Konstantin Momirović, Marijan Gredelj i Maja Herak
Sveučilišni računski centar, Zagreb

COCAIN — ALGORITAM I PROGRAM ZA KANONIČKU KORELACIJSKU ANALIZU

SAŽETAK

Definiran je algoritam i napisan program za kanoničku korelacijsku analizu kojoj je pridružena ortogonalna parsimonijska transformacija kanoničkih dimenzija dva skupa varijabli prethodno transformiranih u standardizirane glavne komponente.

0. UVOD

Najveća poteškoća u primjeni kanoničke korelacijske analize povezana je sa interpretacijom kanoničkih varijabli. Ova interpretacija složena je i onda kada se kanonička korelacijska analiza tretira ne samo kao statistička metoda, nego i kao metoda za analizu podataka koja pripada generaliziranom faktorskom modelu. Statistički značajne kanoničke dimenzije nisu istovremeno i informatički važne varijable što rezultira dijelom poteškoća. Drugi dio neugodnosti proizlazi iz u pravilu nejednostavnog oblika kanoničkih faktora pa je zbog toga interpretacija tih faktora često nesigurna.

Izvjesni problemi, tehničke prirode, mogu se javiti u slučaju skoro singularnih skupova varijabli jer je tada osnovna matrica za određivanje kanoničkih korelacija i vektora transformacije varijabli u kanoničke varijable vrlo slabo određena.

Algoritam COCAIN i njemu pridružen program pokušava u izvjesnoj mjeri reducirati većinu ovih poteškoća. On zadržava kao interpretativno važne samo one kanoničke dimenzije koje su pridružene kanoničkim korelacijama čiji su kvadrati veći od prosjeka koeficijenata determinacije ma kojeg od analiziranih skupova varijabli određenih na temelju drugog skupa.

Osim toga COCAIN transformira zadržane kanoničke dimenzije u ortogonalnu parsimonijsku poziciju što može olakšati interpretaciju kanoničkih faktora jer i nakon te transformacije vektori kanoničkih varijabli razapinju isti zajednički prostor.

COCAIN određuje također kanoničke relacije skupova varijabli nakon transformacije tih skupova u glavne komponente sa nenultim varijancama pa tako izbjegava problem netočno određenih inverza matrice korelacija kod skoro singularnih skupova varijabli i dalje olakšava interpretaciju kanoničkih faktora eksplicitnim određivanjem relacija tih faktora i glavnih komponenata analiziranih skupova.

1. MODEL KANONIČKE KORELACIJSKE ANALIZE

Neka su $P=\{p_j;\ j=1,\dots,m\}$ i $K=(k_l;\ l=1,\dots,r)$; $r \leq m$ dva skupa multivarijatno normalno distribuiranih varijabli sa vektorima aritmetičkih sredina μ i η i matricama kovarijanci Σ i Γ.

Neka su

$$Z^T = (z_1,\dots,\ z_j,\dots,\ z_m)$$

i

$$C^T = (C_1,\dots,\ C_l,\dots,\ C_r)$$

vektori kojima je neki entitet opisan nad skupovima P i K. Neka su

$$B^T_t = (b_{t1},\dots,\ b_{tj},\dots,\ b_{tm}) \qquad t=1,\dots,r$$

$$A^T_t = (a_{t1},\dots,\ a_{tl},\dots,\ a_{tr}) \qquad t=1,\dots,r$$

vektori odabrani tako, da uz uvjete

$$E((Z-\mu)^T B_t,\ (Z-\mu)^T B_t) = 1 \qquad t=1,\dots,r$$
$$E((C-\eta)^T A_t,\ (C-\eta)^T A_t) = 1 \qquad t=1,\dots,r$$
$$E((Z-\mu)^T B_t,\ (Z-\mu)^T B_u) = 0 \qquad t,u=1,\dots,r \quad t\neq u$$
$$E((C-\eta)^T A_t,\ (C-\eta)^T A_u) = 0 \qquad t,u=1,\dots,r \quad t\neq u$$

i

$$E((Z-\mu)^T B_t,\ (C-\eta)^T A_u) = 0 \qquad t,u=1,\dots,r \quad t\neq u$$

zadovoljavaju uvjete

$$\alpha_t = E((Z-\mu)^T B_t,\ (C-\eta)^T A_t) = \max \qquad t=1,\dots,r$$

i

$$\alpha_t > \alpha_{t+1} \qquad t=1,\dots,r-1$$

Varijable

$$g_t = (Z-\mu)^T B_t \qquad t=1,\dots,r$$

i

$$h_t = (C-\eta)^T A_t \qquad t=1,\dots,r$$

definirane su tada kao kanoničke varijable u koje su transformirane varijable iz P i K, a koeficijenti α_t, $t=1,\dots,r$ kao kanoničke korelacije između varijabli g_t i h_t, pa zato i između varijabli iz skupova P i K.

