JÓZSA SÁNDOR

A diád-módszer statisztikai expozíciója

Igen hasznosnak bizonyult Székely Béla [8] dolgozatában kifejtett diád-módszere a gazdasági fejlődés közös trendjeinek vizsgálata céljára (l. Rimler Judit—Székely Béla [7] dolgozatát). A módszer alapgondolata algebrai, amit jelen dolgozatomban a feladat enyhe módosításával statisztikai alapgondolattal helyettesítek. Az így módosított feladat megoldása a legkisebb négyzetek módszere helyett a legnagyobb valószínűség (Maximum Likelihood) elvén alapul. Az eljárás gyakorlati alkalmazására példát mutatok.

1. Bevezetés

A diád-módszer azon $u_i^{(l)}$, $v_j^{(l)}$ (valós) mennyiségek meghatározására ad eljárást, melyekre valamely x_{ij} $(i=1, 2, \ldots, n;\ j=1, 2, \ldots, m)$ adott (valós) mennyiségekre az alábbi rekurzív feltétel teljesül:

$$\begin{aligned} x_{ij} &= \hat{u}_i^{(1)} \hat{v}_j^{(1)} + r_{ij}^{(1)}, \quad \sum_{i,j} (r_{ij}^{(1)})^2 = \text{minimális}, \\ r_{ij}^{(1)} &= \hat{u}_i^{(2)} \hat{v}_j^{(2)} + r_{ij}^{(2)}, \quad \sum_{i,j} (r_{ij}^{(2)})^2 = \text{minimális}, \\ &\ldots\ldots\ldots\ldots\ldots \\ r_{ij}^{(k-1)} &= \hat{u}_i^{(k)} \hat{v}_j^{(k)} + r_{ij}^{(k)}, \quad \sum_{i,j} (r_{ij}^{(k)})^2 = \text{minimális}, \\ &(i=1, 2, \ldots, n;\ j=1, 2, \ldots, m). \end{aligned} \tag{1}$$

Az x_{ij} mennyiségek ilyen értelmű felbontásának megkísérlését bizonyos közgazdasági meggondolások vetik fel. k, a rekurzió lépésszáma úgy választandó, hogy a $\sum_{i,j} (r_{ij}^{(k)})^2$ összeg megfelelően kicsiny legyen. (Itt a $\wedge$ jel használatát a későbbiek indokolják.)

Természetszerűen vetődik fel a kérdés, hogy az (1) feladat megoldásai nem elégítik-e ki egyben az

$$x_{ij} = \hat{u}_i^{(1)} \hat{v}_j^{(1)} + \hat{u}_i^{(2)} \hat{v}_j^{(2)} + \ldots + \hat{u}_i^{(k)} \hat{v}_j^{(k)} + r_{ij}, \quad \sum_{i,j} r_{ij}^2 = \text{minimális}, \tag{1a}$$
$$(i=1, 2, \ldots, n;\ j=1, 2, \ldots, m)$$

feltételt is. Később látni fogjuk, hogy a válasz igenlő. Dolgozatom célja az (1a) feladat statisztikai szemléletű módosítása és megoldása egy speciális esetre.

Az (1a) algebrai feladatot statisztikaivá alakíthatjuk a következő módon. Tekintsük x_{ij}-t bizonyos ξ_{ij} valószínűségi változó realizációjának (azaz statisztikai értelemben mintának), továbbá tegyük fel, hogy ξ_{ij} várhatóértéke $(\mathcal{E}\xi_{ij})$ az alábbi alakban írható:

$$\mathcal{E}\xi_{ij} = u_i^{(1)} v_j^{(1)} + u_i^{(2)} v_j^{(2)} + \ldots + u_i^{(k)} v_j^{(k)} \tag{2}$$

$$(i = 1, 2, \ldots, n; \; j = 1, 2, \ldots, m).$$

Az x_{ij} realizációkból a ξ_{ij} valószínűségi változók együttes sűrűségfüggvénye típusának ismeretében a Maximum Likelihood elv alapján [1] szándékozunk becsülni az $u_i^{(l)}$, $v_j^{(l)}$ paramétereket, pontosabban a (2) jobboldalán álló összeget — hiszen bármely mátrix többféleképpen bontható diádok összegére [4]. A legkisebb négyzetek elvének alkalmazása itt (általában) nem indokolt, gondoljunk pl. arra az esetre, amikor a ξ_{ij} változók különböző szórásúak, vagy arra, amikor nem függetlenek.

A feladatot alább arra az esetre fogalmazzuk meg, amikoris a ξ_{ij} *valószínűségi változók együttes eloszlása (több dimenziós) normális eloszlás*. A 3. pontban egy további speciális eloszlás mellett elő is állítjuk a (2) jobboldalának becslését.

