	.1233	-.0886	.4721

Tabela 2 - RAZLIKE KOORDINATA (t - (t+1)) TRANZITIVNIH STANJA

	PRX 1	PRX 2	PRX 3	PRX 4
1 CUC	.0517	.0953	.0070	-.0929
1 DSM	-.0205	-.0032	.1285	-.1214
1 DTK	-.1272	-.0544	-.0436	.1209
1 SKL	-.0165	-.0324	-.0250	.1059
1 ZGIB	-.1842	-.0732	-.0138	.1374
1 NABAC	-.1225	-.1400	-.0480	.1503

Tabela 3 - KOSINUSI KUTEVA PARTICIJA PROMAX FAKTORA

	PRX 1	PRX 2	PRX 3	PRX 4
PRX 1	.9661	.0213	-.1397	-.0490
PRX 2	.0131	.9783	-.0206	.0693
PRX 3	.0069	.1851	.9882	-.0303
PRX 4	.0933	-.0098	.1291	.9645

Tabela 4 - STRUKTURA PROMAX FAKTORA

	PRX 1	PRX 2	PRX 3	PRX 4
1 CUC	.1005	.3375	.0299	.5947
1 DSM	.0614	.8829	.1310	.2097
1 DTK	.8235	.0201	.2476	.2877
1 SKL	.4206	-.0061	.5245	.7873
1 ZGIB	.1952	.0583	.8415	.2937
1 NABAC	-.0266	.1365	-.0144	.5353
2 CUC	.0730	.2651	.0313	.6477
2 DSM	.0689	.9168	.0450	.3008
2 DTK	.9322	.0442	.3030	.2288
2 SKL	.4161	.0003	.5249	.7011
2 ZGI	.3451	.0973	.8835	.2306
2 NABAC	.0704	.2391	.0344	.4678

Tabela 5 - KORELACIJE OBLIMIN FAKTORA

	PRX 1	PRX 2	PRX 3	PRX 4
PRX 1	1.0000			
PRX 2	-.0033	1.0000		
PRX 3	.3709	-.0099	1.0000	
PRX 4	.2850	.2432	.2875	1.0000

LITERATURA

1. Bosnar, K., F. Prot, K. Momirović, V. Lužar i V. Dobrić. *Algoritam za procjenu pseudokanoničkih faktora.* Kineziologija, 13(1982)29-34
2. Cattel, R.B.. *Handbook of multivariate experimental psychology.* McNally, Chicago, 1966
3. Fulgosi, A. *Faktorska analiza.* Školska knjiga, Zagreb, 1979
4. Harman, H.H.. *Modern factor analysis.* University of Chicago Press, Chicago, 1960
5. Harris, C.W. *Canonical factor models in the description of change (IN:* C.W. Harris, *Problems in measuring change,* University of Wisconsin Press, Madison, 1963*)*
6. Hošek-Momirović, A. *Utjecaj socioloških karakteristika na motoričke sposobnosti.* Kineziologija, 9(1979)107-124
7. Kaiser, H. *The varimax criterion for analytic rotations in factor analysis.* Psychometrika,23(1958)187-200
8. Momirović, K. *Metode za transformaciju i kondenzaciju kinezioloških informacija.* Institut za kineziologiju, Zagreb, 1972
9. Momirović, K. i J. Štalec. *DMEAN i DMAX kriteriji za odredjivanje broja značajnih image faktora pri analizi zadataka u psihologijskim testovima.* Stručni skupovi "Dani Ramira Bujasa", 1970 i 1972, Društvo psihologa Hrvatske, Zagreb, 1973(95).
10. Momirović, K.,J. Štalec and E. Zakrajšek. *A programming language for multivariate data analysis.* COMPSTAT 82. I. Proceedings in computational statistics, Phisica Verlag, Wien, 1982, 87-95
11. Mulaik, S.A. *The foundations of factor analysis.* McGraw-Hill, New York, 1972
12. Zakrajšek, E., J. Štalec i K. Momirović. *SS-Programski jezik za multivarijatnu analizu podataka.* Zbornik simpozija "Kompjuter na sveučilištu", Zagreb, 1974, C8.1-16

2.9 ALGORITAM I PROGRAM QUQULELE

Asortiman sadržaja sportskog treninga višestruko je ograničen, bilo da su ograničenja uvjetovana mjestom održavanja treninga, raspoloživom opremom, ili vremenom trajanja i sl., pa su treneri često suočeni sa problemom odabira odredjenog broja iz ograničenog izbora sadržaja treninga.

Objektivna taksonomizacija sadržaja transformacijskih operatora, ako se njihov odredjeni broj odabire iz nekog ograničenog broja mogućih sadržaja, znatno olakšava rješenje problema strukturiranja sadržaja treninga u danim tipičnim uvjetima. Time se znatno povećava operativna efikasnost i kreativna uloga trenera, a znatno smanjuje vjerojatnost utjecaja pogreške subjektivne procjene.

Algoritam i pridruženi program QUQULELE efikasno rješava problem taksonomizacije sadržaja transformacijskih operatora opisanih fiksnim brojem iz ograničenog broja atributa.

```
*
*    NAZIV PROGRAMA
*
*              *** QUQULELE ***
*
*
*    AUTORI
*
*              F. PROT
*              K. BOSNAR
*              K. MOMIROVIC
*
*
*    FUNKCIJA
*
*              PROGRAM QUQULELE IZVODI TAKSONOMSKU ANALIZU NEKOG SKUPA
*         ENTITETA KOJI SU OPISANI IZBOROM FIKSNOG BROJA OBJEKATA IZ NEKOG
*         SKUPA. MJERE SLICNOSTI ENTITETA DEFINIRANE SU U NORMIRANOJ HAMMING
*         METRICI, KAO SKALARNI PRODUKTI VEKTORA ENTITETA JEDINICNE DUZINE,
*         OPISANIH NAD OBJEKTIMA IZBORA. OSNOVNA SOLUCIJA JE ODREDJENA
*         KARAKTERISTICNIM VEKTORIMA MATRICE MJERA SLICNOSTI, NORMIRANIM
*         NA PRIPADNE SVOJSTVENE VRIJEDNOSTI. BROJ TAKSONOMSKIH DIMENZIJA
*         DEFINIRAN JE BROJEM SVOJSTVENIH VRIJEDNOSTI VECIH OD PROSJEKA
*         NENULTIH SVOJSTVENIH VRIJEDNOSTI ILI UNAPRIJED FIKSIRANIM BROJEM
*         DIMENZIJA POD NEKOM HIPOTEZOM. TAKSONOMSKE DIMENZIJE FORMIRANE SU
*         ORTONORMALNIM TRANSFORMACIJAMA DIMENZIJA KOJE TVORE OSNOVNI KOO-
*         RDINATNI SUSTAV. PROGRAM ODREDJUJE MATRICE SKLOPA, STRUKTURE I
*         RELACIJA TAKSONOMSKIH DIMENZIJA. PROGRAM IZVODI DVIJE REGRESIJSKE
*         ANALIZE, POD MODELOM NAJMANJIH KVADRATA, ANALIZU OBJEKATA IZBORA
*         U PROSTORU TAKSONOMSKIH DIMENZIJA I ANALIZU TAKSONOMSKIH DIMENZIJA
*         U PROSTORU IZBORA.
*
*
*
*    ALGORITAM, OPIS PROGRAMA I PRIMJER PRIMJENE OVOG PROGRAMA JE U
*    RADU
*
*         PROT, F., A.A. ZENKIN, K. BOSNAR, K. MOMIROVIC I Z. KNAP
*         "AN ALGORITHM FOR TAXONOMIC ANALYSIS OF DATA OBTAINED IN
*         FIXED NUMBER FORCED CHICE EXPERIMENTS"
*         "C O M P S T A T - 84", PRAG 1984.
*         ALGORITAM I PROGRAM SU REZULTAT ZAJEDNICKOG RADA MATEMATICARA
*
*         I INFORMATICARA IZ INSTITUTA ZA SOCIOLOGIJO U LJUBLJANI, SVEUCILIS-
*         NOG RACUNSKOG CENTRA U ZAGREBU, FAKULTETA ZA FIZICKU KULTURU  U
*         ZAGREBU I VYCISLITELJNOGO CENTRA AKADEMII NAUK SSSR, MOSKVA.
*
*         ZA ALGORITAM I KOD SU ODGOVORNI KNAP I MOMIROVIC,
*         A ZA ODRZAVANJE I KONZULTACIJE PROT I BOSNAR.
*
*    PRIPREMA PODATAKA
*
*              KONTROLNI ZAPIS I PODACI U FILEU DATA.
*
*              OVA VERZIJA PROGRAMA QUQULELE MOZE ANALIZIRATI DO 250
*         ENTITETA OPISANIH NAD DO 250 OBJEKATA UZ UVJET DA JE FIKSNI BROJ
*         IZBORA VECI OD 1 A MENJI OD M-1, GDJE JE M BROJ VARIJABLI.
*
*
*    IZVRSNI DIO PROGRAMA
*
*      BLOK 1. PRETHODNE TRANSFORMACIJE I FORMIRANJE MATRICE
*              MJERA SLICNOSTI.
*
```

```
OUTPUT(DEVICE=PR)
HEADING(TEXT=Q U Q U L E L E,T)
HEADING(TEXT=PROGRAM ZA TAKSONOMIJU ENTITETA S FIKSNIM BROJEM IZBORA)
HEADING(TEXT=VERZIJA 1=0.0)
TEXT(TEXT=             QUQULELE)
TEXT(TEXT=MATRICA   MJERA     SLICNOSTI)
INPUT(DATA=DATA)
SORT(IN=SCORE,OUT=B)
MULT(A=B,B=B.TB,M=SS)
SCALE(C=SS,R=S)
              DELETE(MATRIX=SS)
PRINT(MATRIX=S,TEXT=MATRICA MJERA SLICNOSTI)
*
*       BLOK 2. OSNOVNE ORTOGONALNE DIMENZIJE. U NAREDBI
*               HOTELLING, MIN=1/M, M=BROJ VARIJABLI.
*
TEXT(TEXT=ORTO-     GONALNE   DIMENZIJE)
DIAGONALISATION(R=S)
HOTELLING(NUM=3,F=HT)
HOTELLING(NUM=1,F=WO)
              DELETE(MATRIX=LAMBDA)
              DELETE(MATRIX=X)
TRANSPOSE(OLD=HT,NEW=H)
MULT(A=HT,B=H,M=LAMBDA)
PRINT(MATRIX=H,TEXT=ORTOGONALNE DIMENZIJE)
RESIDUAL(R=S,F=HT)
PRINT(MATRIX=RES,TEXT=REZIDUALI MJERA SLICNOSTI,D)
*
*       BLOK 3. TAKSONOMSKE DIMENZIJE.
*
TEXT(TEXT=KOSE     TAXONOMSKEDIMENZIJE)
VARIMAX(F=HT,TAU=QT,FN=TT,N)
TRANSPOSE(OLD=QT,NEW=Q)
MULT(A=QT,B=LAMBDA,M=QTL)
MULT(A=QTL,B=Q,M=REL)
SCALE(C=REL,R=M)
              DELETE(MATRIX=QTL)
PRINT(MATRIX=TT,T,TEXT=TAKSONOMSKE DIMENZIJE)
PRINT(MATRIX=M,TEXT=KOSINUSI KUTEVA TAKSONOMSKIH DIMENZIJA)
*
*       BLOK 4. SKLOP I STRUKTURA TAKSONOMSKIH DIMENZIJA.
*
TEXT(TEXT=SKLOP  I  STRUKTURA TAKSONA)
MULT(A=TT,B=B,M=VOLINA)
MULT(A=B,TA,B=B,M=SVINJA)
DIAGMULT(A=VOLINA,T,D=SVINJA,C=-0.5,L,M=KRAVA)
DIAGMULT(A=KRAVA,T,D=REL,C=-0.5,L,M=F)
INVERSION(R=M,RINV=TINV)
MULT(A=F,TA,B=TINV,M=A)
PRINT(MATRIX=A,TEXT=SKLOP TAKSONOMSKIH DIMENZIJA)
PRINT(MATRIX=F,T,TEXT=STRUKTURA TAKSONOMSKIH DIMENZIJA)
*
*       BLOK 5. REGRESIJSKA ANALIZA VARIJABLI U PROSTORU
*               TAKSONOMSKIH DIMENZIJA.
*
TEXT(TEXT=REGRESISKAANALIZA   VARIJABLI)
HADMULT(A=WO,C=0.0,M=ZZZ)
TRANSPOSE(OLD=ZZZ,NEW=ZVIZ)
TRANSPOSE(OLD=TT,NEW=TTT)
MERGE(IN1=ZVIZ,IN2=TTT,OUT=JTT)
MULT(A=JTT,TA,B=JTT,M=JTJ)
INVERSION(R=JTJ,RINV=IJTJ)
MULT(A=IJTJ,B=JTT,TB,M=IJTT)
MULT(A=IJTT,B=B,M=BETA)
```

```
MULT(A=B,TA,B=JTT,M=BJTT)
MULT(A=BJTT,B=BETA,M=ROVAR)
MULT(A=JTT,B=BETA,M=JTTZ)
MULT(A=JTTZ,TA,B=JTTZ,M=BJEL)
DIAGMULT(A=ROVAR,D=SVINJA,C=-0.5,L,M=ROVA)
DIAGMULT(A=ROVA,D=BJEL,C=-0.5,R,M=ROV)
PRINT(MATRIX=JTTZ,TEXT= REPRODUKCIJA PRISILNIH IZBORA)
PRINT(MATRIX=BETA,TEXT= REGRESIJSKI KOEFICIJENTI U PROSTORU TAX DIMENZIJA)
PRINT(MATRIX=ROVAR,TEXT=MULTIPLE KOVARIANCE VARAIABLI)
PRINT(MATRIX=ROV,TEXT=MULTIPLE KORELACIJE VARIJABLI)
*
*      BLOK 6. REGRESIJSKA ANALIZA TAKSONOMSKIH DIMENZIJA U
*              PROSTORU VARIJABLI
*
MERGE(IN1=ZVIZ,IN2=B,OUT=JB)
MULT(A=JB,TA,B=JB,M=JBBJ)
INVERSION(R=JBBJ,RINV=JIIJ)
MULT(A=JIIJ,B=JB,TB,M=JIIB)
MULT(A=JIIB,B=TT,TB,M=BETAV)
MULT(A=TT,B=JB,M=TTJB)
MULT(A=TTJB,B=BETAV,M=ROTAX)
MULT(A=JB,B=BETAV,M=JBZ)
MULT(A=JBZ,TA,B=JBZ,M=CRN)
DIAGMULT(A=ROTAX,D=REL,C=-0.5,L,M=ROTA)
DIAGMULT(A=ROTA,D=CRN,C=-0.5,R,M=ROT)
PRINT(MATRIX=JBZ,TEXT= REPRODUKCIJA TAKSONOMSKIH SKOROVA)
PRINT(MATRIX=BETAV,TEXT=REGRESIJSKI KOEFICIJENTI U PROSTORU VARIJABLI)
PRINT(MATRIX=ROTAX,TEXT= MULTIPLE  KOVARIJANCE  TAX DIMENZIJA)
PRINT(MATRIX=ROT,TEXT= MULTIPLE KORELACIJE TAX DIMENZIJA)
*
*
*         KRAJ PROGRAMA   G U G U L E L E
*
```

ALGORITAM ZA TAKSONOMSKU ANALIZU MATRICA PODATAKA DOBIJENIH U EKSPERIMENTIMA S FIKSNIM BROJEM PRISILNIH IZBORA*

Prot, F., A.A. Zenkin, K. Bosnar, K. Momirović i Ž. Knap

SAŽETAK

Predložen je algoritam i napisan program koji provodi taksonomsku analizu skupa entiteta opisanih izborom fiksnog broja objekata iz nekog skupa. Bazična solucija odredjena je svojstvenim vektorima matrice mjera sličnosti formirane skalarnim produktima normiranih vektora entiteta u Hammingovoj metrici. Finalna solucija odredjena je ortonormalnim transformacijama dimenzija bazične solucije.