2. ALGORITAM

Neka je $P = \{p_j;\ j=1,\dots,m\}$ neki skup varijabli i neka je $K = \{k_l;\ l=1,\dots,r\}$ neki drugi skup varijabli izabran tako da je $P\ K = 0$. Pretpostavimo da su sve varijable iz P i K multivarijatno normalno distribuirane sa vektorima aritmetičkih sredina μ i η matricama kovarijanci Σ i Γ.

Kako je, za uspostavljanje kanoničkih relacija između skupova P i K irelevantna metrika varijabli, mogu se te varijable, bez gubitka značajnih informacija, standardizirati tako da su, za varijable iz P i K, $\mu=0$ i $\eta=0$, a $\Sigma=\mathrm{P}$ i $\Gamma=\mathrm{M}$, gdje su P i M matrice korelacija između tih varijabli.

Neka je $U = \{u_i;\ i=1,\dots,n\}$ uzorak entiteta izabran iz neke populacije Π. Neka je Z matrica tipa $\mathrm{Z}=(z_{ij})$, $i=1,\dots,n;\ j=1,\dots,m$ koja sadrži standardizirane varijable iz P dobijena na U, i neka je C matrica tipa $\mathrm{C}=(c_{il})$, $i=1,\dots,n;\ l=1,\dots,r$ koja sadrži standardizirane varijable iz K dobijene na U. Pretpostavimo da vrijedi $r \leqq m$.

Procjena korelacijske matrice P bit će korelacijska matrica, dobijena pod kriterijem najveće vjerodostojnosti.

$$\mathrm{R} = \mathrm{Z}^T \mathrm{Z} \frac{1}{n}$$

a procjena korelacijske matrice Γ, također pod kriterijem najveće vjerodostojnosti, korelacijska matrica

$$\mathrm{S} = \mathrm{C}^T \mathrm{C} \frac{1}{n}$$

Označimo sa $\Lambda=(\lambda_p)$, $p=1,\dots,q \leqq m$ dijagonalnu matricu nenultih svojstvenih vrijednosti matrice R, a sa $\Delta=(\delta_t)$, $t=1,\dots,w \leqq r$ matricu nenultih svojstvenih vrijednosti matrice S, i pretpostavimo da sada vrijedi $w \leqq q$. Matricu svojstvenih vektora pridruženih svojstvenim vrijednostima iz Λ označimo sa X, i standardizirajmo ih tako da je $\mathrm{X}^T\mathrm{X} = \mathrm{I}_q$, a matricu jednako normiranih svojstvenih vektora pridruženih svojstvenim vrijednostima iz Δ označimo sa Y, tako da je $\mathrm{Y}^T\mathrm{Y} = \mathrm{I}_w$.

U matricama

$$\mathrm{K}^* = \mathrm{Z}\,\mathrm{X}$$

i

$$\mathrm{L}^* = \mathrm{C}\,\mathrm{Y}$$

bit će glavne komponente varijabli iz P i K sa matricama kovarijanci Λ i Δ, a u matrici

$$\mathrm{Q}^* = \mathrm{K}^{*T}\mathrm{L}^* \frac{1}{n} = \mathrm{X}^T\mathrm{V}\mathrm{Y}$$

gdje je $\mathrm{V}=\mathrm{Z}^T\mathrm{C}\frac{1}{n}$, matrica kroskorelacija varijabli iz P i K, kovarijance glavnih komponenata iz K^* L^*.