2. A feladat megfogalmazása normális együttes eloszlás esetén

A ξ_{ij} valószínűségi változók sűrűségfüggvényének felírásához képezzük a $\Xi = (\xi_{ij})$ $n \cdot m$-es valószínűségi mátrixot. Ξ oszlopvektorainak kiterítésével nyerjük az $n \cdot m$ komponensű ξ valószínűségi vektorváltozót. Tegyük fel, hogy ξ kovariancia-mátrixa S (tehát $S = \mathcal{E}(\xi - \mathcal{E}\xi)(\xi - \mathcal{E}\xi)'$), ismert $nm \cdot nm$-es nem szinguláris mátrix. Jelölje Q az S mátrix inverzét. ξ sűrűségfüggvényében a változókat (megkülönböztetésül az x_{ij} realizációktól) félkövér $\mathbf{x}_{ij}$-vel jelöljük, az ezekből összeállított nm komponensű vektorváltozót $\mathbf{x}$-szel.

Feltesszük, hogy ξ normális eloszlást követ, azaz sűrűségfüggvénye a bevezetett jelölésekkel:

$$f(\mathbf{x}) = c \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathcal{E}\xi)'Q(\mathbf{x} - \mathcal{E}\xi)\right\},$$

(a $'$ jel transzponálást jelöl e dolgozatban), ahol c ismert konstans

$$(c = 1/\sqrt{(2\pi)^{nm} \det S}).$$

(2)-re tekintettel $f(\mathbf{x})$ tartalmazza az ismeretlen u, v, paramétereket. Ha most $\mathbf{x}$ helyébe az x_{ij} realizációkból képzett x vektort helyettesítjük, $f(x)$ már csak a keresett, u, v paraméterektől függ. A Maximum Likelihood elv alkalmazása abban áll, hogy (2) jobboldalának becsléséhez olyan $\hat{u}$, $\hat{v}$ értékeket keresünk, melyek mellett $f(x)$ maximális.

$f(x)$ maximumát ott veszi fel, ahol exponense maximális, azaz

$$\mathfrak{L}(u, v) = (x - \mathcal{E}\xi)'\,Q(x - \mathcal{E}\xi)$$

minimális. E kifejezés kezelhetőbb alakba való átírása céljából jelöljük az $X = (x_{ij})$ mátrix j-ik oszlopvektorát x_j-vel ($j = 1, 2, \ldots, m$), továbbá legyenek

$$u_l = (u_1^{(l)}, u_2^{(l)}, \ldots, u_n^{(l)})'$$
$$v_l = (v_1^{(l)}, v_2^{(l)}, \ldots, v_m^{(l)})' \qquad (l = 1, 2, \ldots, k)$$
$$w_j = (v_j^{(1)}, v_j^{(2)}, \ldots, v_j^{(k)})' \qquad (j = 1, 2, \ldots, m)$$

paraméter-vektorok, és

$$U = (u_1, u_2, \ldots, u_k)$$
$$V = (v_1, v_2, \ldots, v_k) = (w_1, w_2, \ldots, w_m)'$$

a paraméterekből összeállítot $n \times k$, ill. $m \times k$ típusú mátrixok. Ezekkel, tekintettel (2)-re

$$\mathcal{E}\Xi = UV' = U(w_1, w_2, \ldots, w_m) = (Uw_1, Uw_2, \ldots, Uw_m),$$

$\mathcal{E}\xi$ pedig az utóbbi mátrix oszlopaiból kiterített vektor. Végül particionáljuk a Q mátrixot $n \times n$-es blokkokra:

$$Q = \begin{pmatrix} Q_{11} & Q_{12} & \ldots, & Q_{1m} \\ Q_{21} & Q_{22} & \ldots, & Q_{2m} \\ \cdot & \cdot & \cdot & \cdot \\ Q_{m1} & Q_{m2} & \ldots, & Q_{mm} \end{pmatrix}$$

Ezekután a fenti $\mathfrak{L}$ kifejezés az alábbi alakot ölti:

$$\sum_{j=1}^{m} \sum_{j^*=1}^{m} (x_j - Uw_j)' Q_{jj^*} (x_{j^*} - Uw_{j^*}). \tag{3}$$

Tekintsük most e kifejezést az UV' mátrix elemei függvényének, azaz ebben a szemléletben UV' *nem a* Ξ *valószínűségi mátrix tényleges várható-értéke, hanem tetszőleges valós elemű* U, V' *mátrixok szorzata.*

A feladat tehát UV' olyan becslésének megkeresése, mely(ek)nél (3) minimális. Ebben a még mindig eléggé általános esetben (3)-ból olyan egyenletrendszer adódott az $\hat{U}$, $\hat{V}$ becslésekre, melynek explicit megoldását nem sikerült találnom. A kapott egyenletrendszer megoldását illetően iterációs eljárás látszik célravezetőnek.

