

INDIFG – ALGORITAM I PROGRAM ZA SPEKTRALNU ANALIZU PROMJENA STANJA JEDNOG OBJEKTA

Franjo Prot
Konstantin Momirović
Ksenija Bosnar

Fakultet za fizičku kulturu
Sveučilišta u Zagrebu

Originalni znanstveni članak
UDK: 519.2

SAŽETAK

Definiran je algoritam i napisan makro program INDIFG koji analizira matricu trajektorija promjena jednog skupa varijabli na skupu vremenskih točaka koja je dobivena operacijom sukcesivnih mjerenja na jednom objektu istraživanja. Analiza matrice podataka izvedena je spektralnom dekompozicijom. Taksonomska analiza projekcija vremenskih točaka na lijeve svojstvene vektore, i taksonomska analiza projekcija varijabli na desne svojstvene vektore izvedena je ortonormalnim transformacijama maksimiziranjem Kaiserove varimax funkcije.

0. UVOD

Registracija i analiza promjena većeg broja karakteristika jednog objekta istraživanja u nizu ekvidistantnih vremenskih točaka neizbježna je procedura pri praćenju spontanijih procesa, kao i procesa koji su posljedica sustavno apliciranih tretmana u kineziologiji, psihologiji, medicini, defektologiji, i mnogim drugim znanostima i strukama.

Klasični radovi primjene faktorskih tehnika u analizi trendova (Tucker, 1962; Cattell, Choan, 1966; Momirović, 1972; Karaman i Momirović, 1982; Pavičić, Karaman i Momirović, 1985; Choutka, 1985) pokazali su sve nedostatke univarijatnog tretmana analize promjena, i ukazali na nužnost multivarijatnog pristupa.

Predloženi algoritam analizira matricu trendova spektralnom dekompozicijom, nakon čega izvodi taksonomizaciju vremenskih točaka ortonormalnom transformacijom lijevih svojstvenih vektora matrice po-

dataka i taksonomizaciju varijabli ortonormalnom transformacijom desnih svojstvenih vektora matrice podataka. Ovim postupkom integrirani su pristupi analizi promjena ranije rješavani separatnim dokompozicijama matrice sličnosti varijabli i matrice sličnosti vremenskih točaka (O i P analiza, Catell, 1966).

Ovakav analitički pristup osigurava simultano praćenje i prepoznavanje taksona točaka mjerenja i taksona varijabli što omogućava neposredno povezivanje sadržaja apliciranih tretmana sa efektima na trendove kontrolnih mjera.

1. ALGORITAM INDIFG

Neka je $V = \{v_j; j=1, \dots, m\}$ skup kvantitativnih varijabli, neka $T = \{t_i; i=1, \dots, n\}$ skup ekvidistantnih uređenih vremenskih točaka i neka je e objekt istraživanja, pripradnih neke populacije P.

Operacija

$$P^* \leftarrow T \otimes V$$

formira matricu

$$P^* = (p_{ij}^*) \quad \begin{matrix} i = 1, \dots, n \\ j = 1, \dots, m \end{matrix}$$

kao matricu trajektorija promjena m varijabli u n vremenskih točaka na objektu istraživanja e .

Neka je $\mu = (\mu_{ij}); i=1, j=1, \dots, m$ rezultat u m varijabli u prvoj vremenskoj točki (odnosno prvi redak matrice P^*), $i \phi = (\phi_i); i=1, \dots, n, \phi_i=1.0$ sumacioni vektor od n elemenata.

$$\text{Oparacija} \quad P = P^* - \phi \mu T$$

centrira matricu podataka na vrijednosti rezultata prvog mjerenja.

Spektralnom dekompozicijom matrice

$$P = \sum_{g=1}^m Y_g X_g^T \delta_g$$

određeni su generatori trajektorija promjena koje je sukladno doprinosu moguće reducirati na željeni broj prema unaprijed određenom kriteriju. Kriterij mogu biti:

- (1) $k =$ unaprijed zadani broj
- (2) $k = m$
- (3) $k = \text{num} \left(\delta_g \leq \sum_{g=1}^m \delta_g \frac{1}{m} \right)$

Neka se formira

$$\Delta = (\delta_g) \quad g = 1, \dots, k$$

matrica prvih k svojstvenih vrijednosti;
 $\delta_g \geq \delta_{g+1}$

i matrica $Y = (y_g)$

njima pridruženih lijevih svojstvenih vektora matrica podataka P i matrica

$$X = (x_g)$$

prvih k desnih svojstvenih vektora matrice podataka P .