** Ovaj rad je izvorno napisan na engleskom jeziku i bit će prezentiran na Medjunarodnom simpoziju računarske statistike, COMPSTAT '84, Prag, 1984.*

1. UVOD

Algoritam i program QUQULELE izvode taksonomsku analizu nekog skupa entiteta koji su opisani izborom fiksnog broja objekata iz nekog skupa. Mjere sličnosti entiteta definirane su u normiranoj Hamming metrici, kao skalarni produkti vektora entiteta jedinične dužine, opisanih nad objektima izbora. Osnovna solucija je odredjena karakterističnim vektorima matrice mjera sličnosti, normiranim na pripadne svojstvene vrijednosti. Broj taksonomskih dimenzija definiran je brojem svojstvenih vrijednosti većih od prosjeka nenultih svojstvenih vrijednosti ili unaprijed fiksiranim brojem dimenzija pod nekom hipotezom. Taksonomske dimenzije formirane su ortonormalnim transformacijama dimenzija koje tvore osnovni koordinatni sustav. Program odredjuje matrice sklopa, strukture i relacija taksonomskih dimenzija. Program izvodi dvije regresijske analize, pod modelom najmanjih kvadrata, analizu objekata izbora u prostoru taksonomskih dimenzija i analizu taksonomskih dimenzija u prostoru izbora.

2. ALGORITAM

Neka je $E = \{e_i;\ i=1,...,n\}$ skup takmičara iz neke populacije P, i neka je $O = \{o_j;\ j=1,...,m\}$ skup konačnog broja sadržaja transformacijskih operatora izvučen iz nekog univerzuma U, u jednoj fazi trenažnog ciklusa T.

Neka je za svakog $e_i \in E$ izabran fiksni broj k; $1 \leq k \leq m-1$ sadržaja tranzitivnih operatora $o_j \in O$. Neka je rezultat operacije izbora kodiran tako da je svakom e_i pridružen opis definiran vektorom $B^T=(b_{ij})=\{b_{i1},...,b_{im}\}$. Za svaki od k odabranih sadržaja o_j u vektoru B_i^T neka je b =1, a za ostalih (m-k) neizabranih sadržaja neka je, u vektoru B_i^T $b_{ij}=0$.

Organizirajmo vektore kodiranih izbora u matricu

$$B = (b_{ij}); \; i=1,\ldots,n, \; j=1,\ldots,m,$$

tako da vektori B_i^T čine retke formirane matrice. Matrica B je potpun opis trenažnog ciklusa T, u okviru kojeg je nad svakim takmičarem $e_j \in E$ primijenjen fiksni broj k konačnog broja sadržaja $o_j \in O$.

2.1 Mjera sličnosti

Normiranom Hammingovom metrikom tj. skalarnim produktom vektora entiteta jedinične dužine, opisanih objektima izbora

$$\hat{C} = D^{-1}C^*D^{-1}$$

gdje je

$$C^* = BB^T$$

$$D^2 = \text{diag } C^*$$

odredjena je prirodna mjera sličnosti entiteta $e_i \in E$ u prostoru izbora fiksnog broja sadržaja $o_j \in O$.

2.2 Osnovna ortogonalna solucija

Neka se matrica mjera sličnosti dekomponira na produkt

$$\hat{C} = \hat{Y}\,\hat{\Lambda}\,\hat{Y}^T$$

u kojem su u matrici $\hat{Y}=(\hat{y}_i)$; $i=1,\ldots,n$ lijevi svojstveni vektori, po retcima normirane, matrice B, a u dijagonalnoj matrici $\hat{\Lambda}=(\hat{\lambda}_i)$; $i=1,\ldots,n$, njima pridružene svojstvene vrijednosti.

Matrica $\hat{C}$ se može razložiti na dva sumanda

$$\hat{C} = C + \tilde{C}.$$

Prvi sumand definiran je produktom

$$C = Y \Lambda Y^T$$

gdje je $\Lambda=(\lambda_p)$; $p=1,\ldots,q$ matrica prvih q svojstvenih vrijednosti, poredanih po veličini, a $Y=(y_p)$ matrica njima pridruženih lijevih svojstvenih vektora po retcima normirane matrice B. Broj q može biti odredjen ovako:

a/ q = broj svojstvenih vrijednosti većih od prosjeka nenultih svojstvenih vrijednosti

ili

b/ q = apriori fiksiran u skladu s nekom hipotezom, u pravilu s nekom hipotezom o taksonomiji sadržaja tranzitivnih operatora.

Drugi sumand je odredjen produktom

$$\tilde{C} = \tilde{Y} \tilde{\Lambda} \tilde{Y}^T$$

gdje je $\tilde{\Lambda}=(\tilde{\lambda}_t)$; $t=q+1,\ldots,n$ matrica preostalih svojstvenih vrijednosti i njima pridruženih svojstvenih vektora $\tilde{Y}=(\tilde{y}_t)$.

Koordinate entiteta na ortogonalnim dimenzijama odredjuje operacija

$$H = Y \Lambda^{1/2}.$$

Matricom reziduala koeficijenta sličnosti

$$W = C - HH^T = \tilde{C}$$

predočen je onaj dio sličnosti entiteta koji se ne može reproducirati sa q zadržanih dimenzija.

2.3 Finalna neortogonalna solucija

Transformacija

$$HQ = T = (t_{ip});\ i=1,\ldots,n,\quad p=1,\ldots,q$$

uz uvjet

$$Q^TQ = QQ^T = I$$

koja minimizira parsimonijsku funkciju

$$\omega = n \sum_{i=1}^{n} \sum_{p=1}^{q} t_{ip}^4 - \sum_{p=1}^{q} \Big(\sum_{i=1}^{n} t_{ip}^2 \Big)^2$$

formira kosokutni taksonomski koordinatni sustav entiteta, gdje je Q, ortonormalna transformacijska matrica, istovremeno matrica relacija ortogonalnog i kosog koordinatnog sustava.

U matrici

$$M = (\text{diag } T^TT)^{-1/2}T^TT(\text{diag } T^TT)^{-1/2}$$

nalaze se relacije kosih taksonomskih dimenzija, definirane kosinusima kuteva izmedju taksonomskih dimenzija.

2.4 Struktura i sklop taksonomskih dimenzija

Matricom

$$F = (\mathrm{diag}\ T^T T)^{-1/2} T^T B (\mathrm{diag}\ B^T B)^{-1/2}$$

odredjene su ortogonalne koordinate vektora objekata izbora u prostoru kosih taksonomskih dimenzija, dakle struktura taksona.

Sklop taksonomskih dimenzija definiran paralelnim projekcijama vektora objekata izbora na kose taksonomske dimenzije odredjen je operacijom

$$A = F\, M^{-1}.$$

2.5 Regresijska analiza izbora sadržaja u prostoru taksonomskih dimenzija

Neka je matrica Ψ formirana tako da je matrici T pridružen vektor sa n jedinica, tj.

$$\Psi = \begin{bmatrix} 1 & \vdots & T \end{bmatrix}$$

i neka je

$$\Psi\beta = B + E \quad \Big| \quad tr(E^T E) = \min$$

gdje su u matrici β regresijski koeficijenti, a u matrici E reziduali. Pod modelom najmanjih kvadrata, regresijski koeficijenti

$$\beta = (\Psi^T \Psi)^{-1} \Psi^T B$$

omogućava reprodukciju izbora sadržaja

$$B^* = \Psi\beta = \Psi(\Psi^T\Psi)^{-1}\Psi^T B$$

u prostoru taksonomskih dimenzija.

Multiple korelacije prisilnih izbora i taksonomskih dimenzija odredjene su operacijom

$$\tilde{\rho} = (\text{diag } B^T B)^{-1/2} B^T\Psi(\Psi^T\Psi)^{-1}\Psi^T B \; (\text{diag } B^{*T}B^*)^{-1/2}.$$

2.6 <u>Regresijska analiza taksonomskih dimenzija u prostoru izbora sadržaja</u>

Neka je matrica S formirana tako da je matrici B pridružen vektor sa n jedinica tj.

$$S = \begin{bmatrix} 1 & \vdots & B \end{bmatrix}$$

i neka je

$$S\beta^* = T + \varepsilon \;\Big|\; tr(E^T E) = min$$

gdje su u matrici β^* regresijski koeficijenti, a u matrici ε reziduali. Pod modelom najmanjih kvadrata, regresijski koeficijenti

$$\beta^* = (S^T S)^{-1} S^T T$$

omogućavaju reprodukciju taksonomskih vrijednosti entiteta operacijom

$$T^* = S\beta^* = S(S^T S)^{-1} S^T T,$$

Multiple korelacije taksonomskih dimenzija i izbora sadržaja odredjene su operacijom

$$\rho^* = (\text{diag } T^T T)^{-1/2} T^T S(S^T S)^{-1} S^T T (\text{diag } T^{*T} T^*)^{-1/2}.$$

3. PROGRAM

Prezentirani algoritam je inplementiran u programu QUQULELE napisanom u verziji 5.2/M programskog sistema SS. Program je pohranjem u bibliotekama FFK*WORK i SRCE*SS-MAKRO. Aktiviranje programa i neophodan opis podataka nalaze se u priručniku za korištenje programskog sistema SS (Štalec, Momirović i Zakrajšek, 1983).

QUQULELE može analizirati do 250 entiteta opisanih ograničenim brojem izbora iz kolekcije do 250 objekata.

4. NUMERIČKI PRIMJER

Sa ciljem provjere efikasnosti algoritma i programa, konstruiran je primjer gdje je matrica podataka formirana tako da sadrži tri taksona od po 10 entiteta, te 10 entiteta-distraktora sa izborima koji ne odgovaraju definiciji niti jednog od tri taksona.

Kao što je vidljivo usporedbom rezultata u tabelama 1 i 2, programom su ispravno identificirana sva tri taksona i izvršena je ispravna klasifikacija entiteta, što govori u prilog valjanosti algoritma i programa.

Tabela 1 - ULAZNI PODACI; sa DIT su označeni distraktori

	OBJ1	OBJ2	OBJ3	OBJ4	OBJ5	OBJ6	OBJ7	OBJ8	OBJ9	OBJ10
DIT001	1	0	0	0	0	0	1	0	0	1
DIT002	1	0	0	1	0	0	1	0	0	0
DIT003	0	0	1	0	0	0	0	1	0	1
DIT004	0	1	0	0	1	0	0	1	0	0
DIT005	0	0	0	0	0	1	1	0	0	1
DIT006	0	0	1	0	0	1	0	0	1	0
DIT007	0	1	0	0	1	0	0	0	0	1
DIT008	1	0	0	1	0	0	0	1	0	0
DIT009	0	0	0	1	0	0	1	0	0	1
DIT010	1	0	0	0	1	0	0	0	1	0
MIKI01	0	0	0	1	1	1	0	0	0	0
MIKI02	0	0	0	1	1	1	0	0	0	0
MIKI03	0	0	0	1	1	1	0	0	0	0
MIKI04	0	0	0	1	1	1	0	0	0	0
MIKI05	0	0	0	1	1	1	0	0	0	0
MIKI06	0	0	0	1	1	1	0	0	0	0
MIKI07	0	0	0	1	1	1	0	0	0	0
MIKI08	0	0	0	1	1	1	0	0	0	0
MIKI09	0	0	0	1	1	1	0	0	0	0
MIKI10	0	0	0	1	1	1	0	0	0	0
NINA01	0	0	0	0	0	0	1	1	1	0
NINA02	0	0	0	0	0	0	1	1	1	0
NINA03	0	0	0	0	0	0	1	1	1	0
NINA04	0	0	0	0	0	0	1	1	1	0
NINA05	0	0	0	0	0	0	1	1	1	0
NINA06	0	0	0	0	0	0	1	1	1	0
NINA07	0	0	0	0	0	0	1	1	1	0
NINA08	0	0	0	0	0	0	1	1	1	0
NINA09	0	0	0	0	0	0	1	1	1	0
NINA10	0	0	0	0	0	0	1	1	1	0
PEPE01	1	1	1	0	0	0	0	0	0	0
PEPE02	1	1	1	0	0	0	0	0	0	0
PEPE03	1	1	1	0	0	0	0	0	0	0
PEPE04	1	1	1	0	0	0	0	0	0	0
PEPE05	1	1	1	0	0	0	0	0	0	0
PEPE06	1	1	1	0	0	0	0	0	0	0
PEPE07	1	1	1	0	0	0	0	0	0	0
PEPE08	1	1	1	0	0	0	0	0	0	0
PEPE09	1	1	1	0	0	0	0	0	0	0
PEPE10	1	1	1	0	0	0	0	0	0	0

Tabela 2 - FINALNA TAKSONOMSKA SOLUCIJA - REZULTATI ENTITETA NA TAKSONOMSKIM DIMENZIJAMA

	VRX 1	VRX 2	VRX 3
DIT001	.0456	.3698	.4060
DIT002	.3515	.3399	.3707
DIT003	.0120	.3616	.3699
DIT004	.3298	.3448	.3090
DIT005	.3632	.0136	.3865
DIT006	.3082	.3100	.3033
DIT007	.3664	.3676	.0008
DIT008	.3382	.3514	.3560
DIT009	.3736	.0240	.3978
DIT010	.3383	.3520	.3302
MIKI01	.9980	-.0002	.0000
MIKI02	.9980	-.0002	.0000
MIKI03	.9980	-.0002	.0000
MIKI04	.9980	-.0002	.0000
MIKI05	.9980	-.0002	.0000
MIKI06	.9980	-.0002	.0000
MIKI07	.9980	-.0002	.0000
MIKI08	.9980	-.0002	.0000
MIKI09	.9980	-.0002	.0000
MIKI10	.9980	-.0002	.0000
NINA01	-.0051	-.0030	.9957
NINA02	-.0051	-.0030	.9957
NINA03	-.0051	-.0030	.9957
NINA04	-.0051	-.0030	.9957
NINA05	-.0051	-.0030	.9957
NINA06	-.0051	-.0030	.9957
NINA07	-.0051	-.0030	.9957
NINA08	-.0051	-.0030	.9957
NINA09	-.0051	-.0030	.9957
NINA10	-.0051	-.0030	.9957
PEPE01	-.0033	.9978	-.0016
PEPE02	-.0033	.9978	-.0016
PEPE03	-.0033	.9978	-.0016
PEPE04	-.0033	.9978	-.0016
PEPE05	-.0033	.9978	-.0016
PEPE06	-.0033	.9978	-.0016
PEPE07	-.0033	.9978	-.0016
PEPE08	-.0033	.9978	-.0016
PEPE09	-.0033	.9978	-.0016
PEPE10	-.0033	.9978	-.0016

LITERATURA

1. Cattell, R.B.:
Handbook of multivariate experimental psychology, Rand mcNally, Chicago, 1966.

2. Hamming, R.W.:
Coding and information theory, Prentice-Hall, New York, 1980.

3. Horst, P.:
Factor analysis of data matrices, Holt, Rinehart and Winston, New York, 1965.

4. Zakrajšek, E., J. Štalec i K. Momirović
SS-programski sistem za multivarijatnu analizu podataka, 1st International symposium "Computer at the University", Proceedings, Zagreb, 1974, C8, 1-18.