Megjegyzés: Ha ξ degenerált eloszlású, akkor S szinguláris, nem invertálható. Ilyenkor (3)-ban a Q mátrix szerepét S-nek a Moore—Penrose-féle általánosított inverze [3] tölti be. Legyen ugyanis S rangja $p < n \cdot m$. Ekkor ξ sűrűségfüggvénye egy az $\mathbf{x} = \mathcal{E}\xi$ ponton átmenő p-dimenziós H_p hipersíkon kívül eltűnik, magán H_p-n pedig egy nem degenerált eloszlású p-dimenziós ξ^* valószínűségi vektorváltozó sűrűségfüggvényével adható meg. $\varrho(S) = p$ ugyanis azt jelenti [1], hogy létezik ξ^*, p-dimenziós nem degenerált normális eloszlású valószínűségi vektorváltozó és A, $nm \times p$-es mátrix úgy, hogy $\xi = A\xi^*$, $A'A = E$ ($p \times p$-es egység mátrix). Ebből $\xi^* = A'\xi$, $\mathcal{E}\xi^* = A'\mathcal{E}\xi$ és ξ^* kovariancia-mátrixa $S^* = A'SA$, nem szinguláris, inverze legyen Q^*. A ξ^* sűrűségfüggvénye legyen $f^*(\mathbf{x}^*)$ ahol $\mathbf{x}^* = A'\mathbf{x}$ ($\mathbf{x} \in H_p$). $\mathbf{x} \in H_p$-re ezekből következően fennáll:

$$f(\mathbf{x}) = f^*(\mathbf{x}^*) = c^* \exp\left\{-\frac{1}{2}(\mathbf{x}^* - \mathscr{E}\xi^*)' Q^*(\mathbf{x}^* - \mathscr{E}\xi^*)\right\} =$$

$$= c^* \exp\left\{-\frac{1}{2}(A'\mathbf{x} - A'\mathscr{E}\xi)'(A'SA)^{-1}(A'\mathbf{x} - A'\mathscr{E}\xi)\right\} =$$

$$= c^* \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathscr{E}\xi)' A(A'SA)^{-1}A'(\mathbf{x} - \mathscr{E}\xi)\right\}.$$

Az exponensben Q szerepét tehát az $A(A'SA)^{-1}A'$ mátrix tölti be, ami valóban az S mátrix Moore—Penrose-féle általánosított inverze [3].

3. UV′ becslése egy speciális kovariancia-mátrix esetén

A. 4. pont példájánál, ami konkrét gyakorlati igény kapcsán adódott, a Ξ mátrix oszlopvektorai egymástól független azonos eloszlásúak. Ebben az esetben az S kovariancia-mátrix és annak Q inverze is diagonál-hipermátrix, nevezetesen:

$$Q_{jj^*} = \begin{cases} Q_0, & \text{ha } j = j^* \\ 0, & \text{ha } j \neq j^*. \end{cases}$$

Q_0, mint kovariancia-mátrix, pozitív definit vagy pozitív szemidefinit [1]. A (3) kifejezés ekkor az alábbi módon egyszerűsödik:

$$\sum_{j=1}^{m} (x_j - Uw_j)' Q_0 (x_j - Uw_j). \tag{3a}$$

A feladat tehát olyan $\hat{U}\hat{V}'$ mátrix(ok) keresése, mely(ek)re (3a) minimális vagy — (3a)-t „nyom” formában írva:

$$\mathrm{Sp}(X - UV')' Q_0 (X - UV') = \text{minimális}. \tag{4}$$

E speciális feladat megoldását adja a következő

1. *TÉTEL. Legyen X $n \times m$-es valós mátrix, Q_0 $n \times n$-es valós, szimmetrikus, pozitív (szemi) definit s rangú mátrix. Jelöljön $Y_k (k = 1, 2, \ldots, s)$ tetszőleges (valós), legfeljebb k-ad rangú $n \times m$-es mátrixot és tekintsük az alábbi függvényt:*

$$d(Y_k) = \mathrm{Sp}\,(X - Y_k)' Q_0 (X - Y_k).$$

Áll: d minimumát adó (bármely) Y_k^* mátrixra

$$Y_k^* = M_k = \sqrt{\lambda_1} f_1 g_1' + \sqrt{\lambda_2} f_2 g_2' + \ldots + \sqrt{\lambda_k} f_k g_k', \quad (k = 1, 2, \ldots, s),$$

ahol $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s$ az $X'Q_0X$ mátrix pozitív sajátértékei, $g_1, g_2, \ldots, g_s$ (valamely) megfelelő ortonormált sajátvektor rendszere és $f_l = Xg_l/\sqrt{\lambda_l}$ $(l = 1, 2, \ldots, s)$. A minimumra továbbá:

$$\min d(Y_k) = d(M_k) = \mathrm{Sp}\, X' Q_0 X - \sum_{l=1}^{k} \lambda_l \quad (= \lambda_{k+1} + \ldots + \lambda_s),$$

