IMRÉNYI BÉLA

A kanonikus korrelációszámítás

A kanonikus korrelációszámítás a többváltozós statisztikai elemzés témakörébe tartozik. A vizsgálatra kiválasztott valószínűségi változókat két csoportra osztjuk, valamilyen — a változók természetéből adódó — szempont szerint. Az így kapott két változócsoportot felfoghatjuk úgy is, mint két valószínűségi vektorváltozót. Az elméletben levezetésre kerülő „trace korrelációs együttható" a közönséges korrelációs együttható általánosítása a két vektorváltozó között. Bemutatjuk emellett az elmélet egy gyakorlati alkalmazását: egy becslési módszert a többváltozós ökonometriai modell endogén (függő) változóira. Az összefüggések levezetése során alkalmazzuk a matematikai statisztika és a mátrixszámítás egyes alapvető tételeit, és az anyag megértése szempontjából ezek ismerete az olvasó részéről is szükséges.

1. A kanonikus korrelációszámítás problémája

Legyen X a p komponensű $x' = (x_1, \ldots, x_p)$ valószínűségi vektorváltozóból vett T elemű minta (vagy T elemű idősor) mátrixa:

$$X = \begin{pmatrix} x_{11} & x_{12} \ldots x_{1p} \\ x_{21} & x_{22} \ldots x_{2p} \\ \vdots & \\ x_{T1} & x_{T2} \ldots x_{Tp} \end{pmatrix} \tag{1.1}$$

és Y a q komponensű $y' = (y_1, \ldots, y_q)$ valószínűségi vektorváltozóból vett T elemű minta (v. T elemű idősor) mátrixa:

$$Y = \begin{pmatrix} y_{11} & y_{12} \cdots y_{1q} \\ y_{21} & y_{22} \cdots y_{2q} \\ \vdots & \\ y_{T1} & y_{T2} \cdots y_{Tq} \end{pmatrix} \tag{1.2}$$

Az általánosság megszorítása nélkül feltehetjük, hogy $q \leq p$.

A vektorok komponenseit a változók jelentése szerint állítjuk össze. Például az x vektor tartalmazhat közgazdasági mutatókat, az y pedig szociológiaikat, számszerű fiziológiai, illetve pszichológiai kísérletek eredményeit, ökonometriai modell predeterminált, illetve endogén változóit stb.

Vizsgálni akarjuk, hogy milyen mértékű sztochasztikus kapcsolat van a két vektorváltozó között. Nagyon áttekinthető volna ez az összefüggés, ha x_1 és y_1 korrelációja nagy volna, ugyanígy x_2 és y_2-é, . . ., x_q és y_q-é. Ugyanakkor mindegyik változó kizárólag a „saját párjával" volna korrelált, az x_i és y_j függetlenek lennének, ha $i \neq j$.

A változók összefüggéseit leíró lineáris modell ekkor csakis az

$$\begin{aligned} y_1 &= x_1 \beta_1 + w_1 \\ y_2 &= x_2 \beta_2 + w_2 \\ &\vdots \\ y_q &= x_q \beta_q + w_q \end{aligned} \tag{1.3}$$

alakot öltheti (itt feltettük, hogy x_i, y_i és w_i várható értéke zérus, és w_i a sztochasztikus kapcsolattal együttjáró disturbancia).

Természetesen egy statisztikai adatfelvétel sem fog ilyen összefüggést tartalmazó adatokat adni. Ellenben bevezethetünk olyan változókat, amelyek mutatják a fenti tulajdonságokat, és ugyanakkor tartalmazzák az összes információt is, amely a statisztikai adatokban van. Ezek a változók az x, illetve y vektorok komponenseinek különböző lineáris kombinációi lesznek, maguk is valószínűségi vektorváltozók:

$$\begin{aligned} u_1 &= a_1' x \\ u_2 &= a_2' x \\ &\vdots \\ u_q &= a_q' x \end{aligned} \tag{1.4}$$