2.10 ALGORITAM I PROGRAM LAV

Izmedju različitih metoda za analizu krivulja vježbe u posljednje vrijeme, zbog svoje jednostavnosti i neposredne povezanosti sa većinom drugih metoda za analizu kinezioloških informacija, faktorska analiza multivarijatnih trajektorija postaje sve popularnija u analizi podataka dobijenih registracijom promjena koje nastaju u toku treninga.

Ove promjene, medjutim, mogu u mnogim slučajevima biti maskirane djelovanjem slučajnih varijabli pogreške. Zbog toga je predložen algoritam i napisan program koji omogućava analizu kvantitativnih promjena u toku treninga, ili analizu promjena nekog kvantitativnog obilježja uopće, nakon eliminacije varijabli pogreške. Iako mu je osnovna namjena analiza krivulja vježbe, može se primijeniti i u analizi razvoja bilo kojih karakteristika, pa stoga i karakteristike koje opisuju stanje i dinamiku promjena u području fizičke kulture.

```
*
*   NAZIV PROGRAMA
*
*
*                        ***************
*                        *    L A V    *
*                        ***************
*
*
*   AUTORI
*
*
*            L. PAVICIC
*            Z. KARAMAN
*            K. MOMIROVIC
*
*
*   FUNKCIJA
*
*
*   LAV JE PROGRAM ZA IMAGE ANALIZU MULTIVARIANTNIH TRAJEKTORIJA BAZIRAN
*   NA BENTLEROVOJ(BENTLER, 1969) GENERALIZACIJI IMAGE ANALIZE.
*
*
*
*
*   METODA I ALGORITAM OPISANI SU U
*
*            PAVICIC, L., Z. KARAMAN I K. MOMIROVIC:
*            IMAGE ANALYSIS OF MULTIVARIATE TRAJECTORIES,
*            PROCEEDINGS OF 5TH INTERNATIONAL SYMPOSIUM
*            "COMPUTER AT THE UNIVERSITY",CAVTAT,1983,PP.491-497.
*
*   PRIPREMA PODATAKA
*
*   U OVOJ VERZIJI SS JEZIKA, DOZVOLJEN JE EFEKTIV UZORKA DO
*   10 000, BROJ VREMENSKIH TOCAKA DO 250.
*
*
*
*
*
*   PRIJE UPOTREBE PROGRAMA POTREBNO JE PRIREDITI:
*
*   (1) DATOTEKU DATA SA POTPUNIM ORIGINALNIM PODACIMA U SVIM
*       VREMENSKIM TOCKAMA.
*
*   (2) DATOTEKU ATA SA PODACIMA U PRVOJ VREMENSKOJ TOCI.
*
*
*   (3) KONTROLNE ZAPISE PRIDRUZENE PODACIMA POD (1) I (2).
*
*
*
*   IZVRSNI DIO PROGRAMA
*
*
*   BLOK 0.
*   INPUT, CONTROL, SORT AND CONFORM ULAZNIH PODATAKA.
*
OUTPUT(DEVICE=PR)
HEADING(TEXT= L   A   V ,T)
HEADING(TEXT=IMAGE ANLIZA MULTIVARIANTNIH TRAJEKTORIJA)
HEADING(D)
```

```
TEXT(TEXT=                LAV)
INPUT(SCORE=TPA.DATA=DATA)
INPUT(SCORE=TP1.DATA=ATA)
CONFORM(IN1=TPA.IN2=TP1.OUT1=W.OUT2=WT)
         DELETE(MATRIX=TPA)
         DELETE(MATRIX=TP1)
      CORRELATION(SCORE = W, C = CX, R = RX)
      INVERSION(R=RX.RINV=XX.SMC=EF)
      HADMULT(A = EF, C = 0.0, M = F1)
      HADMULT(A = W,  C = 0.0, M = JEDAN)
      DELETE(MATRIX = CX)
      DELETE(MATRIX = RX)
      DELETE(MATRIX = XX)
      HADMULT(A = WT, C = 0.0, M = N1)
      MULT(A = N1, TA, B = WT, M = SUM)
      MULT(A = N1, TA, B = N1, M = EN)
      DIAGMULT(A = F1, D = EN, L , C = -1.0, M = EN1)
      DIAGMULT(A = SUM  , D = EN,   L , C = -1.0, M = AS)
      CORRELATION(SCORE = WT, C = VAR, R = RX)
      DIAG(A = VAR, C = 0.5, D = SD)
      MULT(A = F1, TA, B = AS, M = F1AS)
      MULT(A = F1, TA, B = SD, M = F1SD)
      DIAGMULT(A = JEDAN, D = F1AS, R ,             M = CENT)
      LINEAR(A = W, B = CENT, CB = -1.0, M = WC)
      DIAGMULT(A = WC,    D = F1SD, R , C = -1.0, M = CW)
      DELETE(MATRIX = WT)
      DELETE(MATRIX = N1)
      DELETE(MATRIX = RX)
      DELETE(MATRIX = CENT)
      DELETE(MATRIX = WC)
*
*   BLOK 1.
*
*
*
HEADING(TEXT = RAW.D)
STATISTICS(SCORE = W,  Z = ZX, S)
      DELETE(MATRIX = ZX)
HEADING(TEXT = CENTERED.D)
STATISTICS(SCORE = CW, Z = ZX, S)
      DELETE(MATRIX = ZX)
CORRELATION(SCORE = CW, C = CX, R = R)
HEADING(D)
PRINT(MATRIX = R, TEXT = KORELACIJE MJERENJA)
* FORM H
*=======
      MULT(A = CW, TA, B = CW, M = NH)
DIAGMULT(A = NH,  D = EN1.R , M = H)
      INVERSION(R = H, RINV = IH, SMC = SMCX)
*************************************************************************
* FORM U
*=======
      DIAGMULT(A = IH,    D = IH  , R , C = -1.0,  M = IHD)
MULT(A = CW, B = IHD, M = U)
* FORM C
*=======
LINEAR(A = CW, B = U, CB = -1.0, M = C)
* FORM G
*=======
      MULT(A = C , TA , B = C, M = GN)
DIAGMULT(A = GN,  D = EN1.R , M = G)
* FORM E
*=======
      MULT(A = U, TA, B = U , M = NE)
```

```
DIAGMULT(A = NE, D = EN1, R, M = E)
* FORM A
*=======
*
DIAGMULT(A = IH,    D = IH,    L , R, C = -0.5,  M = A)
* FORM F
*=======
      DIAGONALISATION(R = G, LAMBDA = DELTA, X = X)
HOTELLING(X = X, LAMBDA = DELTA, F = F)
* FORM K
*=======
      MULT(A = F, B = F, TB , M = LL)
      DIAGMULT(A = F, T, D = LL, R, C = -0.5, M = XX)
      MULT(A = C, B = XX, M = K)
***********************************************************************
*
* PRINT : G,K,F
PRINT(MATRIX=G)
PRINT(MATRIX=F,T)
PRINT(MATRIX=XX)
PRINT(MATRIX=K)
*
* CALC SUM OF SQUARES OF ELEMENTS OF: G,E,A,H
      HADMULT(A = G, C = 2.0 , M = G2)
      MULT(A = F1, B = G2, M = G1)
      MULT(A = F1, B = G1, TB,M = SUMG2)
      HADMULT(A = E, C = 2.0 , M = E2)
      MULT(A = F1, B = E2, M = E1)
      MULT(A = F1, B = E1, TB,M = SUME2)
      HADMULT(A = A, C = 2.0 , M = A2)
      MULT(A = F1, B = A2, M = A1)
      MULT(A = F1, B = A1, TB,M = SUMA2)
      HADMULT(A = H, C = 2.0 , M = H2)
      MULT(A = F1, B = H2, M = H1)
      MULT(A = F1, B = H1, TB,M = SUMH2)
*
* CALC ALPHA
      LINEAR(A = SUME2, B = SUMG2, M = EPG)
      DIAGMULT(A = SUME2, D = EPG , L , C = -1.0, M = ALPHA1)
      DIAGMULT(A = ALPHA1, D = ALPHA1, L, C = -1.0, M = ONEA)
LINEAR(A = ONEA, B = ALPHA1, CB = -1.0, M = ALPHA)
PRINT(MATRIX = ALPHA, TEXT = ALPHA)
*
* CALC BETA
      DIAGMULT(A = SUMA2, D = SUMH2, L , C = -1.0, M = BETA1)
      DIAGMULT(A = BETA1, D = BETA1, L, C = -1.0, M = ONEB)
LINEAR(A = ONEB, B = BETA1, CB = -1.0, M = BETA)
PRINT(MATRIX=BETA, TEXT = BETA)
*
*
HEADING(TEXT=KRAJ LAV)
*
```

IMAGE ANALIZA MULTIVARIJATNIH TRAJEKTORIJA*

LEO PAVIČIĆ
Fakultet za fizičku kulturu
Sveučilišta u Zagrebu

ŽIVAN KARAMAN i KONSTANTIN MOMIROVIĆ
Sveučilišni računski centar

SAŽETAK

U ovom radu predložen je model za analizu multivarijatnih trajektorija koji se osniva na Bentlerovoj (Bentler, 1969) generalizaciji image analize. Algoritam za analizu multivarijatnih trajektorija, izradjen na osnovu ovog modela, implementiran je u programskom jeziku SS.

1. UVOD

Nakon klasičnih Tuckerovih radova faktorska analiza vremenskih serija postala je metoda koja se polako, ali sigurno probija u područje analiza stohastičkih procesa. Neke modifikacije Tuckerove osnovne metode (Momirović i Karaman, 1982a; 1982b) nedavno su primijenjene u analizi fizioloških, biokemijskih i psiholoških procesa, i osobito u analizi kinezioloških transformacijskih procesa. Iskustvo je medjutim pokazalo da je u mnogim slučajevima podatke potrebno prethodno osloboditi slučajnih smetnji, kako bi se bitna svojstva trajektorija mogla odrediti na sigurniji i precizniji način.

** Prijevod rada publiciranog u Proceedings of 5th International Symposium "Computer at the University", Cavtat, 1983, pp. 491-497.*

2. MODEL

Neka je $w_j \subset W$, gdje je $W=\{w_j;\ j=1,\ldots,m\}$ skup varijabli kojima se u svakoj vremenskoj točki t, $t=0,1,\ldots,f$ opisuju antropologijske i kriterijske karakteristike na temelju kojih je moguće procijeniti efekte izazvane treningom T_k, $k=1,\ldots,g$ koji je primijenjen na neku grupu sportaša G_k, $k=1,\ldots,g$.

Neka je $U=\{u_i;\ i=1,\ldots,n\}$ uzorak iz neke grupe G_k, i neka su efekti treninga T_k na svakom entitetu u_i; $i=1,\ldots,n$ u kontrolnim točkama t, $t=0,1,\ldots,f$ na manifestnu ili latentnu varijablu w_j opisani vektorom

$$\mathrm{W}_{ij}^{\mathrm{T}} = (w_{ijo},\ldots,\ w_{ijt},\ldots,\ w_{ijf}) \qquad i=1,\ldots,n \quad j=1,\ldots,m.$$

Neka je

$$\mathrm{W}_j = (w_{it}) \qquad j=1,\ldots,m \quad i=1,\ldots,n \quad t=0,1,\ldots,f$$

matrica rezultata entiteta iz U opisanih nad w_j u vremenskim točkama t. Bazičnu strukturu matrice W_j uvijek je moguće definirati kao

$$\mathrm{W}_j * \mathrm{P}_j \Sigma_j \mathrm{Q}_j^{\mathrm{T}} \qquad j=1,\ldots,m$$

gdje je $\Sigma_j = (\sigma_p)_j$ dijagonalna matrica sa članovima $\sigma_j \neq 0$, a P_j i Q_j su ortonormalne matrice tako da je

$$\mathrm{P}_j^{\mathrm{T}}\mathrm{P}_j = \mathrm{I} \qquad j=1,\ldots,m$$

i

$$\mathrm{Q}_j^{\mathrm{T}}\mathrm{Q}_j = \mathrm{I} \qquad j=1,\ldots,m.$$

Pretpostavimo, da se matrice W_j, $j=1,\ldots,m$ mogu dekomponirati na dvije aditivne matrice C_j i U_j tako da je

$$W_j = C_j + U_j \qquad j=1,\ldots,m$$

i da su varijable iz C_j ortogonalne sa varijablama iz U_j, tj. da vrijedi

$$C_j^T U_j = U_j^T C_j = 0 \qquad j=1,\ldots,m$$

i da su varijable u matrici

$$C_j = (c_{it})_j \qquad \begin{aligned} j&=1,\ldots,m \\ i&=1,\ldots,n \\ t&=0,1,\ldots,f \end{aligned}$$

posljedica djelovanja nekih generatora koji se mogu pripisati efektima treninga u cijelom kontinuumu $\{0,f\}$, a varijable u matrici

$$U_j = (u_{it})_j \qquad \begin{aligned} j&=1,\ldots,m \\ i&=1,\ldots,n \\ t&=1,0,\ldots,f \end{aligned}$$

slučajne varijable, nezavisne od ovih generatora.

Očito, matrica relacija izmedju vektora

$$w_{jt}^T = (w_{1jt},\ldots, w_{ijt},\ldots, w_{njt}) \qquad \begin{aligned} j&=1,\ldots,m \\ t&=1,0,\ldots,f \end{aligned}$$

kojima su opisana stanja, definirana varijablom $w_j \in W$ u točkama t, $t=0,1,\ldots,f$ je, pod ovim modelom

$$W_j^T W_j = C_j^T C_j + U_j^T U_j \qquad j=1,\ldots,m.$$

Ali, kako je

$$P_j P_j^T W_j = P_j \Sigma_j Q_j^T = W_j \qquad j=1,\dots,m$$

to je

$$W_j^T W_j = C_j^T P_j P_j^T + C_j^T P_j P_j^T U_j + U_j^T P_j P_j^T C_j + U_j^T P_j P_j^T U_j \qquad j=1,\dots,m.$$

Uočimo da vrijedi

$$C_j^T W_j = C_j^T (C_j + U_j) = C_j^T C \qquad j=1,\dots,m$$

i da su, prema tome, relacije izmedju matrice koja sadrži efekte generatora koji su zajednički, za vrijablu w_j, za cijeli proces u intervalu $\{0,f\}$, i matrice registracije stanja u toj varijabli u toku procesa u intervalu $\{0,f\}$ na uzorku U upravo identični relacijama izmedju varijabli kojima je opisan efekat zajedničkih generatora promjena.