Naravno, varijance glavnih komponenata irelevantne su za kanonički problem. Definirajmo stoga matrice standardiziranih glavnih komponenata

$$\mathrm{K} = \mathrm{Z}\mathrm{X}\Lambda^{-1/2}$$
$$\mathrm{L} = \mathrm{C}\mathrm{Y}\Delta^{-1/2}$$

sa matricama kovarijanci I_q i I_w i matricom kroskorelacija

$$\mathrm{Q} = \Lambda^{-1/2}\mathrm{X}^T\mathrm{V}\mathrm{Y}\Delta^{-1/2}$$

pa se problem određivanja kanoničkih korelacija između sistema P i K svodi na rješavanje sustava karakterističnih jednadžbi

$$(\mathrm{Q}^T\mathrm{Q} - \alpha_t^2\mathrm{I})\mathrm{A}_t = 0 \qquad t=1,\dots,w$$

pri čemu su koeficijenti α_t, $t=1,\dots,w$ kanoničke korelacije između ortogonalno transformiranih glavnih komponenata varijabli iz P i K, pa zato i između kanoničkih varijabli dobijenih iz tih sustava varijabli.

Organizirajmo vektore A_t u matricu $\mathrm{A}=(a_{lt})$ i odredimo matricu

$$\mathrm{B} = \mathrm{Q}\mathrm{A}\alpha^{-1} = (b_{jt}) \qquad j=1,\dots,m \quad t=1,\dots,w$$

gdje je $\alpha=(\alpha_t)$ dijagonalna matrica koja sadrži kanoničke korelacije između varijabli iz P i K.

Matrice B i A sadrže koeficijente transformacije standardiziranih glavnih komponenata iz K i L u kanoničke varijable, sadržane u matricama

$$\mathrm{G} = \mathrm{K}\mathrm{B} = \mathrm{Z}\mathrm{X}\Lambda^{-1/2}\mathrm{B}$$
$$\mathrm{H} = \mathrm{L}\mathrm{A} = \mathrm{C}\mathrm{Y}\Delta^{-1/2}\mathrm{A}$$

tako da je

$$\alpha = \mathrm{G}^T\mathrm{H}\frac{1}{n}$$

a kako je $\mathrm{G}^T\mathrm{G}\frac{1}{n} = \mathrm{I}_w$ i $\mathrm{H}^T\mathrm{H}\frac{1}{n} = \mathrm{I}_w$, to su i matrice B i A ortogonalne transformacijske matrice.

Očito, $\mathrm{K}^T\mathrm{G}\frac{1}{n} = \mathrm{B}$ i $\mathrm{L}^T\mathrm{H}\frac{1}{n} = \mathrm{A}$, pa su matrice B i A istovremeno i faktorske matrice kanoničkih latentnih dimenzija izvedene u prostoru glavnih komponenata.

Faktorske matrice kanoničkih latentnih dimenzija u prostoru varijabli iz P i K su

$$\mathrm{F} = \mathrm{Z}^T\mathrm{G}\frac{1}{n} = \mathrm{X}\Lambda^{1/2}\mathrm{B}$$

i

$$\mathrm{E} = \mathrm{C}^T\mathrm{H}\frac{1}{n} = \mathrm{Y}\Delta^{1/2}\mathrm{A}$$

pa su, prema tome, B i A ortogonalne transformacije glavnih osovina matrica R i S.

Naravno, matrice koje sadrže koeficijente transformacije varijabli iz P i K u kanoničke varijable mogu također biti od značaja za interpretaciju kanoničkih varijabli i za primjenu dobijenih rezultata. To su matrice

$$B^* = X\Lambda^{-1/2}B$$

i

$$A^* = Y\Delta^{-1/2}A$$

pa su očito rezultat ortogonalnih transformacija matrica za izračunavanje standardiziranih glavnih komponenata.

Od nekog interpetativnog značaja, i od nesumnjivog značaja za primjenu rezultata su i matrice kroskorelacija između varijabli iz P i kanoničkih varijabli, dobijenih transformacijom varijabli iz sustava K, a možda i kroskorelacije između varijabli iz K i kanoničkih varijabli, dobijenih transformacijom varijabli iz sustava P. Te su kroskorelacije u matricama

$$F^* = Z^T H \frac{1}{n} = V Y \Delta^{-1/2} A$$

i

$$E^* = C^T G \frac{1}{n} = V^T X \Lambda^{-1/2} B$$

pa kanonička korelacijska analiza, u okviru algoritma koji, prethodno, faktorizira sustave varijabli, omogućava potpuni uvid u strukturalne odnose između skupova varijabli P i K.