és

$$\begin{aligned} v_1 &= b_1' y \\ v_2 &= b_2' y \\ &\vdots \\ v_q &= b_q' y \end{aligned} \tag{1.5}$$

Vektorálisan:

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_q \end{pmatrix} = \begin{pmatrix} a_1' x \\ a_2' x \\ \vdots \\ a_q' x \end{pmatrix} \text{ vagy } u = A' x \tag{1.6}$$

és

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_q \end{pmatrix} = \begin{pmatrix} b_1' y \\ b_2' y \\ \vdots \\ b_q' y \end{pmatrix} \text{ vagy } v = B' y \tag{1.7}$$

A matematikai feladat lesz az A és B mátrixokat meghatározni oly módon, hogy a kapott u és v valószínűségi vektorváltozók mutassák az (1.3) egyenletekkel kapcsolatban említett tulajdonságokat, nevezetesen azokat, hogy u_1 és v_1, u_2 és v_2, . . . u_q és v_q a lehető legszorosabban korreláljon, és az összes többi korreláció zérus legyen.

Most kimondhatjuk a pontos *definíciót:* Ha adott a p komponensű x és a q komponensű y valószínűségi vektorváltozó, akkor képezzük az A és B konstans mátrixokkal az $u = A'x$ és $v = B'y$ valószínűségi vektorváltozókat. Az így kapott $u' = (u_1, \ldots, u_q)$ és $v' = (v_1, \ldots, v_q)$ vektorok komponenseit „kanonikus változóknak" nevezzük, ha az

$$\varepsilon u_i^2 = 1 \tag{1.8}$$

$$\varepsilon v_i^2 = 1 \qquad i = 1, \ldots, q \tag{1.9}$$

feltétel mellett teljesül az, hogy a következő korrelációk zérusok:

$$\varepsilon u_i u_j = 0 \tag{1.10}$$

$$\varepsilon v_i v_j = 0 \tag{1.11}$$

$$\varepsilon u_i v_j = 0 \qquad i = 1, \ldots, q \tag{1.12}$$

ha $i \neq j$; azonkívül ha még teljesül az $\varepsilon u_i v_i$ $(i = 1, \ldots, q)$ korrelációkra az (1.8), (1.9) feltételek mellett az, hogy az $\varepsilon u\, v = a'\varepsilon xy'b$, $p + q$ változós függvény lokális maximumai. Az $\varepsilon u\, v$ függvénynek e lokális maximumait „kanonikus korrelációs együtthatóknak" nevezzük. (A görög eredetű „kanonikus" szó jelentése: szabályos, szabályszerű.)

2. A kanonikus korreláció számítás matematikai alapjai

Tekintsük a $p + q$ komponensű $\begin{pmatrix} x \\ y \end{pmatrix}$ valószínűségi vektorváltozót Σ variancia-kovariancia mátrixszal, amely négy rész-mátrixból áll:

$$\Sigma = \varepsilon \begin{pmatrix} x \\ y \end{pmatrix} (x'\ y') = \varepsilon \begin{pmatrix} xx' & xy' \\ yx' & yy' \end{pmatrix} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}. \tag{2.1}$$

Először azt az esetet vizsgáljuk, amikor $\varepsilon \begin{pmatrix} x \\ y \end{pmatrix} = 0$. Később látni fogjuk, hogy ez a megszorítás nem változtat a kanonikus korrelációs együtthatók értékén. Feltesszük még azt, hogy Σ pozitív definit. Ez utóbbi feltételezéssel nem zártunk ki egyetlen számunkra jelentőséggel bíró esetet sem, csupán azokat az érdektelen eseteket, amelyekben a komponensek lineárisan összefüggőek.

E feltételek mellett a következő *tételt* fogjuk bebizonyítani:

Az x és y vektorokhoz találhatók (1.4) és (1.5) alakban felírt kanonikus változók. Az $r_1, \ldots, r_q$ kanonikus korrelációs együtthatók a

$$\begin{vmatrix} -r\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -r\Sigma_{22} \end{vmatrix} = 0 \tag{2.2}$$

egyenlet pozitív gyökei.