Definirajmo generalizirani inverz matrice W_j kao

$$W_j^- = Q_j \Sigma_j^{-1} P_j^T \qquad j=1,\dots,m$$

i uočimo da vrijedi

$$C_j^T W_j W_j^- = C_j^T P_j P_j^T = C_j^T C_j W_j^- \qquad j=1,\dots,m$$

pa je prema tome teorijske krivulje vježbe, definirane vrijednostima u matrici C_j, moguće odrediti ako je poznata matrica $C_j^T C_j$. Množenjem matrice $C_j^T P_j P_j^T$ sa njenom transpozicijom

$$C_j^T P_j P_j^T C_j = C_j^T C_j (W_j^T W_j)^- C_j^T C_j \qquad j=1,\dots,m.$$

Uočimo, da, čak i ako su varijable iz U_j ortogonalne, tj. $U_j^T U_j$ = diag, matrica $U_j^T P_j P_j U_j$ nije dijagonalna matrica, jer, kako je

$$P_j = W_j Q_j \Sigma_j^{-1}$$

to je

$$U_j^T P_j P_j^T U_j = U_j^T U_j (W_j^T W_j)^{-1} U_j^T U_j \qquad j=1,\ldots,m$$

što nije dijagonalna matrica i ako je $U_j^T U_j$ dijagonalna matrica, tj. i ako su varijable, specifične za pojedine točke $t, t=0,1,\ldots,f$ u kojima je analiziran trening i njegovi efekti opisani varijablom $w_j \in W$, nezavisne.

Pretpostavimo, sada, da je

$$D_j^2 = U_j^T U_j \qquad j=1,\ldots,m$$

zaista dijagonalna matrica. Tada je

$$W_j^T W_j = C_j^T C_j + D_j^2 \qquad j=1,\ldots,m$$

i

$$C_j^T C_j = W_j^T W_j - D_j^2 \qquad j=1,\ldots,m$$

pa je

$$C_j^T P_j P_j^T C_j = (W_j^T W_j - D_j^2)(W_j^T W_j)^{-1}(W_j^T W_j - D_j^2) \qquad j=1,\ldots,m$$

ako pretpostavimo da $W_j^T W_j$ ima regularan inverz.

Ako pri tome definiramo

$$D_j^{-2} = \mathrm{diag}\,(W_j^T W_j)^{-1} \qquad j=1,\ldots,m$$

problem spektralne analize krivulja vježbe može se riješiti u okviru generaliziranog Bentlerovog (Bentler, 1969) image modela.

3. ALGORITAM

Neka je

$$W_j = (w_{it})_j \qquad j=1,\ldots,m \quad i=1,\ldots,n \quad t=0,1,\ldots,f$$

matrica podataka dobijena opisivanjem nekog skupa entiteta $U=\{u_i;\ i=1,\ldots,n\}$ na varijabli w_j, $j=1,\ldots,m$ u vremenskim točkama $t, t=0,1,\ldots,f$.

Definirajmo

$$H_j = W_j^T W_j \frac{1}{n} \qquad j=1,\ldots,m$$

kao matricu necentriranih kovarijanci varijabli iz W_j i pretpostavimo da je H_j, $j=1,\ldots,m$ pozitivno definitna matrica.

Definirajmo

$$D_j^{-2} = \mathrm{diag}\ H_j \qquad j=1,\ldots,m$$

i procijenimo slučajne komponente stanja skupa U u točkama $t, t=0,1,\ldots,f$ opisanog varijablom w_j iz $W=\{w_j;\ j=1,\ldots,m\}$ operacijom

$$U_j = W_j H_j^{-1} D_j^2 \qquad j=1,\ldots,m.$$

Necentrirane kovarijance varijabli iz U_j bit će elementi matrice

$$E_j = U_j^T U_j \frac{1}{n} = D_j^2 H_j^{-1} D_j^2 \qquad j=1,\ldots,m$$

pa će, pod ovim modelom, ocjena varijabli koje ovise od zajedničkih komponenata procesa opisanog varijablom w_j biti

$$C_j = W_j - U_j = W_j(I - H_j^{-1}D_j^2) \qquad j=1,\dots,m$$

sa matricom necentriranih kovarijanci

$$G_j = C_j^T C_j \frac{1}{m} = H_j + D_j^2 H_j^{-1} D_j^2 - 2\,D_j^2 \qquad j=1,\dots,m.$$

Neka su, u dijagonalnoj matrici $\Delta_j = (\delta_p)_j$, $p=1,\dots,f+1$ svojstvene vrijednosti matrice G_j, i neka su u matrici $X_j = (x_{tp})_j$, $t=0,1,\dots,f$; $p=1,\dots,f+1$ njima pridruženi svojstveni vektori, skalirani tako da je $X_j^T X_j = I$, $j=1,\dots,m$.

Projekcije vektora entiteta na ortogonalne komponente matrice C_j bit će

$$V_j = C_j X_j \qquad j=1,\dots,m$$

pa je, na temelju vektora

$$v_{ij}^T = c_{ij}^T X_j = (v_{ip})_j \qquad \begin{array}{l} j=1,\dots,m \\ i=1,\dots,n \\ p=1,\dots,f+1 \end{array}$$

moguće odrediti poziciju svakog entiteta u_i iz U na ortogonalnim komponentama varijabli iz C_j.

Relativni značaj vrijednosti varijable w_j u vremenskim točkama $t, t=0,1,\dots,f$ za odredjivanje ovih komponenata proporcionalan je vrijednostima faktorske matrice

$$F_j = X_j \Delta_j^{1/2} \qquad j=1,\dots,m$$

jer je, očito,

$$F_j F_j^T = G_j \qquad j=1,\dots,m.$$

Mjera adekvatnosti ovog modela za analizu spektra promjena skupa U opisanog nad w_j u vremenskom intervalu $\{0,f\}$ je

$$\alpha_j = \frac{\sum_{t=0}^{f} \sum_{t'=0}^{f} e^2_{tt'}}{\sum_{t=0}^{f} \sum_{t'=0}^{f} g^2_{tt'} + \sum_{t=0}^{f} \sum_{t'=0}^{f} e^2_{tt'}} \qquad j=1,\dots,m$$

ali i mjera

$$\beta_j = 1 - \frac{\sum_{t=0}^{f} \sum_{t'=0}^{f} a^2_{tt'}}{\sum_{t=0}^{f} \sum_{t'=0}^{f} h^2_{tt'}} \qquad j=1,\dots,m$$

gdje su $a_{tt'}$ članovi matrice $A_j = D_j H_j^{-1} D_j$, $j=1,\dots,m$ je u tu svrhu jednako pogodna.

4. TESTOVI ZNAČAJNOSTI

Distribucija mjera α_j i β_j nije, za sada, poznata pa će, za testiranje hipoteza $\alpha_j=\alpha_j^*$ i $\beta_j=\beta_j^*$, $j=1,\ldots,m$, gdje su α_j^* i β_j^* neke specificirane vrijednosti tih mjera biti potrebno, postupkom Mostellera i Tuckeya, odrediti empirijske funkcije raspodjele.

Sve komponente iz V_j nisu, naravno, od jednake važnosti; važnost je tih komponenata proporcionalna svojstvenim vrijednostima δ_{pj} iz Δ_j. Pogodan heuristički postupak za redukciju broja ovih komponenata je

$$c_j = \text{num}\ (\delta_{pj} \geq \sum_{p=1}^{f+1} \delta_{pj} \frac{1}{f+1}) \qquad j=1,\ldots,m$$

pa se analiza može usmjeriti samo na one komponente V_{pj}, $p=1,\ldots,c_j$ sa iznad prosječnom mjerom važnosti. Gubitak informacija kod ove redukcije je, naravno,

$$z_j = \sum_{p=1}^{c_j} \delta_{pj} \Big/ \sum_{p=1}^{f+1} \delta_{pj} \qquad j=1,\ldots,m.$$

Ni distribucija koeficijenata f_{tpj}, $t=0,1,\ldots,f$; $p=1,\ldots,c_j,c_j+1,\ldots,f+1$ u faktorskoj matrici F_j nije, nažalost, poznata, pa će i ovdje biti potrebno da se, postupkom Mostellera i Tuckeya, odrede empirijske distribucije ovih koeficijenata kako bi se mogle testirati hipoteze tipa $f_{tpj}=f^*$ i konstruirati intervali pouzdanosti ovih koeficijenata.

5. PROGRAM

Algoritam za image analizu multivarijatnih trajektorija implemetiran je u SS jeziku (Štalec, Momirović i Zakrajšek, 1983; Momirović, Štalec i Zakrajšek, 1982). Program omogućava analizu uzoraka do 10.000 entiteta opisanih nekom kvantitativnom varijablom u do 250 vremenskih točaka. Postoje dvije verzije ovog programa. Program TALAMBAS II je eksperimentalni program primijenjen u fazi ispitivanja ponašanja algoritma. Program LAV je konačna verzija implementacije predloženog algoritma.

LITERATURA

1. Bentler, P.M.:
   Some extensions of image analysis. Psychometrika (1969), 34, 1, 77-83.

2. Momirović, K., J. Štalec and E. Zakrajšek:
   A programming language for multivariate data analysis. COMPSTAT 82. Proceedings in computational statistics, Physica Verlag, Wien, 1982 (pp. 87-95).

3. Momirović, K. and Ž. Karaman:
   INDIFF - Model, algoritam i program za analizu promjena stanja nekog objekta opisanog nad skupom kvantitativnih varijabli. Kineziologija (1982), 13, 1-2, 5-8.

4. Momirović, K. i Ž. Karaman:
   COLDIFF - Algoritam i program za analizu promjena spektralnom dekompozicijom multivarijatnih trajektorija. Kineziologija (1982), 13, 1-2, 9-11.

5. Štalec, J., K. Momirović i E. Zakrajšek:
   Statistički sistem. Fakultet za fizičku kulturu Sveučilišta u Zagrebu, Zagreb, 1983.

2.11 ALGORITAM I PROGRAM SCHOENEMANN

Kose Prokrustovske transformacije sa ortogonalnim rotacijama spadaju u klasu konfirmativnih metoda, pa će zbog toga biti korištene kao objektivan test nekog modela u sportu.

Koristit će se, prvenstveno, za provjeru latentne strukture subuzorka neke populacije sportaša i za usporedbu sa nekom hipotetskom strukturom, definiranom na temelju neke teorije.

Osim toga, koristiti će se i u svim onim slučajevima kad želimo komparirati latentnu strukturu psihosomatskih potprostora subuzoraka (u ma kojoj referentnoj ili finalnoj točki) sa nekom populacijom sportaša, kako bi mogli optimalizirati proces treninga.