$$(k = 1, 2, \ldots, s)$$

*

A tétel szövegében álló „valamely" szó arra utal, hogy ha a λ_l sajátértékek között többszörös multiplicitású is szerepel, akkor a sajátvektor rendszer többféleképpen is választható. Ilyenkor M_k függhet a választott sajátvektor rendszertől. Erre vonatkozóan a $Q_0 = E$ esetre [5]-ben igazoltam, hogy azokra a k-kra, melyekre $\lambda_k > \lambda_{k+1}$, továbbá $\lambda_n > 0$ esetén $k = n$-re, M_k független a választott sajátvektor rendszertől. Ennek bizonyítása itt hasonlóan végezhető. A tétel szerint viszont már $d(M_k)$ a $\{g_l\}$ rendszer bármely választása mellett $d(Y_k)$ minimumával egyenlő.

A tétel megadja UV' (összes) becslését, nevezetesen (tetszőleges $\{g_l\}$ rendszer mellett)

$$\hat{U}\hat{V}' = M_k$$

UV' (egy) becslése, (és mindegyik becslése ilyen alakban írható). Ebből nyerhetők U, ill. V speciális becslései:

$$\hat{U} = (\sqrt{\lambda_1} f_1, \sqrt{\lambda_2} f_2, \ldots, \sqrt{\lambda_k} f_k) = (f_1, f_2, \ldots, f_k)\langle\sqrt{\lambda_1}, \sqrt{\lambda_2}, \ldots, \sqrt{\lambda_k}\rangle = F_k \Lambda_k^{1/2},$$

(ahol $\langle\lambda_1, \lambda_2, \ldots, \lambda_k\rangle = \Lambda_k$ a feltüntetett elemekből álló diagonális mátrix), és

$$\hat{V} = (g_1, g_2, \ldots, g_k) = G_k$$

Érdemes $\hat{U}\hat{V}'$ ennek megfelelő faktorizációját felírni:

$$\hat{U}\hat{V}' = F_k \Lambda_k^{1/2} G_k'$$

Megjegyzések:

1. Q_0 pozitív (szemi) definit voltából következik, hogy az $X'Q_0X$ mátrix is pozitív (szemi) definit, tehát a tételben, nem említett sajátértékei eltűnnek. Mint ismeretes [4], ekkor $\operatorname{Sp} X'Q_0X = \lambda_1 + \lambda_2 + \ldots + \lambda_s$, ami egyben a tétel utolsó (zárójelben álló) egyenlőségét is igazolja. Speciálisan, $d(M_s) = 0$, úgyhogy a $k > s$ esetek érdektelenek.

2. Az f_l vektorokra fennáll:

$$Q_0XX'Q_0f_l = \lambda_l Q_0 f_l \text{ és } f_l'Q_0f_{l^*} = \begin{cases} 1, \text{ ha } l = l^* \\ 0, \text{ ha } l \neq l^*. \end{cases} \tag{5}$$

Innen, ha Q_0 invertálható, akkor λ_l az $XX'Q_0$ mátrixnak is sajátértéke, f_l a (egy) hozzátartozó sajátvektor. *Speciálisan, ha $Q_0 = E$, akkor λ_l az $X'X$ és egyben az XX' mátrix sajátértéke, g_l, ill. f_l sajátvektorral.* Ezekkel viszont M_k diádjai éppen a Székely Béla által leírt rekurziós eljárással nyerhetőkkel azonosak, azaz *(1) megoldásai kielégítik (1a)-t is.*

(5) első ill. második részének belátásához elég az $X'Q_0Xg_l = \lambda_l g_l$ (g_l definíciója!) egyenlőséget balról $Q_0X/\sqrt{\lambda_l}$-lel, ill. $g_{l^*}'/\sqrt{\lambda_{l^*}}$-gal szoroznunk.

3. Abból, hogy $Q_0^{-1} = S_0$ a Ξ mátrix oszlop-vektorainak közös kovarianciamátrixa, belátható, hogy a $d(M_s) = 0$ egyenlőségen túl $M_s = X$ is áll. A tétel állítását emiatt és 1. miatt a következőképpen interpretálhatjuk:

M_s az X mátrix olyan diád-előállítása, melyben az első diád maximális (százalékban kifejezve $100\lambda_1/\operatorname{Sp} X'Q_0X$ %) információt nyújt X-ről, a megmaradt információ maximális hányadát ($100\,\lambda_2/\operatorname{Sp} X'Q_0X$ %-ot) foglalja magába a második diád, és így tovább.