Parsimonijska transformacija lijevih svojstvenih vektora matrice podataka izvodi se u cilju detekcije srodnih vremenskih točaka mjerenja uz zadržavanje ortogonalnosti generatora i udovoljenje Kaiserove varimax funkcije (Kaiser, 1958)

$$Y W = T \quad \left| \begin{array}{l} n \sum_{i=1}^n \sum_{g=1}^k f_{ig}^4 - \sum_{g=1}^k \left(\sum_{i=1}^n t_{ig}^2 \right)^2 = \max \\ W^T W = W W^T = I. \end{array} \right.$$

Matrica T predstavlja taksone skupa točaka mjerenja, a u matrici W su kosinusi direkcionalnih taksonomskih vektora u odnosu na lijeve vektore matrice trajektorija.

Ortonormalnom transformacijom desnih svojstvenih vektora matrice podataka određuju se taksoni varijabli uz maksimizaciju varimax funkcije i uvjet da i ovaj sistem bude ortogonalan.

$$X Q = F \quad \left| \begin{array}{l} m \sum_{j=1}^m \sum_{g=1}^k f_{jp}^4 - \sum_{g=1}^k \left(\sum_{j=1}^m f_{jq}^2 \right)^2 = \max \\ Q^T Q = Q Q^T = I. \end{array} \right.$$

U matrici F nalaze se koordinate u ortogonalnom taksonomskom sustavu, a u matrici Q kosinusi kuteva između desnih vektora matrice trajektorija i vektora koji određuju taksone varijabli.

2. PROGRAM INDIFG

Program INDIFG napisan je u programskom jeziku GENSTAT, verzija 4.04 a implementiran je i dostupan u javnoj programskoj biblioteci SRCE*GENS-MACRO.

Program je podijeljen u nekoliko sekcija:

- (1) deklaracije,
- (2) centriranje i normiranje matrice podataka,
- (3) spektralna dekompozicija matrice podataka,
- (4) određivanje broja značajnih dimenzija,
- (5) parsimonijske transformacije lijevih i desnih svojstvenih vektora,
- (6) ispis rezultata.

Za aktiviranje makro programa INDIFG korisnik mora iz glavnog programa prenijeti u makro program sljedeće:

- (1) VARIATE strukturu VARSET sa vremenskim točkama u kolonskim vektorima i vrijednostima obilježja u rečima,
- (2) dva POINTER vektora, RNAME s imenima obilježja i CNAM s identifikatorima vremenskih točaka,
- (3) vrijednost za skalar na kojem se određuje kriterij za broj svojstvenih vektora koji će se podvrći transformaciji,
- (4) vrijednosti za skalare, SVC i SVV, kojima se signalizira zahtjev za pohranjivanjem matrica svojstvenih vektora i njihovih transformacija.

Program INDIFG mora se pozvati s USE/R opcijom.

3. NUMERIČKI PRIMJER

PROGRAMOM INDIFG analizirani su prosjeci deset najboljih rezultata muškaraca u 18 atletskih disciplina ostvarenih u periodu 1951 – 1976 godine¹. Analiziran je trend razvoja rezultata ovih disciplina: trčanje na 100m, 200m, 400m, 800m, 1500m, 5000m, 10000m, trčanje s preponama na 110m i 400m, trčanje na 3000m s preprekama, skok uvis, skok s motkom, skok udalj, troskok, bacanje kugle, diska, kladiva i koplja.

Rezultati u disciplinama su centrirani na vrijednosti za početnu 1951. godinu i normirani na jediničnu dužinu, čime su poništeni efekti razlika u inicijalnoj metrici (centimetri i sekunde).

Spektralna dekompozicija centrirane i normirane matrice podataka dala je dvije natprosječne vrijednosti spektra (tabela

- 1). Apsolutne i relativne veličine svojstvenih vrijednosti pokazuju da je prve detektirana krivulja osnovna, a druga ima znatno manji doprinos u objašnjenju promjena rezultata (tabela 2).

Tabela 1

Svojstvene vrijednosti matrice podataka (LK),
relativne veličine svojstvenih vrijednosti (LKR)
i prosjek svojstvenih vrijednosti (MELD)

	LK	LKR
1	4.2071	0.7107
2	0.3781	0.0639
MELD = 0.3289		

Tabela 2

Lijevi svojstveni vektori matrice podataka

	1	2
G1951	0.0000	-0.0000
G1952	0.0311	0.1189
G1953	0.0263	0.3282
G1954	0.0432	0.3713
G1955	0.0761	0.3892
G1956	0.1161	0.1895
G1957	0.1018	0.3009
G1958	0.1197	0.2836
G1959	0.1259	0.1937
G1960	0.1646	0.1751
G1961	0.1427	0.2327
G1962	0.1617	0.1472
G1963	0.1650	0.1484
G1964	0.2001	0.0076
G1965	0.1872	0.0425
G1966	0.2104	0.0802
G1967	0.2199	-0.1523
G1968	0.2662	-0.2961
G1969	0.2386	-0.1185
G1970	0.2455	-0.0723
G1971	0.2622	-0.1224
G1972	0.2851	-0.1792
G1973	0.2712	-0.0056
G1974	0.2777	0.0329
G1975	0.2925	-0.1330
G1976	0.3062	-0.1072

¹ Podaci su preuzeti iz rada Važnog (1978), str. 22–23.