Az i-edik kanonikus változó-pár képzéséhez szükséges a_i, b_i konstans vektorok pedig a

$$\begin{pmatrix} -r_i\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -r_i\Sigma_{22} \end{pmatrix} \cdot \begin{pmatrix} a_i \\ b_i \end{pmatrix} = 0 \tag{2.3}$$

homogén lineáris egyenletrendszer gyökei $i = 1, \ldots, q$ esetén, feltéve, hogy $r_i \neq r_j$, ha $i \neq j$.

Először megkeressük az $u = a'x$ és $v = b'y$ alakú skaláris szorzatok közül azokat, amelyek kielégítik az

$$1 = \varepsilon u^2 = \varepsilon a'xx'a = a'\Sigma_{11}a \tag{2.4}$$

$$1 = \varepsilon v^2 = \varepsilon b'yy'b = b'\Sigma_{22}b \tag{2.5}$$

feltételeket, azaz varianciájuk egységnyi, és e feltételek mellett az εuv függvénynek (amely az u és v változók korrelációs együtthatója) szélsőértéke van. A függvény más alakban:

$$\varepsilon uv = \varepsilon a'xy'b = a'\varepsilon xy'b = a'\Sigma_{12}b \tag{2.6}$$

A Lagrange-féle feltételes szélsőértékszámítás módszerével keressük meg (2.6) szélsőértékeit:

Legyen:

$$\psi = a'\,\Sigma_{12}\,b - \frac{1}{2}\,\lambda(a'\,\Sigma_{11}\,a - 1) - \frac{1}{2}\,\mu(b'\,\Sigma_{22}\,b - 1) \tag{2.7}$$

ahol $\frac{1}{2}\lambda$ és $\frac{1}{2}\mu$ Lagrange-féle multiplikátorok. A ψ függvény szélsőértékei között lesznek az εuv szélsőértékei (2.4) és (2.5) feltételek mellett. Deriváljuk ψ-t a és b komponensei szerint, és zérussal tesszük egyenlővé:

$$\frac{\partial\psi}{\partial a} = \Sigma_{12}\,b - \lambda\Sigma_{11}\,a = 0 \tag{2.8}$$

$$\frac{\partial\psi}{\partial b} = \Sigma_{21}\,a - \mu\Sigma_{22}\,b = 0. \tag{2.9}$$

(2.8)-at szorozzuk balról a'-val, (2.9)-et b'-vel:

$$a'\Sigma_{12}b - \lambda a'\Sigma_{11}a = 0 \tag{2.10}$$

$$b'\Sigma_{21}a - \mu b'\Sigma_{22}b = 0. \tag{2.11}$$

Az egységnyi varianciák következtében $\lambda = \mu = a'\Sigma_{12}b = \varepsilon uv$, tehát λ értéke éppen az u és v változók korrelációs együtthatója. Így (2.8)-at és (2.9)-et a következőképpen írhatjuk át.

$$-\lambda\Sigma_{11}a + \Sigma_{12}b = 0 \tag{2.12}$$

$$\Sigma_{21}a - \lambda\Sigma_{22}b = 0 \tag{2.13}$$

vagy mátrix egyenletben:

$$\begin{pmatrix} -\lambda\,\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = 0. \tag{2.14}$$

mive 14) homogén lineáris egyenletrendszer az ismeretlen $\begin{pmatrix} a \\ b \end{pmatrix}$ vektorra nézve, a nem triviális megoldás létezésének szükséges és elégséges feltétele, hogy mátrixának determinánsa zérus legyen:

$$\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0. \tag{2.15}$$

A (2.14) és (2.15) egyenletekkel megkaptuk a tételben szereplő egyenleteket. Ezzel azt láttuk be, hogy az uv függvény feltételes szélsőérték helyei csakis a (2.14) és (2.15) egyenletek által meghatározott $\begin{pmatrix} a \\ b \end{pmatrix}$ vektorok között lehetnek.