3. TESTOVI HIPOTEZA O PARAMETRIMA

Test hipoteze $\alpha_1=0$ je, istovremeno, i test hipoteze $V=0$ i $Q=0$, dakle test hipoteze o potpunoj nezavisnosti skupova P i K. Ta se hipoteza može testirati na temelju vrijednosti varijable

$$\chi^2_1 = -a \log_e \eta_1$$

gdje je

$$a = (n-1) - (q+w)/2$$

a

$$\eta_1 = \prod_{t=1}^{w} (1-\alpha_t^2)$$

jer je χ^2_1 distribuiran u skladu sa χ^2 raspodjelom za $q * w$ stupnjeva slobode.

Ako se hipoteza $\alpha_1=0$ odbaci, i prihvati alternativna hipoteza da je prvi par kanoničkih varijabli u nenultoj korelaciji, mogu se testirati i hipoteze o povezanosti ostalih kanoničkih varijabli. Te se hipoteze, sukcesivno, mogu testirati na temelju vrijednosti varijabli

$$\chi^2_t = -a \log_e \eta_t \qquad t=1,\ldots,w$$

gdje je

$$\eta_t = \prod_{p=t+k}^{w} (1-\alpha_p) \qquad t=1,\ldots,w$$

$$k=0,1,\ldots,w-1$$

jer varijable χ^2_t imaju i dalje χ^2 raspodjelu sa $(q-k) * (w-k)$ stupnjeva slobode.

Varijance koeficijenata u matricama A, B, F, E, F* i E* su, aproksimativno, $\sigma^2=n^{-1}$, pa ta vrijednost može, uz znatan oprez, biti upotrebljena za testiranje hipoteza o koeficijentima u tim matricama i konstrukciju intervala pouzdanosti.

Broj interpretativno smislenih kanoničkih dimenzija bolje je, međutim, odrediti na temelju kriterija

$$r^* = \text{num}\,(\alpha_t^2 > \alpha^*)$$

gdje je

$$\alpha^* = \sum_{t=1}^{w} \alpha_t^2 \frac{1}{w}$$

dakle prosječan koeficijent determinacije varijabli iz K na temelju sustava varijabli iz P.

4. ANALIZA JEDNOSTAVNE STRUKTURE

Neka su sada

$$F = (f_{jt}) \qquad j=1,\ldots,m \quad t=1,\ldots,r^*$$

i

$$E = (e_{lt}) \qquad l=1,\ldots, \quad t=1,\ldots,r^*$$

kanoničke faktorske matrice koje sadrže samo r* značajnih kanoničkih dimenzija. Nađimo ortonormalne transformacijske matrice T_p i T_k koje, nakon transformacija

$$F_p = F\,T_p = (f^*_{jt}) \qquad j=1,\ldots,m \quad t=1,\ldots,r^*$$

i

$$E_k = E\,T_k = (e_{lt}) \qquad l=1,\ldots,r \quad t=1,\ldots,r^*$$

maksimiziraju funkcije

$$v_p = m \sum_{j=1}^{m}\sum_{t=1}^{r^*} f^{*4}_{jt} - \sum_{t=1}^{r^*}\Big(\sum_{j=1}^{m} f^{*2}_{jt}\Big)^2$$

i

$$v_k = r \sum_{l=1}^{r}\sum_{t=1}^{r^*} e^4_{lt} - \sum_{t=1}^{r^*}\Big(\sum_{l=1}^{r} e^2_{lt}\Big)^2$$

Sada su matrice koeficijenata transformacija varijabli iz P i K u pseudokanonističke varijable

$$B^*_p = X\Lambda^{-1/2}B\,T_p$$

i

$$A^*_k = Y\Delta^{-1/2}A\,T_k$$

a u matrici

$$\beta = T_p^T\,\alpha\,T_k$$

gdje je α matrica koja sadrži samo prvih r* značajnih koeficijenata kanoničke korelacije, bit će koeficijenti korelacije između pseudokanoničkih varijabli dobijenih nakon transformacije strukturalnih matrica u jednostavan oblik.