Az M_s diádösszeget éppen az a tény teszi kitüntetetté, hogy *egyes szeletei UV' becslését tetszőleges k mellett szolgáltatják.*

A tétel bizonyítása:

A tételt fogalmazhattuk volna úgy is, hogy csak a pontosan k-ad rangú Y_k mátrixokra szorítkozunk. Az így módosított állításból ugyanis következik a ténylegesen megadott állítás, hiszen akkor

$$\min_{\varrho(Y_k)=k^*<k} d(Y_k) = d(M_{k^*}) > d(M_k) = \min_{\varrho(Y_k)=k} d(Y_k).$$

Feltehetjük tehát, hogy $\varrho(Y_k) = k$. Ekkor Y_k felbontható k diád összegére [4]: $Y_k = UV'$, ahol U $n\times k$, V $m\times k$ típusú mátrix, (nem azonosak a Ξ várhatóértékében azonosan jelölt U, V mátrixokkal), mindkettő rangja k. Minthogy emiatt $V'V$ nem szinguláris, feltehetjük, hogy $V'V = E$. Ellenkező esetben ugyanis az $\tilde{U} = U(V'V)^{\frac{1}{2}}$, $\tilde{V} = V(V'V)^{-\frac{1}{2}}$ transzformációkkal olyan $\tilde{U}$, $\tilde{V}$ mátrixokat nyerünk, melyekkel egyrészt $\tilde{U}\tilde{V}' = UV'$, másrészt $\tilde{V}'\tilde{V} = E$ már teljesül. Utóbbi következik abból, hogy $V'V = (V'V)^{\frac{1}{2}}(V'V)^{\frac{1}{2}}$.

Egyszerű módon adódik, hogy

$$d(Y_k) = X'Q_0X - (2\,\mathrm{Sp}\, VU'Q_0X - \mathrm{Sp}\, VU'Q_0UV') \tag{6}$$

$d(Y_k)$ minimális, ha a zárójelben álló kifejezés (mondjuk b) maximális. Felhasználva a $\mathrm{Sp}\, AB = \mathrm{Sp}\, AB$ azonosságot [4] és a $V'V = E$ feltételt, írhatjuk:

$$b = 2\,\mathrm{Sp}\, U'Q_0XV - \mathrm{Sp}\, U'Q_0U.$$

Rögzített V mellett deriváljuk b-t rendre U' sorvektorai szerint, a deriváltakat tegyük egyenlővé 0-val. A nyom definíciójából [4] adódik, hogy

$$\frac{\partial b}{\partial u_l'} = \frac{\partial}{\partial u_l'}(2\,u_l'Q_0Xv_l - u_l'Q_0u_l) = 2\,Q_0Xv_l - 2\,Q_0u_l\,.$$

Ily módon, ugyanahhoz az egyenletrendszerhez jutunk, mint ha formálisan U' szerint deriváljuk b-t. Deriválás után a maximumot adó $\hat{U}_V$ mátrixokra az alábbi egyenletet kapjuk:

$$Q_0XV = Q_0\hat{U}_V \tag{7}$$

Ebből következik, hogy $\hat{U}_V'Q_0XV = \hat{U}_V'Q_0\hat{U}_V$, azaz b rögzített V melletti b_V maximális értékénél

$$b_V = \mathrm{Sp}\, \hat{U}_V'Q_0XV,$$

amiből (7) szerinti helyettesítéssel nyerjük:

$$b_V = \mathrm{Sp}\, V'X'Q_0XV = \sum_{l=1}^{k} v_l'X'Q_0Xv_l. \tag{8}$$

A b_V és egyben b maximumát szolgáltató $\hat{v}_l$ vektorokat kell még megkeresnünk. Ehhez, tekintettel a $v_l'v_l = 1$ feltételre, a $b_V + \sum\limits_{l=1}^{k}\mu_l(1 - v_l'v_l)$ kifejezés v_l' szerinti deriváltját kell 0-val egyenlővé tenni ($l = 1, 2, \ldots, k$), ahol a μ_l konstansok a Lagrange-féle multiplikátorok. Deriválással nyerjük:

$$X'Q_0X\,\hat{v}_l = \mu_l\hat{v}_l \quad (l = 1, 2, \ldots, k) \tag{9}$$

azaz μ_l az $X'Q_0X$ mátrix sajátértéke, $\hat{v}_l$ (egy) hozzátartozó sajátvektor. (9)-et balról v_l'-lel szorozva és (8)-ba helyettesítve, b maximális értékére a

$$b_{\max} = \sum_{l=1}^{k} \mu_l \tag{10}$$

összeget kapjuk. Tekintettel arra, hogy a $\hat{v}_l$ sajátvektoroknak ortogonálisaknak kell lenniök egymásra, (10) maximumát akkor veszi fel, ha $\mu_1, \mu_2, \ldots, \mu_k$ az $X'Q_0X$ mátrix k legnagyobb sajátértéke, azaz $\mu_l = \lambda_l$, amikoris $v_l = g_l$ írható ($l = 1, 2, \ldots, k$); tömören: $\hat{V} = G_k$.