Predloženi algoritam i njemu pridruženi program SCHOENEMANN namijenjen je rješavanju ove klase problema.

```
*
*     NAZIV PROGRAMA
*
*
*              *** S C H O E N E M A N N ***
*
*
*
*     AUTORI
*
*              N. VISKIC-STALEC
*              J. STALEC
*              K. MOMIROVIC
*
*
*     FUNKCIJA
*
*
*       SCHOENEMANN JE IMPLEMENTACIJA UZ NEZNATNE MODIFIKACIJE METODE
*       ZA ORTOGONALNU PROCRUST TRANSFORMACIJU SCHOENEMANNA (1966).
*       TA SE MODIFIKACIJA SASTOJI U SLIJEDECEM:
*       (1)  SCHOENEMANNOVA PROCEDURA SE PRIMJENJUJE NA VLASTITE VEKTORE
*            REDUCIRANE MATRICE KORELACIJA, A NE NA FAKTORE TE MATRICE
*       (2)  KOMUNALITETI SE ODREDJUJU NA TEMELJU BROJA FAKTORA I
*            HIPOTETSKIH KOMUNALITETA DOBIJENIH IZ CILJNE MATRICE
*       (3)  PREDLOZENA PROCEDURA DAJE KOSE FAKTORE U PROSTORU VEKTORA
*            ENTITETA I ORTOGONALNE FAKTORE U PROSTORU VARIJABLI. U TOM
*            SLUCAJU SOLUCIJA JE SLICNA HARRIS-KAISER NEZAVISNOJ KLASTER
*            ORTHOBLIQUE SOLUCIJI (HARRIS I KAISER, 1964).
*       (4)  PREDLOZENA METODA PRIPADA FAKTORSKOM MODELU, NE KOMPONENTNOM
*            MODELU FAKTORSKE ANALIZE TAKO DA SU ZA ODREDJIVANJE
*            FAKTORSKIH SKOROVA PRIMJENJENE METODA NAJMANJIH KVADRATA
*            I REGRESIJA.
*
*
*     METODA I ALGORITAM OPISANI SU U
*
*            VISKIC-STALEC, N., J. STALEC I K. MOMIROVIC:
*            OBLIQUE PROCRUSTES TRANSFORMATION BY ORTHOGONAL ROTATION,
*            NEPUBLICIRANI RAD
*
*
*     PRIPREMA PODATAKA
*
*            (1)  MATRICA PODATAKA, PRETHODNO OPISANA S SEQUENCE
*                 NAREDBOM I ZAHTJEVANIM BROJEM VARIABLE NAREDBAMA,
*                 MORA BITI PRIPREMLJENA U SISTEMSKOJ INPUT DATOTECI
*            (2)  MATRICA CILJA MORA BITI PRIPREMLJENA U DRUGOJ SISTEMSKOJ
*                 DATOTECI. IMENA REDAKA CILJNE MATRICE MORAJU BITI
*                 IDENTICNA IMENIMA VARIJABLI MATRICE PODATAKA I UREDJENA
*                 U ISTOM REDOSLJEDU KAO I IMENA VARIJABLI U MATRICI
*                 PODATAKA. IMENA HIPOTETSKIH FAKTORA SU ZADANA S
*                 VARIABLE NAREDBAMA.
*            (3)  CILJNA MATRICA MORA BITI DEFINIRANA KAO ORTOGONALNA
*                 MATRICA STRUKTURE. TO ZNACI DA SU ELEMENTI CILJNE
*                 MATRICE HIPOTETSKE KORELACIJE VARIJABLI S ORTOGONALNIM
*                 LATENTNIM DIMENZIJAMA. ORTOGONALNOST CILJNE MATRICE
*                 PO KOLONAMA NIJE NUZNA.
*
*            ***  UPOZORENJE  ***
*
*            AKO BROJ HIPOTETSKIH FAKTORA PREMASUJE BROJ ZNACAJNIH
*            GLAVNIH KOMPONENATA PREMA KRITERIJU GUTTMAN-KAISER, REZULTATI
*            MOGU BITI BESMISLENI. U SVAKOM SLUCAJU, BROJ HIPOTETSKIH
```

```
*             FAKTORA MORA BITI MANJI OD POLOVINE BROJA VARIJABLI.
*
*
*     IZVRSNI DIO PROGRAMA
*
*     BLOK 0.
*     UCITAVANJE I USKLADJIVANJE PODATAKA
*
OUTPUT(DEVICE=PR)
HEADING(TEXT= S C H O E N E M A N N , T)
HEADING(TEXT=KOSE PROKRUSTOVE TRANSFORMACIJE S ORTHOGONALNIM ROTACIJAMA)
HEADING(TEXT=MODIFIKACIJA SCHOENEMANNOVE METODE - 1966.D)
TEXT(TEXT=KOSE     PROKRUSOVE ROTACIJE )
*
*     BLOK 1.
*     PROCJENA KOMUNALITETA I OSNOVNO RJESENJE PO NAJMANJIM KVADRATIMA
*
HEADING(TEXT=POCETNO RJESENJE.D)
INPUT.
INPUT(SCORE=T)
CORRELATION(SCORE=SCORE)
TRANSPOSE(OLD=T.NEW=TT)
ITERCOMMUNALITIES(F=TT.FN=H.LAMBDA=L.ITER=99)
DELETE(MATRIX=SCORE)
DELETE(MATRIX=L)
RESIDUAL(F=H)
PRINT(MATRIX=R.TEXT=MATRICA INTERKORELACIJA)
PRINT(MATRIX=H.T.TEXT=GLAVNE KOMPONENTE REDUCIRANE MATRICE KORELACIJA)

*
*     BLOK 2.
*     RACUNANJE SCHOENEMANNOVE MATRICE TRAMSFORMACIJE
*
HEADING(TEXT=MATRICA TRANSFORMACIJE.D)
MULT(A=TT.B=T.M=LT)
DIAGMULT(A=T.D=LT.C=-0.5.R.M=Y)
MULT(A=H.B=H.TB.M=L)
DIAGMULT(A=H.T.D=L.C=-0.5.R.M=X)
MULT(A=X.TA.B=Y.M=G)
MULT(A=G.TA.B=G.M=GG)
DIAGONALISATION(R=GG.LAMBDA=LMB.X=V)
MULT(A=G.B=G.TB.M=GGT)
DIAGONALISATION(R=GGT.LAMBDA=LLMB.X=W)
MULT(A=W.B=V.TA.M=Q)
DELETE(MATRIX=LT)
DELETE(MATRIX=G)
DELETE(MATRIX=GG)
DELETE(MATRIX=LMB)
DELETE(MATRIX=V)
DELETE(MATRIX=GGT)
DELETE(MATRIX=LLMB)
DELETE(MATRIX=W)
PRINT(MATRIX=Q.TEXT=SCHOENEMANNOVA TRANSFORMACIJASKA MATRICA)
*
*     BLOK 3.
*     KOSO PROKRUSTOVO RJESENJE
*
HEADING(TEXT=KOSO PROKRUSTOVO RJESENJE.D)
MULT(A=X.B=Q.M=XQ)
DIAGMULT(A=Q.T.D=L.R.M=QTL)
MULT(A=QTL.B=Q.M=C)
SCALE(C=C.R=M)
DIAGMULT(A=XQ.D=C.C=0.5.R.M=A)
MULT(A=A.B=M.M=F)
```

```
PRINT(MATRIX=A,TEXT=PROCRUSTOVA MATRICA SKLOPA)
PRINT(MATRIX=M,TEXT=MATRICA KORELACIJA PROKRUSTOVIH FAKTORA)
PRINT(MATRIX=F,TEXT=PROCRUSTOVA MATRICA STRUKTURE)
DELETE(MATRIX=QTL)
*
*    BLOK 4.
*    PROCJENA FAKTORSKIH SKOROVA
*
HEADING(TEXT=PROCJENA FAKTORSKIH SKOROVA:D)
DIAGMULT(A=XQ,D=C,C=-0.5,R,M=BETA)
MULT(A=R,B=BETA,M=RBETA)
MULT(A=BETA,TA,B=RBETA,M=M1)
DIAG(A=M,D=I)
DIAG(A=M1,D=SMC1)
LINEAR(A=I,B=SMC1,CB=-1.0,M=E1)
DIAG(A=E1,C=0.5,D=ERROR1)
PRINT(MATRIX=BETA,TEXT=TEORETSKI REGRESIJSKI PONDERI PO NAJMANJIM KVADRATIMA)
PRINT(MATRIX=M1,TEXT=TEORETSKE KORELACIJE FAKTORA PO NAJMANJIM KVADRATIMA)
PRINT(MATRIX=ERROR1,TEXT=TEORETSKE GREŠKE PROCJENE FAKTORSKIH SKOROVA)
DELETE(MATRIX=RBETA)
DELETE(MATRIX=SMC1)
DELETE(MATRIX=E1)
INVERSION,
MULT(A=RINV,B=F,M=GAMA)
MULT(A=F,TA,B=GAMA,M=M2)
DIAG(A=M2,D=SMC2)
LINEAR(A=I,B=SMC2,CB=-1.0,M=E2)
DIAG(A=E2,C=0.5,D=ERROR2)
PRINT(MATRIX=GAMA,TEXT=STANDARDNI REGRESIJSKI KOEFICIENTI)
PRINT(MATRIX=M2,TEXT=KORELACIJE STANDARDNIH FAKTORA)
PRINT(MATRIX=ERROR2,TEXT=GRESKE PROCJENE STANDARDNIH FAKTORA)
*
HEADING(TEXT=KRAJ PROGRAMA SCHOENEMANN)
```

KOSE PROKRUSTOVE SOLUCIJE POMOĆU ORTOGONALNIH TRANSFORMACIJA*

Viskić-Štalec, N., J. Štalec i K. Momirović
Fakultet za fizičku kulturu
Sveučilišta u Zagrebu

SAŽETAK

Predložena je takva modifikacija Schönemannove (1966) metode koja daje kosu Prokrust soluciju s ortogonalnim transformacijama. Modifikacija se sastoji od primjene Schönemannove procedure na vlastite vektore reducirane korelacijske matrice s komunalitetima koji su procijenjeni iterativnim procesom na osnovu broja faktora i početnih komunaliteta izvedenih iz ciljne matrice sklopa postupkom sličnim Guttmanovoj faktorizaciji kvadratne simetrične matrice. Predložen postupak daje kose faktore u prostoru vektora ispitanika i ortogonalne faktore u prostoru vektora varijabli.

---

* *Ovaj rad izvorno je napisan na engleskom jeziku i bit će prezentiran na Medjunarodnom simpoziju računarske statistike COMPSTAT '84, Prag, 1984.*

1. UVOD

1962. godine Hurley i Cattell prvi put su upotrijebili naziv Prokrustove transformacije u faktorskoj analizi, kojim su označili komparaciju empirijske faktorske matrice i hipotetske ili ciljne matrice, a koja služi za provjeru znanstvenih modela.

Utvrdjivanje da li empirijski podaci potvrdjuju predvidjanja dana na osnovu neke teorije, svodi se na transformaciju empirijske faktorske matrice, da bi se postigla što bolja aproksimacija ciljne matrice, pri čemu konačno rješenje može biti ortogonalno ili kosokutno.

Schönemann je 1966. godine predložio ortogonalnu soluciju Prokrustovog problema pri kojem empirijska i ciljna matrica ne moraju imati isti rang (za razliku od rješenja kojeg je predložio Green 1952., prema kojem matrice moraju imati kompletan rang), pa je tako moguća singularna transformacija neke inicijalne solucije prema ciljnoj matrici manjeg kolonskog ranga.

Na osnovu rada Cvitaševe i Momirovića (1984) predložena je modifikacija Schönemannove metode i izradjen algoritam i program. Sastoji se u tome da se Schönemannov postupak primjenjuje na vlastite vektore reducirane korelacijske matrice. Komunaliteti se procjenjuju iterativno, na osnovu broja faktora i početnih komunaliteta ciljne matrice, tako da predložena metoda pripada faktorskom, a ne komponentnom modelu faktorske analize. U prostoru vektora ispitanika metoda daje kosokutne faktore, a u prostoru vektora varijabli ortogonalne. Faktorski skorovi odredjuju se na temelju kriterija najmanjih kvadrata i regresijskim postupkom.

2. METODE I ALGORITAM

Neka je $E=\{e_i;\ i=1,\ldots,n\}$ slučajni uzorak iz populacije P i $V=\{v_j;\ j=1,\ldots,m\}$ grupni uzorak iz univerzuma varijabli U, definiran u skladu s nekom odredjenom taksonomijom varijabli iz U.

Neka je $Z=(z_{ij});\ i=1,\ldots,n,\ j=1,\ldots,m$ matrica podataka dobijena deskripcijom E na V, i standardizirana tako da je $Z^T 1=0$ i diag $(Z^T Z)=I\,n$, sa $1(n \times 1)$ vektorom jedinica, i I matricom identiteta reda m.

Neka je $T=(t_{jk});\ j=1,\ldots,m;\ k=1,\ldots,q$ hipotetska matrica sklopa, koja na neki način odražava taksonomiju varijabli iz U na varijablama u V; uz pretpostavku faktorskog modela

$$Z = \Phi T^T + E \qquad \left|\ \begin{array}{l} \Phi^T\Phi\,\frac{1}{n} = M \\ \text{diag}\ M = I \\ \Phi^T E = 0 \\ E^T E\,\frac{1}{n} = U^2 = \text{diag} \end{array}\right.$$

sa $\Phi=(\phi_{ik});\ i=1,\ldots,n;\ k=1,\ldots,q$ nepoznatom matricom faktorskih skorova i $E=(e_{ij});\ i=1,\ldots,n;\ j=1,\ldots,m$ nepoznatom matricom uniknih skorova prihvatljivom za $E \otimes V$*.

Tada je

$$R = Z^T Z\,\frac{1}{n} = T M T^T + U^2$$

i procjena Φ metodom najmanjih kvadrata (Holzinger i Harman, 1941; Horst, 1965; Harris, 1967; McDonald i Bun, 1967) je

$$\Phi^* = Z T (T^T T)^{-1}$$

---

* *To je u stvari model II McDonalda i Buna (1967)*

s matricom kovarijanci

$$M^* = \Phi^{*T}\Phi^* \frac{1}{n} = (T^TT)^{-1}T^T R\, T(T^TT)^{-1}$$

(Harris, 1967).

Uvodjenjem ograničenja diag $M = I$, i označavanjem sa V^{*2}=diag M^* dijagonalne matrice varijanci procijenjenih faktorskih skorova,

$$\Phi = \Phi^* V^{*-1}$$

i

$$M = \Phi^T\Phi \frac{1}{n} = V^{*-1}M^*V^{*-1}.$$

Neka je

$$M^{-1/2} = Y\Delta^{-1/2}Y^T$$

sa Δ dijagonalnom matricom karakterističnih vrijednosti matrice M, i Y, $Y^TY=I$, $YY^T = I$, matrica pripadajućih vlastitih vektora.

Tada je ortogonalizacija Φ na osnovu kriterija najmanjih kvadrata*

$$\tilde{\Phi} = \Phi M^{-1/2}$$

i

$$S = Z^T\tilde{\Phi} \frac{1}{n} = R\, T(T^TT)^{-1}V^{*-} M^{-1/2}$$

je ortogonalna matrica strukture izvedena iz T tako da je

---

* *Green, 1952; ova procedura je u suštini transformacija Φ u Mahalanobisov oblik (vidi, npr., Rudan, Szirovicza i Momirović, 1979).*

odredjivanje komunaliteta pod hipotezom zadanom sa T*

$$H_o^2 = \mathrm{diag}(SS^T).$$

Sada je iterativna procedura definirana slijedećim operacijama

$$R_a = R - I + H_a^2 \qquad a=0,1,\ldots$$

$$R_a = \sum_{k=1}^{q} \lambda_{ak} x_{ak} x_{ak}^T + E \qquad a=0,1,\ldots \quad k=1,\ldots,q$$

$$H_{a+1}^2 = \mathrm{diag}\left(\sum^{q} \lambda_{ak} x_{ak} x_{ak}^T\right) \qquad a=0,1,\ldots$$

sa λ_{ak} prvih q vlastitih vrijednosti R_a, i x_{ak} pripadajućih vlastitih vektora, poslije konvergencije definirane sa uvjetom $|H_a^2 - H_{a+1}^2| \leq \varepsilon I$. Kad je uvjet konvergencije zadovoljen, R_a je definirana dijagonalnom matricom vlastitih vrijednosti $\Lambda=(\lambda_k)$ i matricom pripadajućih vlastitih vektora $X=(x_k)$, tako da je

$$R = X \Lambda X^T + U^2$$

i

$$U^2 = I - H^2$$

sa $H^2=H_a^2$ dobijen u zadnjoj iteraciji.

Definirajmo sada reparametriziranu ciljnu matricu sa

$$G = T(\mathrm{diag}(T^T T))^{-1/2}$$

---

* *Naravno, H_o^2 može biti odredjena jednostavnije sa diag(TMT^T); ali to nije pogodnost u smislu SS jezika, jer, u naredbi ITERCOMMUNALITIES, SS prihvaća inicijalnu ortogonalnu matricu strukture.*

i riješimo problem

$$XQ = G + N \quad \left| \begin{array}{l} tr(N^T N) = min \\ Q^T Q = Q Q^T = I. \end{array} \right.$$

Kako je pokazao Schönemann* (Schönemann, 1966; Mulaik, 1972; Fulgosi, 1979) i za sličan primjer Cvitaš i Momirović (1984),

$$Q = L D^T$$

sa L matricom lijevih vlastitih vektora od $X^T G$, i D matricom desnih vlastitih vektora od $X^T G^*$.

Bazični faktorski model može sada biti napisan u obliku

$$Z = \Psi Q^T X^T + E$$

tako da su latentne dimenzije procijenjene pod modelom najmanjih kvadrata

$$\Psi^* = (Z - E) X Q$$

sa matricom kovarijanci

$$C = \Psi^{*T} \Psi^* = Q^T X^T (R - U^2) X Q = Q^T \Lambda Q.$$

Definirajmo $W^2 = \mathrm{diag} C$ kao dijagonalnu matricu varijanci latentnih dimenzija. Standardizirane latentne dimenzije, suglasno s uvjetom $\mathrm{diag}\, M = I$, su

$$\Psi = (Z - E) X Q W^{-1}$$

---

* *Originalni izvod Schoenemanna je, kako su pokazali Cvitaš i Momirović, bespotrebno zakompliciran; u stvari, taj problem može biti riješen spektralnom dekompozicijom.*

s korelacijskom matricom

$$M = \Psi^T\Psi \frac{1}{n} = W^{-1}Q^T\Lambda QW^{-1}.$$

Matrica strukture, definirana korelacijama zajedničkog dijela standardiziranih podataka i standardiziranih latentnih dimenzija, je

$$F = (Z - E)^T(Z - E)XQW^{-1} \frac{1}{n} = (R - U^2)XQW^{-1} = X\Lambda QW^{-1}$$

tako da je matrica sklopa

$$A = F M^{-1} = X Q W.$$

Očito, ta metoda daje kosa rješenja u prostoru vektora entiteta, dobijena s ortogonalnim rotacijama (reparametrizirane) ciljne matrice. To rješenje je analogno nezavisnom cluster modelu orthoblique transformacija (Harris i Kaiser, 1964); jer su

$$A^TA = W^2,$$

koordinatne osi ortogonalne u prostoru vektora varijabli.