5. PROGRAM

Program za kanoničku korelacijsku analizu osnovan na algoritmima opisanim u sekcijama 2., 3. i 4. napisan je u SS jeziku i omogućava analizu do 10000 entiteta opisanih nad skupom od 250 eksplanatornih i 250 kriterijskih varijabli.

```
*
*    *** COCAIN ***
*
* THIS PROGRAM IS WRITTEN IN 4.7/M VERSION
* OF SS LANGUAGE FOR UNIVAC 1100 COMPUTERS.
* A SIMILAR PROGRAM EXIST FOR CDC CYBER
* COMPUTERS, WRITTEN IN 3.0 VERSION OF SS.
*
* THE PROGRAM PERFORMES CANONICAL COR-
* RELATION ANALYSIS UNDER THE ALGORYTHM
* ORIGINALLY DESCRIBED IN MULAIK (1972), RE-
* VISED BY MOMIROVIĆ (1979). THE FINAL VER-
* SION OF THE ALGORYTHM IS DEVELOPED BY
* MOMIROVIĆ FOR RESEARCH AND TEACHING
* PURPOSES IN THE INSTITUTE FOR KINESIO-
* LOGY OF LJUBLJANA DURING THE LECTURES
* ON QUANTITATIVE METHODS FOR GRADUATE
* STUDENTS.
*
* THE PROGRAM ANALYSES THE CANONICAL RE-
* LATIONSHIPS OF THE TWO SETS OF VARIATES
* TRANSFORMED TO THE STANDARDISED PRIN-
* CIPAL COMPONENTS. THE NUMBER OF SIGNI-
* FICANT AND IMPORTANT CANONICAL VARIAB-
* LES IS DETERMINED BY MEIG CRITERION.

OUTPUT (DEVICE=PR3)
TEXT (TEXT=              COCAIN)
HEADING (TEXT= COCAIN — THE PROGRAM FOR THE CANONICAL CORRELATION ANALYSIS)
INPUT (DATA=DATA, SCORE=SC1)
REWIND (FILE=DATA)
INPUT (DATA=DATA, SCORE=SC2)
TEXT (TEXT=STATISTICS FIRST SET)
STATISTICS (SCORE=SC1, CLASS=9, Z=Z1, S)
TEXT (TEXT=STATISTICS SECOND SET)
STATISTICS (SCORE=SC2, CLASS=9, Z=Z2, S)
TEXT (TEXT=COMPOPNENTS FIRST SET)
CORRELATION (SCORE=Z1, R=R11)
PRINT (MATRIX=R11, TEXT=CORRELATIONS IN FIRST SET)
INVERSION (R=R11, RINV=R11I)
PRINT (MATRIX=R11I, TEXT=INVERSION OF R11)
DIAG (A=R11I, C=—1.0, D=U11)
MULT (A=U11, B=R11I, M=BIK)
MULT (A=BIK, B=U11, M=E11)
DELETE (MATRIX=BIK)
PRINT (MATRIX=E11, TEXT=ANTIIMAGE COVARIANCES IN FIRST SET, D)
DIAGONALISATION (R=R11, LAMBDA=L1, X=A1T, NAMEX=NMX1)
TRANSPOSE (OLD=A T, NEW=A1)
MULT (A=A1T, B=R11, M=BIK)
STATISTICS (SCORE=K1, S, CLASS=9, Z=LL1)
```

```
DELETE (MATRIX=L1)
MULTA (A=A1T, B=R11, M=BIK)
MULT (A=BIK, B=A1, M=L1P)
DELETE (MATRIX=BIK)
DIAG (A=L1P, D=L1)
DELETE (MATRIX=L1P)
DIAG (A=L1, D=L1P, C=0.5)
MULT (A=A1, B=L1P, M=H1)
PRINT (MATRIX=H1, TEXT=PRINCIPAL AXES IN FIRST SET)
TEXT (TEXT=COMPONENTS SECOND SET)
CORRELATION (SCORE=Z2, R=R22)
PRINT (MATRIX=R22, TEXT=CORRELATIONS IN SECOND SET)
INVERSION (R=R22, RINV=R22 I)
PRINT (MATRIX=R22 I, TEXT=INVERSION OF R22)
DIAG (A=R22 I, C=—1.0, D=U22)
MULT (A=U22, B=R22 I, M=BIK)
MULT (A=BIK, B=U22, M=E22)
PRINT (MATRIX=E22, TEXT=ANTI IMAGE COVARIANCES IN SECOND SET, D)
DELETE (MATRIX=SC2)
DELETE (MATRIX=BIK)
DELETE (MATRIX=U22)
DIAGONALISATION (R=R22, LAMBDA=L2, X=A2T, NAMEX=NMX2)
TRANSPOSE (OLD=A2T, NEW=A2)
MULT (A=Z2, B=A2, M=K2)
STATISTICS (SCORE=K2, S, CLASS=9, Z=LL2)
DELETE (MATRIX=L2)
MULT (A=A2T, B=R22, M=BIK)
MULT (A=BIK, B=A2, M=L2P)
DELETE (MATRIX=BIK)
DIAG (A=L2P, D=L2)
DELETE (MATRIX=L2P)
DIAG (A=L2, D=L2P, C=0.5)
MULT (A=A2, B=L2P, M=H2)