Prvi lijevi svojstveni vektor matrice podataka pokazuje kontinuirani porast rezultata uz neznatne oscilacije objašnjive olimpijskim ciklusima. Drugi vektor diferencira godine do 1966 od godina nakon 1966. Najveću (negativnu) projekciju ima 1968. godina. Drugi vektor očito ocrta prelom u razvoju atletike, jer se nakon parcijalizacije općeg pozitivnog trenda počinje manifestirati limit u veličinama prirasta rezultata.

Taksonomizacija vremenskih točaka (tabela 3) rezultirala je grupiranjem godina prema njihovom položaju na generatorima krivulja promjena. Prva taksonomska dimenzija definira dva perioda razvoja atletike, uz male oscilacije u vrijednostima olimpijskih, a u novije vrijeme, predolimpijskih godina. Druga taksonomska dimenzija naglašeno izdvaja 1968. i nekoliko godina u toj fazi, od ranijih godina razvoja atletike.

Prvi desni svojstveni vektor matrice podataka (tabela 4) sadrži gotovo identične vrijednosti za sve discipline i dokazuje da je opravdano promatrati razvoj atletike kao cjeline. Drugi vektor razlikuje trčanja na kraćim dionicama, skok s motkom, i bacanje kugle i diska od trčanja na dužim dionicama i bacanja koplja, tj. razlikuje discipline sa manjom relativnom promjenom od onih s većom, nakon što je parcijalizirana osnovna varijanca promjena.

Taksonomizacija disciplina s obzirom na krivulje promjena dala je dimenzije koje se mogu interpretirati veličinama relativnih promjena u disciplinama. Prvi taksonomsku dimenziju definiraju discipline s većim relativnim prirastom, a drugu dimenziju discipline s manjim relativnim prirastom u periodu od 1951 – 1976. godine.

Tabela 3

	Sklop taksona vremenskih točaka	
	1	2
G1951	0.0000	-0.0000
G1952	-0.0184	0.1215
G1953	-0.1055	0.3119
G1954	-0.1070	0.3582
G1955	-0.0838	0.3876
G1956	0.0318	0.2199
G1957	-0.0254	0.3167
G1958	-0.0021	0.3079
G1959	0.0391	0.2277
G1960	0.0821	0.2259
G1961	0.0392	0.2702
G1962	0.0904	0.1991
G1963	0.0930	0.2015
G1964	0.1808	0.0861
G1965	0.1552	0.1129
G1966	0.1616	0.1568
G1967	0.2622	-0.0531
G1968	0.3615	-0.1668
G1969	0.2660	-0.0146
G1970	0.2541	0.0305
G1971	0.2893	-0.0089
G1972	0.3327	-0.0520
G1973	0.2513	0.1020
G1974	0.2421	0.1399
G1975	0.3213	-0.0066
G1976	0.3236	0.0224

TRANSFORMACIJSKA MATRICA

	1	2
1	0.9187	0.3951
2	-0.3951	0.9187

Tabela 4

Desni svojstveni vektori matrice podataka

	1	2
A100M*	-0.2333	0.3598
A200M*	-0.2354	0.2521
A400M*	-0.2356	0.1549
A800M*	-0.2349	-0.3663
A1500M*	-0.2349	-0.3576
A5000M*	-0.2365	-0.1629
A10000M*	-0.2369	-0.0751
A110MP*	-0.2339	0.3061
A400MP*	-0.2361	-0.2089
A3000MP*	-0.2359	-0.2920
ASKOKVIS	0.2367	-0.0065
ASKOKMOT	0.2330	-0.4259
ASKOKDAL	0.2365	-0.0655
ATROSKOK	0.2369	0.0486
KUGLA	0.2368	-0.1115
DISK	0.2365	-0.1264
KLADIVO	0.2370	0.0788
KOPLJE	0.2360	0.2085