Belátjuk, hogy (2.15)-nek λ-ra q számú pozitív gyöke van. (2.15) egy $(p+q)$-adfokú algebrai egyenlet, mert $(-\lambda)^{p+q}$ együtthatója $(\Sigma_{11}) \cdot (\Sigma_{22})$, ahol egyik tényező sem zérus, és a determináns kifejtésében az összes többi tagban λ kitevője $(p+q)$-nál kisebb. Ezért $(p+q)$ számú gyöke van az egyenletnek. Az összes gyök valós, mert ha a (2.14) egyenletet balról szorozzuk az $(\bar{a}', \bar{b}')$ vektorral, azt kapjuk, hogy

$$\lambda = \bar{a}'\Sigma_{12} b \quad \text{és} \quad \lambda = \bar{b}'\Sigma_{21} a$$

azaz λ egyenlő a saját konjugáltjával.

Vizsgáljuk meg a (2.15) egyenletet a gyökök szempontjából. Nem változtat a gyökökön, ha az egyenletet megszorozzuk a

$$\frac{|\Sigma_{22}^{-1}|}{|\Sigma_{11}|} \text{ konstanssal:}$$

$$\frac{|\Sigma_{22}^{-1}|}{|\Sigma_{11}|} \cdot \begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0 \tag{2.16}$$

Az egyenlet bal oldalát a lehető legegyszerűbb alakra hozzuk:

$$\begin{aligned}
&\frac{|\Sigma_{22}^{-1}|}{|\Sigma_{11}|}(-\lambda\Sigma_{11}) \cdot |-\lambda\Sigma_{22} - \Sigma_{21}(-\lambda\Sigma_{11})^{-1}\Sigma_{12}| = \\
&= (-\lambda)^p (\Sigma_{22})^{-1} |-\lambda\Sigma_{22} - \Sigma_{21}(-\lambda\Sigma_{11})^{-1}\Sigma_{12}| = \\
&= (-\lambda)^p \left| -\lambda I - \Sigma_{22}^{-1}\Sigma_{21}\frac{1}{-\lambda}\Sigma_{11}^{-1}\Sigma_{12} \right| = \\
&= (-\lambda)^p \left| \frac{1}{-\lambda} \cdot (-\lambda^2 \cdot I + \Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) \right| = \\
&= \frac{(-\lambda)^p}{(-\lambda)^q} |\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \lambda^2 I| = \\
&= (-\lambda)^{p-q} |\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \lambda^2 I| = 0
\end{aligned} \tag{2.17}$$

Láthatjuk, hogy a (2.15) egyenletben λ^{p-q} gyöktényezőt kiemelhetjük, tehát az egyenlet gyökei között $p-q$ számú zérus van. A további $2q$ számú gyök a

$$|\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \lambda^2 I| = 0 \tag{2.18}$$

egyenlet gyökei. Itt λ^2 lehetséges értékei a $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ mátrix saját értékei, mivel ez a mátrix nem szinguláris, nincs zérus sajátértéke. Fent láttuk, hogy minden λ valós értékű, ezért minden λ^2 pozitív. (Ezzel közvetve beláttuk azt is, hogy a mátrix pozitív definit.) A nagyság szerint sorbarendezett $p+q$ számú gyök a következőképpen írható fel:

$$\lambda_1 > \lambda_2 > \ldots > \lambda_q > 0 = \ldots = 0 > -\lambda_q > \ldots > -\lambda_1.$$