PRINT (MATRIX=H2, TEXT=PRINCIPAL AXES IN SECOND SET)
TEXT (TEXT=CANONICAL RELATIONS SET     SET2)
CROSSCORRELATION (P1=Z1, P2=Z2)
PRINT (MATRIX=R12, TEXT=CROSSCORRELATIONS AMONG VARIABLES)
CROSSCORRELATION (P1=LL1, P2=LL2, R12=G12)
PRINT (MATRIX=G12, TEXT=CROSSCORRELATIONS AMONG COMPONENTS)
TRANSPOSE (OLD=G12, NEW=G21)
MULT (A=G21, B=G12, M=G)
DIAGONALISATION (R=G, X=YT, LAMBDA=R02T)
TRANSPOSE (OLD=YT, NEW=Y)
MULT (A=YT, B=G, M=VOL)
MULT (A=VOL, B=Y, M=R02P)
DELETE (MATRIX=VOL)
DIAG (A=R02P, D=R02)
PRINT (MATRIX=R02, TEXT=ROOTS OF CANONICAL EQUATION)
DIAG (A=R02, C=0.5, D=RO)
DIAG (A=RO, C=—1.0, D=ROM)
MULT (A=G12, B=Y, M=SLON)
MULT (A=SLON, B=ROM, M=X)
PRINT (MATRIX=RO, TEXT=CANONICAL CORRELATIONS)
PRINT (MATRIX=X, TEXT=COMPONENT STRUCTURE IN SET1)
PRINT (MATRIX=Y, TEXT=COMPONENT STRUCTURE IN SET2)
MULT (A=H1, B=X, M=F1)
MULT (A=H2, B=Y, M=F2)
PRINT (MATRIX=F1, TEXT=CANONICAL FACTORS IN SET1)
PRINT (MATRIX=F2, TEXT=CANONICAL FACTORS IN SET2)
DIAG (A=L1P, C=—1.0, D=L1MP)
DIAG (A=L2P, C=—1.0, D=L2MP)
```

```
MULT (A=A1, B=L1MP, M=GAMA1)
MULT (A=A2, B=L2MP, M=GAMA2)
MULT (A=GAMA1, B=X, M=BETA1)
MULT (A=GAMA2, B=Y, M=BETA2)
PRINT (MATRIX=BETA1, TEXT=CANONICAL COEFFICIENTS IN SET1)
PRINT (MATRIX=BETA2, TEXT=CANONICAL COEFFICIENTS IN SET2)
TRANSPOSE (OLD=R12, NEW=R21)
MULR (A=R12, B=BETA2, M=FI1)
MULT (A=R21, B=BETA1, M=FI2)
PRINT (MATRIX=FI1, TEXT=CROSS STRUCTURE IN SET1)
PRINT (MATRIX=FI1, TEXT=CROSS STRUCTURE IN SET2)
MULT (A=Z1, B=BETA1, M=C1)
MULT (A=Z2, B=BETA2, M=C2)
STATISTICS (SCORE=C1, S, CLASS=9, Z=ZC1)
STATISTICS (SCORE=C2, S, CLASS=9, Z=ZC2)
TEXT (TEXT=SIMPLE      STRUCTURE ANALYSIS)
HOTELLING (X=YT, LAMBDA=R02T, F=HAHA, NAMEF=KRAVA)
TRANSPOSE (OLD=HAHA, NEW=H2R)
MULT (A=HAHA, B=H2R, M=LUL)
DIAG (A=LUL, C=—0.5, D=LULA)
MULT (A=H2R, B=LULA, M=YR)
TRANSPOSE (OLD=YR, NEW=YRT)
MULT (A=YRT, B=G, M=VOLR)
MULT (A=VOLR, B=YR, M=R02R)
DIAG (A=R02R, C=—0.5, D=RORR)
MULT (A=G12, B=YR, M=SLONR)
MULT (A=SLONR, B=RORR, M=XR)
MULT (A=H1, B=XR, M=F1R)
MULT (A=H2, B=YR, M=F2R)
DIAG (A=R02R, C=0.5, D=ROR)
TRANSPOSE (OLD=F1R, NEW=F1T)
TRANSPOSE (OLD=F2R, NEW=F2T)
VARIMAX (F=F1T, FN=V1T, TAU=TAU1, NAMEF=KOZA1)
VARIMAX (F=F2T, FN=V2T, TAU=TAU2, NAMEF=KOZA2)
TRANSPOSE (OLD=V1T, NEW=V1)
TRANSPOSE (OLD=V2T, NEW=V2)
PRINT (MATRIX=TAU1, TEXT=TRANSFORMATION, MATRIX IN SET1, T)
PRINT (MATRIX=V1, TEXT=SIMPLE STRUCTURE IN SET1)
PRINT (MATRIX=TAU2, TEXT=TRANSFORMATION MATRIX IN SET2, T)
PRINT (MATRIX=V2, TEXT=SIMPLE STRUCTURE IN SET2,)
MULT (A=R11I, B=V1, M=Q1)
MULT (A=R22I, B=V2, M=Q2)
PRINT (MATRIX=Q1, TEXT=TRANSFORMED COEFFICIENTS IN SET1)
MULT (A=Z1, B=Q1, M=VC1)
STATISTICS (SCORE=VC1, S, CLASS=9, Z=ZVC1)
PRINT (MATRIX=Q2, TEXT=TRANSFORMED COEFFICIENTS IN SET2)
MULT (A=Z2, B=Q2, M=VC2)
STATISTICS (SCORE=VC2, S, CLASS=9, Z=ZVC2)
TRANSPOSE (OLD=TAU2, NEW=TAU2T)
MULT (A=TAU1, B=ROR, M=ZVIZ)
MULT (A=ZVIZ, B=TAU2T, M=ROT)
PRINT (MATRIX=ROT, TEXT=CORRELATIONS OF TRANSFORMED CANVARIABLES)
HEADING (TEXT=END OF COCAIN)
```