Tabela 5

Sklop taksona atletskih disciplina

	1	2
A100M	0.0849	0.4204
A200M	0.0081	0.3449
A400M	-0.0600	0.2755
A800M	-0.4240	-0.0975
A1500M	-0.4180	-0.0913
A5000M	-0.2830	0.0490
A10000M	-0.2218	0.1120
A110MP	0.0469	0.3823
A400MP	-0.3149	0.0158
A3000MP	-0.3728	-0.0437
ASKOKVIS	0.1646	-0.1702
ASKOKMOT	-0.1314	-0.4674
ASKOKDAL	0.1233	-0.2122
ATROSKOK	0.2033	-0.1310
KUGLA	0.0913	-0.2453
DISK	0.0806	-0.2558
KLADIVO	0.2245	-0.1094
KOPLJE	0.3145	-0.0160

TRANSFORMACIJSKA MATRICA

	1	2
1	0.7147	-0.6994
2	0.6994	0.7147

* Kod disciplina označenih zvjezdicom rezultat se izražava vremenskim jedinicama gdje je manji broj bolji rezultat

Ovaj numerički primjer izveden je jednim 'REFERENCE' programom INDIFG. Ovaj program ima zadatak da pročita makro naredbu INDIFG § iz systemske biblioteke makro programa, da postavi vrijednosti ulaznih parametara:

- pointer RNAM = imena varijabli
- pointer CNAM = imena vremenskih točaka
- skalar NC = izbor kriterija redukcije
- skalari SVC i SVV = izbor procedura pohrane rezultata, te da nakon učitava-

vanja podataka 'USE/R' naredbom GENSTAT jezika aktivira makro program INDIFG §.

'REFERENCE' programu su pridružene i neophodne naredbe operacionog sistema OS1100/37R2D računala UNIVAC 1100/42. Ovaj skup naredbi pridružuje systemsku biblioteku makro programa (SRCE*GENS-MACRO), systemsku datoteku iz koje aktivira GENSTAT programski sistem (GEN404*USERFILE) i korisničku datoteku sa podacima (FFK*WORK).

```

∂ASG,A SRCE*GENS—MACRO.                G1954,D1955,G1956,G1957,G1958,
∂USE 27., SRCE*GENS—MACRO.                G1959,G1960,G1961,G1962,G1963,
∂ASG,A GEN404*USERFILE                    G1964,G1965,G1966,G1967,G1968,
∂GEN404*USERFILE.GENSTAT,TIME            G1969,G1970,G1971,G1972,G1973,
    „50                                     G1974,G1975,G1976
‘REFERENCE(NID=100,NUNN=100’              ‘SCAL’ NC = 2
    INDIFG                                  ‘SCAL’ SVC = 0.0
‘GET/FILE=3’ INDIFG § INDIFG             ‘SCAL’ SVV = 0.0
”                                           ‘SET/POIN=S’ VARSET = CNAM
    ATLETIKA 1951 – 1976                    ‘R’
”                                           ‘READ/P,NUN=v,PRINT=Z’ VARSET §§
‘POIN’ RNAM=A100M,A200M,A400M,           ,7,/,7,/,7,/,5,/
    A800M,A1500M,A5000M,A10000M,          ‘RUN’
    A110MP,A400MP,A3000MP,                ∂ADD FFK*WORK.PROMJENE/INDIF
    ASKOKVIS,ASKOKMOT,                     ‘EOD’
    ASKOKDAL,ATROSKOK,KUGLA,              ‘USE/R,PRINT=Y’ INDIFG §
    DISK,KLADIVO,KOPLJE                    ‘R’
‘POIN’ ONAM=G1951,G1952,G1953,           ‘CLOSE’
                                           ‘STOP’

```

LITERATURA

1. CATTELL, R.B. (1966). Handbook of multivariate experimental psychology. Rand McNally, Chicago.
2. MOMIROVIĆ, K. i Ž. KARAMAN (1982). INDIFG: Model, algoritam i program za analizu promjena stanja nekog objekta opisanog nad skupom kvantitativnih varijabli. Kineziologija, 13, 2:5–8.
3. MOMIROVIĆ, K., F. PROT, D. DUGIĆ, Z. KNEZOVIĆ, K. BOSNAR, N. ERJAVEC, M. GREDELJ, J. KERN, V. DOBRIĆ i J. RADAKOVIĆ (1986). Metode, algoritmi i programi za analizu kvalitativnih i kvantitativnih promjena. Institut za kineziologiju, Zagreb (elaborat)

INDIFG – AN ALGORITHM AND A PROGRAM FOR SPECTRAL ANALYSIS OF STATE CHANGES OF ONE OBJECT

Summary

An algorithm is defined and a macro program (INDIFG) able to analyse matrix of trajectories of changes of one set of variables on a set of time points obtained by successive measures on one investigated object. Analysis of the data matrix is performed with spectral decomposition. Taxonomic analysis of time points projections on the left characteristic vectors and taxonomic analysis of variables projections on the right characteristics vectors is performed with orthonormal transformations maximizing Kaiser's varimax function.