Naravno, matrica E je nepoznata, i faktorski skorovi moraju biti procijenjeni nekom metodom iz standardiziranih podataka u matrici Z. Uzmimo u obzir dvije standardne procedure, metodu najmanjih kvadrata (ili "idealne varijable") predloženu od Holzingera i Harmana (1941) i Horsta (1965), definiranu estimacijskom matricom

$$\Gamma = A(A^TA)^{-1} = X Q W^{-1}$$

i regresijsku metodu (Thurstone, 1935; Thomson, 1936), s estimacijskom matricom

$$\beta = R^{-1}F = R^{-1}X \Lambda Q W^{-1}.$$

Skorovi, odredjeni sa

$$\Phi_1 = Z\Gamma$$

imaju matricu kovarijanci

$$C_1 = \Gamma^T R \Gamma = W^{-1}Q^T X^T RXQW^{-1}$$

i skorovi, odredjeni sa

$$\Phi_2 = Z\beta$$

imaju matricu kovarijanci

$$C_2 = \beta^T R\beta = F^T R^{-1} F.$$

Vidimo da je matrica kroskovarijanci izmedju Φ_1 i Φ_2

$$\Phi_1^T \Phi_2 \frac{1}{n} = \Gamma^T R\,\beta = W^{-1}Q^T \Lambda\, Q\, W^{-1} = M,$$

teoretska korelacijska matrica latentnih dimenzija.

Procjena metodom najmanjih kvadrata daje matricu kroskovarijanci izmedju teoretskih latentnih dimenzija Ψ i Φ_1

$$\Psi^T \Phi_1 \frac{1}{n} = W^{-1}Q^T X^T (Z - E)^T ZXQW^{-1} \frac{1}{n} = W^{-1}Q^T X^T (R - U^2)XQW^{-1} =$$

$$= W^{-1}Q^T \Lambda QW^{-1} = M,$$

ponovo matricu interkorelacija teoretskih latentnih dimenzija. Kroskovarijance medju teoretskim latentnim dimenzijama u Ψ i latentnim dimenzijama u Φ_2 procijenjenim regresijskom metodom su

$$\Psi^T \Phi_2 \frac{1}{n} = W^{-1}Q^T X^T (Z - E)^T ZR^{-1}F = \Gamma^T (I - U^2 R^{-1})F = F^T R^{-1} F = C_2,$$

jednake matrici kovarijanci latentnih dimenzija procijenjenih regresijskom metodom; obratimo pažnju na image oblik $F^T(I - R^{-1}U^2)\Gamma$, i uočimo da je ova forma analogna McDonaldovoj formalizaciji relacija faktorskog i image modela (McDonald, 1975).

3. PROGRAM

Predložena metoda je gotovo doslovno implementirana u kompjuterskom programu, napisanom u SS jeziku (Štalec, Momirović i Zakrajšek, 1983). Program SCHOENEMANN se sastoji od 189 linija koda, uključujući komentare koji sadrže upute o svojstvima i korištenju programa. U ovoj verziji SS (verzija 5,2/M), SCHOENEMANN dozvoljava do 10.000 entiteta, do 250 varijabli i do 110 hipotetskih faktora. U standardnoj verziji korisnik SCHOENEMANN-a priprema matricu podataka i bilo koji tip ciljne matrice; s malim izmjenama u pojedinim linijama koda, program prihvaća kao ulaz matricu korelacija. Imena varijabli moraju biti, naravno, ista u matrici podataka (ili korelacija) i ciljnoj matrici, ali nije neophodno da budu i u istom redoslijedu, tako da se mogu testirati različite hipoteze na istoj matrici podataka ili matrici korelacija, uključujući i hipotezu bez nekih varijabli. Slobodne vrijednosti u ciljnoj matrici nisu dozvoljene, ali vrijednosti mogu biti specificirane na bilo koji način, uključujući jednostavne binarne selektorske matrice.

SCHOENEMANN je pohranjen u zatvorenoj biblioteci FFK*LIB i javnoj biblioteci SRCE*SS-MAKRO sa malim numeričkim primjerom (15 morfoloških mjera i 3 faktora ciljne matrice).

LITERATURA

1. Cvitaš, M. and K. Momirović:
Note on some properties of oblique Procrustes transformations by orthogonal rotations: Unpublished paper, 1984.

2. Fulgosi, A.:
Faktorska analiza. Školska knjiga, Zagreb, 1979.

3. Green, B.T.:
The orthogonal approximation of an oblique structure in factor analysis. Psychometrika, 17 (1952), 429-440.

4. Harris, C.W. and H.F. Kaiser:
Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 29 (1964), 347-362.

5. Harris, C.W.:
On factors and factor scores. Psychometrika, 32 (1967), 363-379.

6. Holzinger, K.J. and H.H. Harman:
Factor analysis - a synthesis of factorial methods. University of Chicago Press, Chicago, 1941.

7. Horst, P.:
Factor analysis of data matrices. Holt, Rinehart and Winston, New York, 1965.

8. McDonald, R.P. and E.J.Bun:
A comparison of four methods of constructing factor scores. Psychometrika, 32 (1967), 381-401.

9. McDonald, R.P.:
Descriptive axioms for common factor theory, image theory and component theory. Psychometrika, 40 (1975), 137-152.

10. Mulaik, S.A.:
The fondations of factor analysis. McGraw-Hill, New York, 1972.

11. Rudan, P., L. Szirovicza and K. Momirović:
The application of an algorithm based on Mahalanobis angles and iterative Q-method of taxonomic analysis in the study of micro-evolution. Periodicum Biologorum, 81 (1979), 583-589.

12. Thomson, G.H.:
Some points of mathematical technique in the factorial analysis of ability. Journal of Edducational Psychology, 27 (1936), 37-54.

13. Thurstone, L.L.:
The vectors of mind. University of Chicago Press, Chicago, 1935.

14. Schönemann, P.H.:
A generalized solution of orthogonal Procrustes problem, Psychometrika, 31 (1966), 1-10.

15. Štalec, J., K. Momirović i E. Zakrajšek:
Programski sistem SS. Fakultet za fizičku kulturu Sveučilišta u Zagrebu, Zagreb, 1983.

2.12 ALGORITAM I PROGRAM QUAQUA

Nizom dosadašnjih istraživanja potvrdjeno je postojanje kvantitativnih i kvalitativnih razlika u psihofizičkim karakteristikama sportaša pripadnika različitih sportskih grana koje se ne mogu zadovoljavajuće aproksimirati modelima dobivenim na osnovu karakteristika opće populacije.

Precizna specifikacija razvoja sportske forme, na osnovu adekvatnih modela, neophodan je osnov za modeliranje kompleksnih procesa treninga. Stoga je mogućnost provjere sukladnosti karakteristika teoretskih modela sa podacima o objektivnoj strukturi empirijskih podataka od presudnog značaja za razvoj teorije i prakse.

U svrhu komparacije empirijskih i teoretskih struktura modela predložen je algoritam i program QUAQUA kojim se utvrdjuje struktura osobina uzorka, npr. iz populacije vrhunskih sportaša, i usporedjuje sa teoretskom strukturom tih osobina, odredjenom prema nekom teoretskom modelu.

```
NAZIV PROGRAMA


        *********************
        *** Q U A Q U A  ***
        *********************


AUTORI

        K. MOMIROVIC
        J. STALEC
        F. PROT
        K. BOSNAR


FUNKCIJA


QUAQUA OBAVLJA OVE POSLOVE I ZADATKE:
(1) ODREDJUJE FAKTORSKU STRUKTURU JEDNOG SKUPA
    KVANTITATIVNIH VARIABLI, POD FAKTORSKIM MODELOM,
    DEFINIRANU QUASIORTHOBLIQUE TRANSFORMACIJOM UZ
    ZADANI BROJ FAKTORA
(2) ODREDJUJE FAKTORSKU STRUKTURU TOG SKUPA VARIJABLI
    TAKO DA FORMIRA KOSU SOLUCIJU ORTONORMALNOM
    TRANSFORMACIJOM SVOJSTVENIH VEKTORA SCHOENEMANNOIDNIM
    POSTUPKOM USMJERENIM NA NORMIRANU MATRICU CILJA
(3) USPOREDJUJE OVE SOLUCIJE RACUNANJEM KOVARIJANCI
    LATENTNIH DIMENZIJA I ODREDJIVANJEM KONGRUENCIJE
    MATRICA SKLOPA
(4) ZA CIJELO VRIJEME BODRI, POTICE, SAVJETUJE I NA
    DRUGE NACINE GNJAVI VRLOG I VEOMA POSTOVANOG KORISNIKA.


PRIPREMA PODATAKA


DA BI QUAQUA MOGAO DA RADI UCENI KORISNIK ILI NJEMU PRIDRUZENI
NISTAVNI INFORMATICAR TREBA DA IZVRSE OVE RADNJE:
(1) DA PRIPREME MATRICU PODATAKA KOJOJ
    PREDHODI KONTROLNI ZAPIS NA NACIN
    KAKO PISE U SS PRIRUCNIKU
(2) DA PRIPREME HIPOTETSKU MATRICU SKLOPA
    U TRANSPONIRANOM OBLIKU TAKO DA KAO
    IMENA ENTITETA NAVEDU IMENA HIPOTETSKIH
    FAKTORA, A KAO IMENA VARIJABLI ISTA IMENA
    I U ISTOM REDOSLJEDU KAKO JE TO UCINJENO
    KOD MATRICE PODATAKA
(3) DA SVE TO PRIDRUZE PROGRAMU QUAQUA,
    AKTIVIRAJU PROGRAM, ODU DA POPIJU
    KAVU ILI NEKU DRUGU TEKUCINU
(4) DA POKUPE LISTINE S REZULTATIMA I VIDE
    DA LI VIDE ONO STO VIDE, ILI VIDE ONO
    STO NE VIDE
(5) DA NI PO KOJU CIJENU NE KONSULTIRAJU
    AUTORE PROGRAMA, JER SU ONI, BUDUCI
    KRATKE PAMETI, VEC ZABORAVILI STA
    QUAQUA RADI
(6) DA NE PSUJU NITI GOVORE RUZNE I
    NEPRILICNE RIJECI, JER JE QUAQUA BEZGRESAN
    PROGRAM, I SVE NEVOLJE DOLAZE OD
    PODATAKA, ILI MODELA.
```

```
*
*    UPOZORAVA SE MUDRI KORISNIK DA JE QUAQUA VRLO SKROMAN
*    I STEDLJIV PROGRAM, I DA ZBOG TOGA (A I ZATO STO ZNA
*    DA SE MNOGI LJUDI, PA STOGA I KORISNICI RACUNALJKI,
*    ZBUNJUJU OD PREVELIKOG BROJA INFORMACIJA) EMITIRA
*    SAMO MATRICE SKLOPA, KOVARIJANCI I KROSKORELACIJA
*    LATENTNIH DIMENZIJA, I KOEFICIJENATA KONGRUENCIJE,
*    A SVE OSTALO HITRO BRISE.
*
*
*    IZVRSNI DIO PROGRAMA
*
OUTPUT(DEVICE=PR)
HEADING(TEXT=Q U A Q U A,T)
HEADING(TEXT= PRORTHOBLIQUE - ORTHOBLIQUE)
HEADING(TEXT=USPOREDBA EKSPLORATIVNE I KONFIRMATIVNE ORTHOBLIQUE SOLUCIJE,D)
TEXT(TEXT=                QUAQUA)
*
*
*                                       "PA JA SAM VIDEO STO SAM VIDEO,
                                         I VIDEO SAM STO NISAM VIDEO",
*                                        RECE VRCOPIVAC TVRDOGLAVO,
*
*                                       "A JA KAZEM DA TU IMA NEKA
*                                        GRESKA", PONOVI MASLOCVOR.
*
*
*                                            TOLKIN,
*                                            GOSPODAR PRSTENOVA
*                                            (PREVOD Z. STANOJEVIC,
*                                             NOLIT, BEOGRAD, 1981
*                                             STR. 259)
*
*
*    OVAJ SU PROGRAM NAPISALI, U ZNOJU LICA SVOGA, I NE ZALECI TRUDA
*    I NAPORA, SKROMNI TRUDBENICI ODJELA ZA INFORMATIKU I STATISTIKU
*    PIXI, DIXI, KOSTA I STABI, E DA BI
*    (1) POMOGLI BLIZNJIMA DA UTVRDE DA LI VIDE ONO STO
*        VIDE, ILI VIDE ONO STO NE VIDE
*    (2) SMANJILI MUKE STROJU I SEBI KADA VRLI KORISNIK
*        UPORNO NASTOJI DA DOKAZE DA MU JE MODEL ISPRAVAN
*    (3) ISPUNILI GODISNJU OBAVEZU U DIJELU KOJI SE ODNOSI
*        NA STRUCNI RAD.
*
*    A SAD QUAQUA POCINJE DA RADI, OD SVOG POSLA KORISNIK
*    NECE VIDJETI SKORO NISTA.
*
INPUT(VARNAME=ZUM,DATA=DATA)
        DELETE(MATRIX=ZUM)
INPUT(SCORE=TT,VARNAME=ZUM,DATA=ATA)
MULT(A=TT,B=TT,TB,M=TTT)
DIAGMULT(A=TT,T,D=TTT,C=-0.5,R,M=Y)
INVERSION(R=TTT,RINV=TTTI)
CORRELATION(SCORE=SCORE)
MULT(A=TT,TA,B=TTTI,M=BEDAK)
MULT(A=R,B=BEDAK,M=GLUPAK)
MULT(A=BEDAK,TA,B=GLUPAK,M=TIKVAN)
SCALE(C=TIKVAN,R=KLIPAN)
DIAGONALISATION(R=KLIPAN)
DIAGMULT(A=X,T,D=LAMBDA,C=0.5,R,M=BIK)
MULT(A=BIK,B=X,M=KONJ)
MULT(A=KONJ,B=TT,M=SLON)
*
*    SVE JE U REDU, QUAQUA RADI, BUDITE STRPLJIVI.
```

```
*
         DELETE(MATRIX=TTT)
         DELETE(MATRIX=TTTI)
         DELETE(MATRIX=SCORE)
         DELETE(MATRIX=BEDAK)
         DELETE(MATRIX=GLUPAK)
         DELETE(MATRIX=TIKVAN)
         DELETE(MATRIX=KLIPAN)
         DELETE(MATRIX=X)
         DELETE(MATRIX=LAMBDA)
         DELETE(MATRIX=KONJ)
*
*   QUAQUA JE OBAVIO VELIKO SPREMANJE, USKORO CE STVARNO
*   POCETI DA PROIZVODI. MEDJUTIM, U BLIZINI LUTA HEYVOODOV
*   DUH, SUJEVJERNI KORISNIK NEKA KUCNE U GLAVU NAJBLIZEG
*   INFORMATICARA.
*
ITERCOMMUNALITIES(F=SLON,LAMBDA=LAMBDA)
MULT(A=FC,B=FC,TB,M=L)
DIAGMULT(A=FC,D=L,C=-0.5,L,M=XT)
         DELETE(MATRIX=LAMBDA)
         DELETE(MATRIX=SLON)
         DELETE(MATRIX=FC)
*
*   SVE PRIPREMNE RADNJE  SU ZAVRSENE. OSNOVNA
*   SOLUCIJA JE DOBIJENA, I ODMAH UNISTENA.
*   ZADRZANI SU SAMO SVOJSTVENI VEKTORI I SVOJSTVENE
*   VRIJEDNOSTI REDUCIRANE MATRICE KOREALCIJA.
*   KOMUNALITETI SU ODREDJENI ITERATIVNIM POSTUPKOM.
*
*   QUAQUA SADA FORMIRA ORTHOBLIQUE SOLUCIJU, ALI
*   ZBOG OSOBNOG ANIMOZITETA NE KORISTI ORTHOBLIQUE
*   NAREDBU.
*
VARIMAX(F=XT,TAU=Q,FN=XQ,N)
DIAGMULT(A=Q,D=L,R,M=QTL)
MULT(A=QTL,B=Q,TB,M=C)
SCALE(C=C,R=M)
DIAGMULT(A=XQ,T,D=C,C=0.5,R,M=A)
         DELETE(MATRIX=QTL)
*
*   OVO SU IZMISLILI HARRIS I KAISER, 1964
*
PRINT(MATRIX=A,TEXT= SKLOP)
PRINT(MATRIX=M,TEXT=KORELACIJE FAKTORA,D)
*
*   OVO JE BILA STANDARDNA SOLUCIJA, DOBIJENA EKSTREMIZACIJOM
*   PRISTOJNE PARSIMONIJSKE FUNKCIJE. SIGURNO JE DA JE OVA
*   SOLUCIJA TOCNA, ALI NIJE SIGURNO DA JE ISTINITA.
*
*   QUAQUA SADA FORMIRA SCHOENEMANNOIDNU
*   SOLUCIJU, GADJAJUCI U MATRICU CILJA
*   NORMIRANU PO KOLONAMA.
*
MULT(A=XT,B=Y,M=G)
MULT(A=G,TA,B=G,M=GTG)
MULT(A=G,B=G,TB,M=GGT)
DIAGONALISATION(R=GTG,LAMBDA=DELTA,X=DT)
         DELETE(MATRIX=GTG)
         DELETE(MATRIX=DELTA)
DIAGONALISATION(R=GGT,LAMBDA=DELTA,X=LT)
MULT(A=LT,TA,B=DT,M=T)
         DELETE(MATRIX=GGT)
DIAGMULT(A=T,T,D=L,M=TTL,R)
```

```
MULT(A=TTL,B=T,M=COV)
SCALE(C=COV,R=CC)
MULT(A=XT,TA,B=T,M=XTT)
DIAGMULT(A=XTT,D=COV,C=0.5,R,M=P)
*
*   OVO JE IZMISLIO SCHOENEMANN, 1966,
*
PRINT(MATRIX=P,TEXT= HIPOTETSKI SKLOP)
PRINT(MATRIX=CC,TEXT=KORELACIJE HIPOTETSKIH FAKTORA,D)
*
*   OVO JE KONFIRMATIVNA SOLUCIJA, DOBIJENA
*   GADJANJEM U MATRICU CILJA.
*
*   QUAQUA SADA USPOREDJUJE DOBIJENA RIJESENJA,
*
MULT(A=Q,B=TTL,TB,M=MAJMUN)
DIAGMULT(A=MAJMUN,D=COV,C=-0.5,R,M=HAJVAN)
DIAGMULT(A=HAJVAN,D=C,C=-0.5,L,M=CROSSC)
PRINT(MATRIX=CROSSC,TEXT=KROSKOVARIJANCE FAKTORA,D)
MULT(A=A,TA,B=P,M=ATP)
DIAGMULT(A=ATP,D=COV,C=-0.5,R,M=ESHEK)
DIAGMULT(A=ESHEK,D=C,C=-0.5,L,M=CONGAP)
PRINT(MATRIX=CONGAP,TEXT=KONGRUENCIJA SKLOPOVA,D)
*
*   QUAQUA JE GOTOV.
*
```