$\hat{V}$ ismeretében a keresett $\hat{U}$ mátrixot úgy kell megválasztani, hogy (7) teljesüljön. (7) viszont teljesül, ha $\hat{U} = X\hat{V}$, azaz $\hat{U} = XG_k$. Ezekkel a $d(Y_k)$ függvényre minimumot adó $\hat{U}\hat{V}'$ mátrix(ok)ra:

$$\hat{U}\hat{V}' = XG_k G_k' = \sum_{l=1}^{k} Xg_l g_l' = \sum_{l=1}^{k} \sqrt{\lambda_l}\, f_l\, g_l' = M_k\,.$$

Végül, ha (6) zárójelbe tett kifejezése helyébe (10)-et tesszük, $\mu_l = \lambda_l$ helyettesítés után a tétel állítása bizonyítást nyert.

*

Tekintettel arra, hogy a legnagyobb sajátérték és az ahhoz tartozó sajátvektor megkeresésére jól használható iterációs eljárások ismeretesek, szerencsés volna, ha M_k egyes diádjait az (1) alattihoz hasonló rekurziós eljárással számíthatnánk. Erre vonatkozóan igaz a következő:

2. *TÉTEL. Legyen* $M_0 = 0$, M_{l-1} ($l = 1, 2, \ldots, s$) *az 1. tételben említett diádösszeg valamely* g_1, g_2, g_{l-1} *sajátvektor rendszer mellett. Ekkor az* $(X - M_{l-1})'$ $Q_0(X - M_{l-1})$ *mátrix legnagyobb sajátértéke az 1. tételben definiált* λ_l, *tetszőleges hozzátartozó* g_l^* *normált sajátvektora pedig választható* g_l*-ként, azaz a* g_1, $g_2, \ldots, g_{l-1}$, g_l^* *sorozat az* $X'Q_0X$ *mátrix* $\lambda_1, \lambda_2, \ldots, \lambda_l$ *sajátvektoraihoz tartozó ortonormált sajátvektorai* ($l = 1, 2, \ldots, s$).

A tétel rekurziós eljárást ad M_k diádjainak számítására, a rekurzió (1) analógja a módosított esetre.

Bizonyítás: $l = 1$-re az állítás azonos λ_1 és g_1 definíciójával. Belátjuk, hogy ha valamely l-re áll a tétel, akkor áll $l + 1$-re is. Valóban, az 1. tétel szerint $\mathrm{Sp}\,(X - M_{l-1} - Y_1)'\, Q_0(X - M_{l-1} - Y_1)$ minimumát az $Y_1 = \sqrt{\lambda_l^*}\, f_l^* g_l^{*\prime}$ diádánál veszi fel, melyben λ_l^* az $(X - M_{l-1})'Q_0(X - M_{l-1})$ mátrix legnagyobb sajátértéke, g_l^* (egy) hozzátartozó normált sajátvektor és $f_l^* = (X - M_{l-1})g_l^*/\sqrt{\lambda_l^*}$. Ekkor viszont valamely, az 1. tételnek megfelelő g_l vektorral:

$$M_{l-1} + \sqrt{\lambda_l^*}\, f_l^*\, g_l^{*\prime} = M_l \tag{11}$$

hiszen a fenti nyom ugyancsak az 1. tétel szerint $M_{l-1} + Y_1 = M_l$-nél minimális. (11)-ből kapjuk

$$\sqrt{\lambda_l^*}\, f_l^*\, g_l^{*\prime} = \sqrt{\lambda_l}\, f_l\, g_l'. \tag{12}$$

Balról $f_l'Q_0$-lal szorozva és figyelembe véve, hogy (5) szerint $f_l'Q_0f_l = 1$, kapjuk: $g_l^* = \mathrm{const} \cdot g_l$. Mivel g_l^* és g_l normáltak, ebből $g_l^* = g_l$ következik.

Mármost felhasználva, hogy $M_{l-1}g_l = 0$ (ami M_{l-1} kifejtéséből könnyen belátható), nyerjük:

$$\begin{aligned}\lambda_l^* &= g_l^{*\prime}(X - M_{l-1})' Q_0 (X - M_{l-1}) g_l^* = \\ &= g_l' X' Q_0 X g_l - 2 g_l' X' Q_0 M_{l-1} g_l + g_l' M_{l-1}' Q_0 M_{l-1} g_l = \\ &= g_l' X' Q_0 X g_l = \lambda_l\end{aligned}$$

amivel a 2. tétel bizonyítását befejeztük.

4. Példa a módszer alkalmazására

Vetésforgó kísérletek elemzésénél problémát okoz, hogy a forgóban szereplő különböző növények hozamai és ezek megbízhatóságai eltérő nagyságrendűek. E nehézség leküzdését célozza az alábbi eljárás.