LITERATURA

1. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, Wiley, New York, 1958.
2. Bartlett, M. S., The General Canonical Distribution, Ann. Math. Statist. 18, 1—17, 1949.
3. Cooley, W. W. and P. R. Lohnes, Multivariate Data Analysis, Wiley, New York, 1971.
4. Hotelling, H., Relations Between Two Sets of Variates, Biometrika 28, 321—377.
5. Kendall, M. G. and A. Stuart, The Advanced Theory of Statistics, Charles Griffin, London, 1973.
6. Morrison, D. F., Multivariate Statistical Methods, McGraw-Hill, New York, 1967.
7. Mulaik, S. A., The Foundations of Factor Analysis, McGraw-Hill, New York, 1972.
8. Rao, C. R., Linear Statistical Inference and Its Applications, Wiley, New York, 1973.

COCAIN — ALGORITHM AND PROGRAM FOR CANONICAL CORRELATION ANALYSIS

An algorithm for canonical analysis is poposed and a program is writen. The canonical dimensions for two sets of variables first transformed to the sandardized principal components are rotated by an partimonious ortogonal transformation.

»COCAIH« — АЛГОРИТМ И ПРОГРММА ДЛЯ КАНОНИЧЕСКОГО КОРРЕЛЯЦИОННОГО АНАЛИЗА

Определен алгоритм и написана программа для канонического корреляционного анализа, к которой добавлена ортогональная парсиномическая трансфомация канонических измерений двух групп переменных, которые были вначале трансформированы в стандартизованные главные компоненты.