ALGORITAM I PROGRAM ZA USPOREDBU EKSPLORATIVNE I KONFIRMATIVNE ORTHOBLIQUE FAKTORSKE SOLUCIJE*

PROT, F., K. BOSNAR, J. ŠTALEC i K. MOMIROVIĆ

Odjel za informatiku i statistiku
Fakulteta za fizičku kulturu
Sveučilišta u Zagrebu

SAŽETAK

Definiran je algoritam i napisan program za usporedbu eksplorativne i konfirmativne faktorske solucije skupa kvantitativnih varijabli. Objektivnu faktorsku strukturu, pod faktorskim modelom, algoritam odredjuje orthoblique transformacijom (Harris i Kaiser, 1964) unaprijed zadanog broja karakterističnih vektora reducirane matrice korelacija. Hipotetsku faktorsku strukturu analiziranog skupa varijabli algoritam odredjuje tako da formira kosu soluciju, ortonormalnom transformacijom svojstvenih vektora reducirane matrice korelacija Schönemannovim postupkom (Schönemann, 1966) usmjerenom na normiranu matricu cilja. Relacije dviju dobivenih solucija odredjene su kovarijancama latentnih dimenzija i koeficijentima kongruencije matrica sklopa.

---

* *Rad je izvorno napisan na engleskom jeziku i bit će prezentiran na Medjunarodnom simpoziju "Kompjuter na Sveučilištu", Cavtat, 1984.*

1. UVOD

Matematičko modeliranje bilo kojeg realnog procesa neophodan je osnov za njegovu racionalnu analizu. Modeliranje karakterističnih stanja kinezioloških transformacijskih procesa osnovano na analitičkim postupcima omogućava ne samo stvaranje podesnih obrazaca koji se mogu eksperimentalno ili opservacijom provjeriti, već takodjer i projektiranje optimalnih procedura koje maksimiziraju pozitivne efekte treninga.

S obzirom da se u procesu treninga i obuke u pravilu radi o transformaciji kompleksnih sistema malo je vjerojatno da je zadovoljavajuća reprezentacija cjelokupnog procesa ili pojedinih karakterističnih stanja moguća na osnovu izoliranih pojedinačnih dimenzija. Daleko je vjerojatnije da je strategija zasnovana na multivarijatnom faktorskom modelu optimalan pristup koji može polučiti zadovoljavajuće rezultate. No, kako solucije faktorske analize nisu potpuno nezavisne od postupaka kojima su podaci podvrgnuti (vidi Harris, 1968) optimalno rješenje za sigurnije zaključivanje o valjanosti hipotetskog modela jest provjera uz pomoć različitih faktorskih tehnika i usporedba dobivenih rezultata. Stabilnost rezultata će u tom slučaju biti argument koji ide u prilog postavljenom modelu.

Procedure koje objedinjuju eksplorativne i konfirmativne faktorske tehnike sretna su kombinacija koja istovremeno omogućuje identifikaciju što u podacima o modelu objektivno postoji te provjeru onoga što se pretpostavljalo da je o modelu bilo poznato.

Stoga, predloženi algoritam i program utvrdjuje faktorsku strukturu skupa kvantitativnih varijabli jednim postupkom eksplorativne faktorske analize i jednim postupkom konfirmativne analize. Bazična ortogonalna solucija je u ovom algoritmu definirana pod faktorskim modelom, s iterativnim odredjivanjem komunaliteta, uz unaprijed zadani broj faktora. Eksplorativna solucija odredjena je orthoblique transformacijom (Harris i Kaiser, 1964) sopstvenih vektora reducirane matrice korelacija varijabli. Konfirmativna solucija izvedena je Schoenemanovim postupkom za Prokrustovu transformaciju (Schönemann, 1966) koji obliknu soluciju odredjuje ortonormalnom transformacijom svojstvenih vektora reducirane matrice korelacije, usmjerenom na normiranu matricu hipotetskog sklopa.

Predloženi algoritam i program nastali su u prvom redu zbog potreba istraživanja i testiranja modela u kineziologiji, ali su primjereni primjeni i u drugim područjima gdje se pojavljuje problem utvrdjivanja faktorske strukture skupa kvantitativnih varijabli.

2. ALGORITAM

2.1 Uvodne definicije

Neka je

$$Z = (z_{ij}) \qquad i=1,\ldots,n \quad j=1,\ldots,m$$

matrica podataka dobivena opisom entiteta e_i iz skupa $E=\{e_i;\ i=1,\ldots,n\}$ formiranog nekim slučajnim postupkom iz populacije P, nad skupom standardiziranih kvantitativnih varijabli $V=\{v_j;\ j=1,\ldots,m\}$ odabranog na osnovu nekog teoretskog modela o strukturi latentnih dimenzija nekog latentnog sistema H. Neka je postavljena hipoteza o latentnoj strukturi manifestnih varijabli odredjenoj nad latentnim sistemom $H=\{h_k;\ k=1,\ldots,\ell\}$ i reprezentirana matricom

$$S = (s_{jk}) \qquad j=1,\ldots,m \quad k=1,\ldots,\ell$$

koja predstavlja hipotetsku matricu sklopa formiranu tako da je $s_{jk}=1$ ako je varijabla v_j mjera neke latentne dimenzije h_k ; $s_{jk}=0$ ako varijabla v_j nije direktna mjera dimenzije h_k, uz dodatni uvjet da jedna varijabla iz skupa V može biti označena kao mjera samo jedne dimenzije iz H.

2.2 Bazična ortogonalna solucija

Pod faktorskim modelom

$$Z = \Phi S^T + N \qquad \left| \begin{array}{l} \Phi^T\Phi \frac{1}{n} = M \\ \operatorname{diag} M = I \\ \Phi^T N = 0 \\ N^T N = U^2 = \operatorname{diag} \end{array} \right.$$

gdje su u matrici $\Phi=(\phi_{ik})$; $i=1,\dots,n$, $k=1,\dots,\ell$ nepoznate faktorske vrijednosti entiteta, u matrici $N=(n_{ij})$; $i=1,\dots,n$, $j=1,\dots,m$ nepoznate unikne vrijednosti entiteta za svaku od varijabli, za relacije manifestnih varijabli vrijedi:

$$R = Z^T Z \frac{1}{n} = (\Phi S^T + N)^T(\Phi S^T + N) \frac{1}{n} = SMS^T + U^2.$$

Procjena faktorskih vrijednosti pod modelom najmanjih kvadrata izvedena regresijskom matricom

$$\beta = S(S^T S)^{-1}$$

je

$$L = Z\beta,$$

sa matricom kovarijanci

$$C = L^T L \frac{1}{n} = \beta^T R \beta = (S^T S)^{-1} S^T R S (S^T S)^{-1}.$$

U dijagonali matrice C nalaze se varijance procijenjenih faktorskih vrijednosti

$$V^2 = \text{diag } C,$$

tako da su standardizirane faktorske vrijednosti odredjene sa

$$\Phi = L V^{-1}$$

a korelacije faktorskih vrijednosti sa

$$M = \Phi^T \Phi \frac{1}{n}.$$

Inicijalna procjena komunaliteta h_o^2 se može izvesti i preko matrice strukture varijabli $H_o = Z^T \Phi^* \frac{1}{n}$ u prostoru Mahalanobisove transformacije faktorskih vrijednosti $\Phi^* = \Phi M^{1/2}$

$$h_o^2 = \text{diag } (H_o H_o^T).$$

Konačna procjena komunaliteta je odredjena iterativnim procesom:

$$R_\alpha = R - I + h_{\alpha+1}^2$$

$$(R_\alpha - \lambda_{k\alpha} I) X_{k\alpha} = 0 \qquad k=1,\dots,\ell$$

$$H_\alpha = X_\alpha \Lambda_\alpha^{1/2}$$

$$h_\alpha^2 = \text{diag}(H_\alpha H_\alpha^T)$$

gdje je α oznaka iteracije.

Iterativni proces se zaustavlja kada se zadovolji uvjet

$|h^2_\alpha| - |h^2_{\alpha+1}| < \varepsilon$, gdje je ε proizvoljno mali broj (u programu QUAQUA $\varepsilon=0.005$) ili jednu iteraciju prije nego što nastupi generalizirani Heywoodov slučaj. Na kraju procesa u matrici $H=X\Lambda^{1/2}$, gdje su Λ i X matrice svojstvenih vrijednosti i svojstvenih vektora matrice $R-I+h^2_\alpha$ dobivene u posljednjoj iteraciji, su glavne osovine reducirane matrice korelacija procijenjene pod standardnim faktorskim modelom, pa je

$$R = X\Lambda X^T + U^2 \text{ , a } \qquad U^2 = \mathrm{diag}(H H^T).$$

2.3 <u>Orthoblique solucija</u>

Objektivno parsimonijsko rješenje izvedeno je orthoblique transformacijom (Harris i Kaiser, 1964).

Neka je T; $T^TT=TT^T=I$ ortonormalna matrica koja postmultiplicirajući matricu karakterističnih vektora reducirane matrice korelacija

$$XT = (v_{jk})$$

zadovoljava uvjet

$$m \sum_{j=1}^{m} \sum_{k=1}^{\ell} v^4_{jk} - \sum_{k=1}^{\ell} \left(\sum_{j=1}^{m} v^2_{jp} \right)^2 = \max \qquad \begin{array}{l} j=1,\ldots,m \\ k=1,\ldots,\ell \end{array}$$

tj. maksimizira Kaiserov varimax kriterij.

U matrici

$$A = XTD$$

gdje je $D = (\mathrm{diag}(T^T\Lambda T))^{1/2}$ nalaze se koordinate vektora varijabli u novom koordinatnom sustavu, a u

$$F = X\Lambda T D^{-1}$$

nalaze se korelacije varijabli i orthoblique faktora.