Legyen a vetésforgóban vizsgált kezelések száma $n + 1$, a kísérlet tartama m év. A kezelés ill. év indexe legyen i ill. j, a kontrollra legyen $i = 0$. Éven belül a különböző kezelésekben azonos növény szerepel. A tapasztalat szerint az η_{ij} hozam (mint valószínűségi változó) szórása arányos e hozam várható-értékével. Erre tekintettel az η_{ij}/η_{0j} relatív hozamok gyakorlatilag azonos szórásúak. E hányadosok egy éven belül — az azonos nevező miatt — nem függetlenek, korrelációjuk egyöntetűen 0,5, az egyes évek között azonban korrelálatlanok. Ha továbbá az η_{ij} hozamok egy növénynél eleget tesznek a szokásos varianciaanalízis modellnek, akkor a kölcsönhatásokra tett bizonyos gyenge megszorítás mellett [6] a

$$\xi_{ij} = \frac{\eta_{ij} - \eta_{0j}}{\eta_{0j}}$$

ún. relatív többlethozamokból álló $n \times m$-es Ξ mátrix várható értéke két diád összegére bomlik:

$$\mathscr{E}\Xi = u_1 v_1' + u_2 v_2' \tag{13}$$

Ξ oszlopvektorai közelítőleg normális, azonos eloszlású vektorváltozók, és egymástól függetlenek, közös kovariancia-mátrixuk:

$$S_0 = \sigma^2 \begin{pmatrix} 1 & {}^1\!/_2 & {}^1\!/_2 & \dots & {}^1\!/_2 \\ {}^1\!/_2 & 1 & {}^1\!/_2 & \dots & {}^1\!/_2 \\ . & . & . & . & . \\ {}^1\!/_2 & {}^1\!/_2 & {}^1\!/_2 & \dots & 1 \end{pmatrix}$$

ahol σ^2, a ξ_{ij} változók közös szórása, a kísérlet ismétléseiből becsülhető.

(13) két diádját ezekután az 1. és 2. tétel alapján becsülhetjük a Ξ mátrixnak a kísérleti eredményekből számítható X realizációja ismeretében. Az $\hat{u}$ és $\hat{v}$ vektorok fontos szakmai következtetésekre alkalmasak, komponenseik között az eltérések szignifikanciája megvizsgálható. Az $\hat{u}_1$ ill. $\hat{v}_1$ vektorok komponensei az *elsődleges* kezelési ill. évi + kumulatív hatásokat becsülik. Az $\hat{u}_2$, $\hat{v}_2$ vektorok az ugyanilyen értelemben vett *másodlagos* hatásokat becsülnek. A két diád informáló erejéről a λ_1 ill. λ_2 sajátérték tájékoztat.

A diádok becslése előtt tesztelhető a (13) modell illeszkedése. S_0 inverze:

$$Q_0 = \frac{2}{n+1}\frac{1}{\sigma^2}\begin{pmatrix} n & -1 & -1 & \dots & -1 \\ -1 & n & -1 & \dots & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & -1 & -1 & \dots & n \end{pmatrix} = c_n T_n,$$

ahol a c_n és T_n jelölések nyilvánvalóak.

Legyenek λ_1 és λ_2 az $X'Q_0X$ mátrix két legnagyobb sajátértéke. Belátható, hogy (13) fennállása esetén az $[\mathrm{Sp}\,\Xi'Q_0\Xi - \lambda_1 - \lambda_2]$ valószínűségi változó közelítőleg χ^2 eloszlást követ, várhatóértéke — azaz szabadságfoka:

$$f_1 = nm - 2(n + m - 1) = (n - 2)(m - 2) - 2$$

Innen egyszerűen következik, hogy az $X'T_nX$ mátrix ν_1, ν_2 ($= \lambda_1/c_n$ ill. λ_2/c_n) két legnagyobb sajátértékével képzett

$$\zeta = \mathrm{Sp}\,\Xi' T_n \Xi - \nu_1 - \nu_2 \tag{14}$$

valószínűségi változó (közelítőleg) $\frac{1}{c_n}\chi^2$ eloszlású, f_1 szabadságfokkal. Ebből, ha a (13) hipotézis helyes, akkor

$$s_1^2 = \frac{2}{n+1}\cdot\frac{z}{f_1}$$

σ^2 egy becslése, ahol z-t (14) szerint kapjuk, ha Ξ helyébe X-et teszünk. σ^2-nek az ismétlésekből számított becslését jelöljük s_2^2-tel, ennek szabadságfoka $f_2 = nm\,(r - 1)$, ahol r az ismétlések száma. Jó illeszkedés esetén az

$$\hat{F} = s_1^2/s_2^2$$

hányados nagyságrendje 1, az illeszkedés az f_1 és f_2 szabadságfokú F-eloszlás táblázatának segítségével tesztelhető.

(13) két diádjának becslése az $X'T_nX$ mátrix g_1 és g_2 normált sajátvektorainak (ezek egyben $X'Q_0X$ saját vektorai is) meghatározását kívánja. Figyelemre méltó, hogy a becsléshez nincs szükség σ^2 numerikus értékének ismeretére. Az $\tilde{f}_1 = Xg_1$, $\tilde{f}_2 = Xg_2$ vektorokkal

$$\hat{u}_1\hat{v}_1 = \tilde{f}_1 g_1', \quad \hat{u}_2\hat{v}_2' = \tilde{f}_2 g_2'.$$

Az $\hat{u}$, $\hat{v}$ vektorokat praktikus úgy választani, hogy a $\hat{v}_1$ ill. $\hat{v}_2$ vektor komponenseinek átlaga 1 legyen. Ekkor ugyanis $\hat{u}_1$ ill. $\hat{u}_2$ komponensei közvetlenül az egyes kezelések elsődleges, ill. másodlagos relatív többlethozamait becsülik a teljes vetésforgóra vonatkozóan.

A leírt módon elemeztük a Keszthelyi Agrártudományi Egyetem egy 10 éves tartamú, 13 kezeléses vetésforgó kísérletét. Érdemes néhány szót ejteni a kapott eredményekről, a részletes elemzést a *Növénytermelés* egy későbbi számában közöljük. Az $\hat{u}_1$ vektor komponensei meglepően jól egyeztek a gabonaegységekre átszámított hozamokkal képzett (évekre átlagolt) ,,nyers'' relatív többlethozamokkal, ami alátámasztja az elemzés eddigi gyakorlatának jogosultságát. Az $\hat{u}_2$ vektor komponensei a kezelések bizonyos jellemzőivel mutattak szoros korrelációt. A két diád informáló erejére 82,5%, ill. 10,5% adódott, a maradék 7% hibahatáron belüli érték, ugyanis az F-próba jó illeszkedést mutatott.

A diádfelbontást elvégeztük arra az esetre is, amikor viszonyítási alapként a kezelések hozamainak átlagát választottuk. Ekkor Ξ oszlopain belül az elemek kevésbé korreláltak, Q_0 nem-diagonális elemei relatíve kisebbek, ugyanakkor az egyes diádok „informáló erejé"-nek így objektívebb jelentés tulajdonítható. Az eredmények kielégítően egyeztek az előbb leírt feldolgozásnál kapottakkal.

(Beérkezett: 1972. aug. 28.)

IRODALOM

1. Anderson, T. W.: An introduction to multivariate statistical analysis. New York, 1958. John Wiley and Sons.
2. Bodewig, E.: Matrix calculus. New York, 1959. Interscience Publ.
3. Egerváry J.: Az inverz mátrix általánosítása. A Matematikai Kutató Intézet Közleményei, 1956/3. 315–324. o.
4. Egerváry J.: Mátrixszámítás. Budapest, 1966. Tankönyvkiadó.
5. Józsa S.: Iteráció mátrixok diadikus közelítésére. A Keszthelyi Agrártudományi Egyetem Kiadványai, 1971. XIII. évf. 1. sz.
6. Józsa S.: Kéttényezős kísérletek kontrollhoz mért értékelése. A Keszthelyi Agrártudományi Egyetem Kiadványai, 1971. XIII. évf. 5. sz.
7. Rimler J.—Székely B.: A gazdasági fejlődésre jellemző közös trendek vizsgálata diád-módszerrel. Szigma, 1971. 1–2. sz. 13–33. o.
8. Székely B.: Mátrixok egy speciális diadikus felbontása és ennek néhány alkalmazása az összehasonlító elemzésben. Szigma, 1970. 4. sz. 241–253. o.

STATISTICAL EXPOSITION OF THE METHOD OF DIADS

Maximum Likelihood estimation of the vectors u, v, for the diad-approximation of some real matrix X

$$X = u_1 v_1' + u_2 v_2' + \ldots + u_k v_k'$$

is investigated. The elements x_{ij} of X are here random variables with normal joint distribution, the covariance matrix S being known. In section 3 the author produces the diads for the special case when the columns of X are independent. For the general case the possibility of an iterative procedure emerges. For the illustration of the method a possible way to analyse crop rotation is presented in section 4.

СТАТИСТИЧЕСКАЯ ЭКСПОЗИЦИЯ МЕТОДА ДИАДОВ

В этой статье автор исследует оценку по максимальному правоподобию векторов u, v, для диадического приближения действительной матрицы X

$$X = u_1 v_1' + u_2 v_2' + \ldots + u_k v_k'.$$

Элементы x_{ij} матрицы X являются здесь случайные величины нормального совокупного разделения, с известной матрицей ковариянции S. В пункте 3 он составляет матрицы на специальный случай, если колонные векторы X являются независимыми друг от друга. На общий случай он намечает возможность метода итерации. В пункте 4 он показывает возможный способ анализа севооборотов, в иллюстрации применения метода.