Orthoblique faktorske vrijednosti entiteta odredjuje

$$\Psi = ZXT\,D^{-1}$$

sa njihovim relacijama u

$$W = \Psi^T\Psi\,\frac{1}{n} = D^{-1}T^T\Lambda T\,D^{-1}.$$

2.4 Schoenemannova transformacija ciljana na hipotezu

Neka je hipotetski sklop normiran po kolonama

$$Y = S\,(\mathrm{diag}(S^TS))^{-1/2}$$

Schoenemannov problem (Schönemann, 1966)

$$XQ = Y + E \quad \left|\quad \begin{array}{l} \mathrm{tr}\ E^TE = \min \\ Q^TQ = QQ^T = I \end{array}\right.$$

se može riješiti tako da se matrica dobivena produktom svojstvenih vektora reducirane matrice korelacija i normirane matrice cilja

$$G = X^TY$$

podvrgne spektralnoj dekompoziciji (Cvitaš, Momirović, 1984; Viskić-Štalec, Štalec, Momirović, 1984)

$$G = L^{*}\Delta^{*}D^{*T}$$

pa je Prokrustna transformacijska matrica Q odredjena produktom lijevih (L^{*}), desnih (D^{*}) svojstvenih vektora od G

$$Q = L^{*}D^{*T}.$$

U tom slučaju su faktorske vrijednosti dobivene u okviru osnovnog faktorskog modela

$$Z = \Gamma^{*}P^{*T} + N$$

gdje je $P=XQ$ matrica koordinata varijabli u ovoj soluciji, u tom su slučaju faktorske vrijednosti entiteta pod kriterijem najmanjih kvadrata, odredjene operacijom

$$\Gamma^{*} = ZP^{*}(P^{*T}P^{*})^{-1} = ZXQ.$$

Ako su

$$M^{*} = \Gamma^{*T}\Gamma^{*}\frac{1}{n} = Q^{T}\Lambda Q$$

kovarijance tako dobivenih faktorskih vrijednosti sa pripadnim varijancama u

$$V^{*2} = \mathrm{diag}(Q^{T}\Lambda Q),$$

onda su standardizirane faktorske vrijednosti determinirane sa

$$\Gamma = \Gamma^{*}V^{*-1},$$

a korelacije faktorskih vrijednosti sa

$$M = V^{*-1}Q^T\Delta QV^{*-1}.$$

Matricom

$$K = Z^T\Gamma\,\frac{1}{n} = RXQV^{*-1} = X\Lambda QV^{*-1}$$

odredjene su korelacije varijabli i faktora hipotetske solucije.

Matrica sklopa, odnosno koordinata varijabli u standardiziranom faktorskom prostoru tada je

$$P = KM^{-1} = X\Lambda QV^{*-1}(V^{*-1}Q^T\Lambda QV^{*-1})^{-1} = XQV^*.$$

2.5 Relacije orthoblique i Schönemannove solucije

Provjera sukladnosti objektivne i hipotetske solucije odredjena je kroskovarijancama standardiziranih faktorskih vrijednosti entiteta

$$M_{12} = \Psi^T\Gamma\,\frac{1}{n} = D^{-1}T^T\Lambda QV^{*-1}$$

i koeficijentima kongruencije koordinata varijabli u prostoru faktora obje solucije

$$C_{12} = D^{-1}A^TPV^{*-1}.$$

3. PROGRAM

Program QUAQUA koji realizira predloženi algoritam napisan je u verziji 5.2/M programskog sistema SS. Program ima 67 linija izvršnog koda i 100 linija komentara koje čine program samodokumentiranim. Program je pohranjen u bibliotekama FFK*WORK i SRCE*SS-MAKRO. Cjelokupni način aktiviranja programa, kao i način kontrolnog opisa podataka opisan je u Priručniku za korištenje programskog sistema SS (Štalec, Momirović i Zakrajšek, 1983).

QUAQUA može analizirati skupove od do 250 varijabli na uzorcima od do 10.000 entiteta.

Iz javne biblioteke SRCE*SS-MAKRO program se može aktivirati ovim naredbama

```
∂RUN <RUN-ID, ACC-ID/USER,ID,PROJECT>
∂ASG,A SRCE*SS-MAKRO.
∂ADD,L SRCE*SS-MAKRO. IZVEDI
PROGRAM QUAQUA
PODACI <FILE1, ELT1    FILE2, ELT1>
VARS   <FILE1, ELT2    FILE2, ELT2>
∂FIN
```

Program očekuje da je u elementu 1 File-a 1 matrica podataka, a u elementu 1 File-a 2 selektorska matrica S, napisana u transponiranom obliku. Kontrolni opis za matricu podataka program očekuje u elementu 2 File-a 1, a kontrolni opis za selektorsku matricu program očekuje u elementu 2 File-a 2.

Osim kvantitativnih varijabli QUAQUA može analizirati i skupove ordinalnih varijabli, ako se razumno može pretpostaviti da su te kategorije posljedica djelovanja nekog filterskog sustava na izlaze iz nekog (nepoznatog) multivarijatnog generatora. U takovom slučaju treba matricu ulaznih podataka neposredno nakon učitavanja normalizirati aktiviranjem naredbe NORMALISATION, tj. naredbu INPUT zamijeniti naredbama

```
INPUT(SCORE=SCORR)
NORMALISATION(OLDSC=SCORR,NEWSC=SCORE)
DELETE(MATRIX=SCORR)
```

4. NUMERIČKI PRIMJER

Na uzorku od 90 ispitanika izmjeren je rezultat u 9 laboratorijskih testova vožnje biciklergometra. Rezultat je definiran ukupnim brojem okretaja pedala biciklergometra u zadanom vremenu, pod zadatim opterećenjem, iz kojeg je parcijaliziran efekat morfoloških karakteristika. Ispitanici su izveli slijedeće testove:

AD30 - vožnja u trajanju od 30 sekundi, pod konstantnim otporom biciklergometra

AD60 - vožnja u trajanju od 60 sekundi pod konstantnim otporom biciklergometra

AD90 - vožnja u trajanju od 90 sekundi pod konstantnim otporom biciklergometra

RD05 - vožnja u trajanju od 5 minuta pod konstantnim otporom biciklergometra

RD10 - vožnja u trajanju od 10 minuta pod konstantnim otporom biciklergometra

RD15 - vožnja u trajanju od 15 minuta pod konstantnim otporom biciklergometra

SD05 - vožnja u trajanju od 5 minuta pod zadanim variranjem otpora biciklergometra

SD09 - vožnja u trajanju od 9 minuta pod zadanim variranjem otpora biciklergometra

SD12 - vožnja u trajanju od 12 minuta pod zadanim variranjem otpora biciklergometra.

Pretpostavljeno je da u osnovi efikasnosti u ovih devet laboratorijskih situacija egzistiraju tri latentna mehanizma. Prvi odgovoran za efikasnost izvršenja kratkotrajnog rada, sa konstantnim opterećenjem, drugi odgovoran za efikasnost izvršenja dugotrajnog rada sa konstantnim opterećenjem i treći, odgovoran za efikasnost izvršenja rada u varijabilnim uvjetima. Ova je hipoteza predstavljena u obliku selektorske matrice predočene u tabeli 1. Sklop objektivne orthoblique solucije predočen je u tabeli 2, a sklop hipotetske solucije predočen je u tabeli 3. U tabeli 4 su predočene kroskovarijance faktorskih vrijednosti entiteta u orthoblique i hipotetskoj soluciji, i kongruence matrica sklopa ove dvije solucije.

Na osnovu dobivenih rezultata može se zaključiti da je algoritam QUAQUA u obje solucije zadovoljavajuće reproducirao odnos varijabli i dimenzija predstavljen hipotezom, pa se u tabeli 4 mogu prepoznati korespondentni faktori objektivne i hipotetske solucije što ide u prilog hipoteze o egzistenciji tri latentna mehanizma odgovorna za efikasnost u ovom skupu laboratorijskih testova.

Tabela 1 - HIPOTEZA

	HIP 1	HIP 2	HIP 3
AD30	1	0	0
AD60	1	0	0
AD90	1	0	0
RD05	0	1	0
RD10	0	1	0
RD15	0	1	0
SD05	0	0	1
SD09	0	0	1
SD12	0	0	1

Tabela 2 - SKLOP OBJEKTIVNE ORTHOBLIQUE SOLUCIJE

	OBQ 1	OBQ 2	OBQ 3
AD30	.57	.11	.07
AD60	1.04	-.04	-.17
AD90	.71	-.04	.09
RD05	.12	-.08	.74
RD10	.02	-.05	.81
RD15	-.05	.09	.69
SD05	.02	.85	-.01
SD09	-.01	.80	.11
SD12	.04	.54	-.08

Tabela 3 - SKLOP HIPOTETSKE SOLUCIJE

	PRX 1	PRX 2	PRX 3
AD30	-.05	.64	.04
AD60	-.41	1.01	.20
AD90	-.10	.74	.20
RD05	.62	.27	.29
RD10	.71	.19	.25
RD15	.65	.13	.09
SD05	.16	.30	-.75
SD09	.26	.29	-.67
SD12	.03	.21	-.48

Tabela 4 - KOVARIJANCE ORTHOBLIQUE I HIPOTETSKE SOLUCIJE

	PRX 1	PRX 2	PRX 3
OBQ 1	.11	(.95)	.27
OBQ 2	.46	.38	(-.99)
OBQ 3	(.94)	.57	.14

KONGRUENCE ORTHOBLIQUE I HIPOTETSKOG SKLOPA

	PRX 1	PRX 2	PRX 3
OBQ 1	-.27	(.94)	.28
OBQ 2	.22	.29	(-.93)
OBQ 3	(.94)	.20	.28

LITERATURA

1. Cvitaš, M. and K. Momirović:
Note on some properties of oblique Procrustes transformations by orthogonal rotations. Unpublished paper, 1984.

2. Fulgosi, A.:
Faktorska analiza. Školska knjiga, Zagreb, 1979.

3. Green, B.T.:
The orthogonal approximation of an oblique structure in factor analysis. Psychometrika, 17 (1952), 429-440.

4. Harris, C.W. and H.F. Kaiser:
Oblique factor analytic solutions by orthogonal transformations. Psychometrika, 29 (1964), 347-362.

5. Harris, C.W.:
On factors and factor scores. Psychometrika, 32 (1967), 363-379.

6. Holzinger, K.J. and H.H. Harman:
Factor analysis - a synthesis of factorial methods. University of Chicago Press, Chicago, 1941.

7. Horst, P.:
Factor analysis of data matrices. Holt, Rinehart and Winston, New York, 1965.

8. McDonald, R.P: and E.J. Bun: A comparison of four methods of constructing factor scores. Psychometrika, 32 (1967), 381-401.

9. McDonald, R.P.:
Descriptive axioms for common factor theory, image theory and component theory. Psychometrika, 40 (1975), 137-152.

10. Mulaik, S.A.:
The fondations of factor analysis. McGraw-Hill, New York, 1972.

11. Rudan, P., L. Szirovicza and K. Momirović:
The application of an algorithm based on Mahalanobis angles and iterative Q-method of taxonomic analysis in the study of micro-evolution. Periodicum Biologorum, 81 (1979), 583-589.

12. Schönemann, P.H.:
A generalized solution of orthogonal Procrustes problem. Psychometrika, 31 (1966), 1-10.

13. Štalec, J., K. Momirović i E. Zakrajšek:
Programski sistem SS. Fakultet za fizičku kulturu Sveučilišta u Zagrebu, Zagreb, 1983.

14. Viskić-Štalec, N., J. Štalec, and K. Momirović:
Oblique Procrustes solutions by orthogonal rotations. International Symposium on Computational Statistics COMPSTAT '84, Prague, 1984.

3. NAČIN AKTIVIRANJA PROGRAMA

Izvodjenje obrade sa nekim od programa provodi se na slijedeći način:

```
∂RUN        RUN-ID,ACC-NO/USER-ID,PROJEKT
∂ASG,A      SRCE*SS-MAKRO.
∂ADD,L      SRCE*SS-MAKRO.IZVEDI
PROGRAM     <IME-PROGRAMA>*
PODACI      <FILE1,EL1   FILE2,EL2   FILE3,EL3   FILE4,EL4>
NUMERUS     N
VARS        <FILE1,EL1   FILE2,EL2   FILE3,EL3   FILE4,EL4>
BRFAKTORA   K1, K2
MIN         'M1'  'M2'
TRACE       T1  T2
PRINT
∂FIN
```

Od parametarskih kartica (4-11) obavezne su samo "PROGRAM", "PODACI" i "VARS" kartice.

Kartica "NUMERUS" obavezna je samo za programe za analizu nenumeričkih podataka, i tada se umjesto N upisuje broj podataka.

Kartice "BRFAKTORA", "MIN" i "TRACE" navode se prilikom aktiviranja onih programa kod kojih se kriterij za odredjivanje broja faktora izvana zadaje, tj. ovisi o dimenzijama podataka (M,N) ili ako se želi promijeniti standardno ugradjeni kriterij.

Kartica "BRFAKTORA" obavezna je samo za programe sa unaprijed definiranim brojem faktora i tada se umjesto K1 (K2) upisuje broj koliko se faktora (komponenata) želi zadržati (npr. za program QUAQUA). Pri tome se K1 odnosi na prvu logičku grupu podataka, a K2 na drugu. Ako postoji samo jedna skupina podataka navodi se samo K1.

Kartica "MIN" obavezna je samo za programe sa kriterijem zavisnim o dimenzijama podataka (M,N). Tada se umjesto M1 (M2) upisuje najmanja relativna veličina varijance glavne komponente

---

* *Unutar oznaka < > nalaze se opcionalni parametri, tj. parametri s promjenjivim vrijednostima.*

prvog skupa podataka (M1), odnosno drugog skupa podataka (M2). Umjesto na skupove podataka, M1 i M2 se mogu odnositi i na različite metode za transformaciju istog skupa podataka unutar jednog programa.

Kartica "TRACE" se navodi onda kada se želi promijeniti standardno definirani kriterij za odredjivanje broja faktora. Tada se umjesto T1 (T2) navodi dio traga (u postocima) koji mora biti objašnjen zadržanim vektorima. Pri tome se T1 odnosi na prvi skup (ili metodu za transformaciju podataka), a T2 na drugi skup ili drugu metodu.

Kartica "PRINT" navodi se ako se želi štampati sve rezultate, osim onih standardnih koji se dobivaju samim pozivom programa. Na kartici "PROGRAM" treba umjesto ime programa utipkati ime programa koji se želi izvesti.

Na kartici "PODACI" treba navesti imena file-ova i elemenata u kojima se nalaze podaci i to: umjesto FILE1,EL1 treba utipkati imena file-a i elementa u kojima se nalazi prva grupa podataka. Umjesto FILE2,EL2 imena file-a i elementa koji sadrže drugu grupu podataka, itd. Zavisno o broju logičkih grupa podataka koje navedeni program zahtijeva.

Umjesto FILEn,ELn (n=1, 2, 3, 4) na kartici "PODACI" može se navesti samo FILEn, ako su podaci u file-u FILEn, a ne u elementu programskog FILE-a FILEn.

Na kartici "VARS" treba navesti imena file-ova i elemenata koji sadrže VARS kartice (utipkane prema uputama za SS jezik) na potpuno isti način kao i na kartici "PODACI".

Parametarske kartice se navode proizvoljnim redoslijedom.

Npr. sa RUN-om:

```
∂RUN
∂ASG,A   SRCE*SS-MAKRO.
∂ADD     SRCE*SS-MAKRO,IZVEDI
PROGRAM QCCR
PODACI  AA*POD1,ANT  BB*POD2
VARS    CC*POD3,VARS1  AA*POD3,VARS2
∂FIN
```

Izvesti će se program QCCR sa prvom grupom podataka spremljenom u katalogiziranom file-u AA*POD1, elementu ANT i drugom grupom podataka spremljenom u file-u BB*POD2. VARS kartice za prvu grupu podataka nalaze se u katalogiziranom file-u CC*POD3, elementu VARS, za drugu u file-u AA*POD2, elementu VARS2.

Prije aktiviranja pojedinog programa potrebno je pročitati opis programa, te provjeriti da li su obavezne parametarske kartice "NUMERUS", "BRFAKTORA", "MIN" ili "TRACE" te da li program dozvoljava izmjene kriterija za broj zadržanih faktora.