A (2.18) egyenletet más módon is levezethetjük. Tekintsük a (2.18) és (2.13) egyenleteket. (2.12)-ből $\lambda\Sigma_{11}a = \Sigma_{12}b$. Ezt balról szorozva Σ_{11}^{-1}-vel

$$\lambda a = \Sigma_{11}^{-1}\Sigma_{12}\, b \tag{2.19}$$

$$a = \frac{1}{\lambda}\Sigma_{11}^{-1}\Sigma_{12}\, b. \tag{2.20}$$

(2.13)-ból $\Sigma_{21}a = \lambda\Sigma_{22}\, b$. Ezt Σ_{22}^{-1}-nel balról szorozva:

$$\Sigma_{22}^{-1}\Sigma_{21}\, a = \lambda b \tag{2.21}$$

egyenletet kapjuk. (2.21) mindkét oldalát szorozzuk λ-val, és λa helyébe (2.16) jobb oldalát helyettesítjük:

$$\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b = \lambda^2 b \tag{2.22}$$

vagy

$$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b = \lambda^2\Sigma_{22}\, b. \tag{2.22/a}$$

Eszerint a b vektorok valóban a $\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ mátrix sajátvektorai, λ_i^2 $(i = 1, \ldots q)$ pozitív valós számok a sajátértékek.

(2.22) levezetésével lépésről-lépésre megegyező levezetéssel beláthatjuk, hogy

$$\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\cdot a = \lambda^2 a \tag{2.23}$$

vagy

$$\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\cdot a = \lambda^2\Sigma_{11}\, a. \tag{2.23/a}$$

(2.22) és (2.23) egyenletek között az a különbség, hogy míg az előbbi együtthatómátrixa $q\times q$, az utóbbi $p\times p$ méretű $(p\geq q)$. A mátrixszorzat tényezőinek rangjai között a legkisebb q, így a (2.23) mátrix szinguláris, és sajátértékei között $p-q$ számú zérus van.

Ezekből az egyenletekből láthatjuk azt is, hogy miért volt szükség arra a feltételre, hogy a különböző sorszámú kanonikus korrelációs együtthatók különbözzenek egymástól: ha két együttható egyenlő volna — pl. $\lambda_i = \lambda_{i+1}$ —, akkor a λ_i^2 sajátértékhez végtelen sok sajátvektor tartozna, amelyek egy kétdimenziós lineáris alteret feszítenek ki. Ekkor a kanonikus változók előállítása már nem egyértelmű. (Bár a „trace korreláció" értelmezhető.)

Most belátjuk, hogy ha $\lambda_i \neq \lambda_j$, akkor $\varepsilon v_i v_j = 0$. Ugyanis: $\varepsilon v_i v_j = b_i'\Sigma_{22}b_j$.

Itt (2.22)-ből behelyettesítjük $b_j = \frac{1}{\lambda_j^2}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b_j$ értékét:

$$\varepsilon v_i v_j = b_i'\Sigma_{22}\frac{1}{\lambda_j^2}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b_j = \frac{1}{\lambda_j^2}b_i'\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b_j$$

$$\lambda_j^2\,\varepsilon\, v_i v_j = b_i'\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b_j. \tag{2.24}$$

Ugyanígy kapjuk:

$$\lambda_i^2\,\varepsilon\, v_j v_i = b_j'\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\, b_i. \tag{2.25}$$

Mivel $\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ szimmetrikus, és (2.24) és (2.25) jobb oldala skalár, ezért azok egyenlők. A két egyenletet kivonva egymásból:

$$(\lambda_j^2 - \lambda_i^2)\,\varepsilon v_i v_j = 0 \tag{2.26}$$

egyenletet kapjuk. A feltétel szerint $\lambda_i \neq \lambda_j$, így $\varepsilon v_i v_j = 0$. Ugyanígy beláthatjuk, hogy a zérustól és egymástól különböző sajátértékekhez tartozó a_i, a_j sajátvektorokból képzett u_i, u_j kanonikus változók is korrelálatlanok. Ebből már az is következik, hogy $\varepsilon u_j v_i = 0$, ha $\lambda_j \neq 0$, mert (2.24) felhasználásával:

$$\varepsilon v_i v_j = \frac{1}{\lambda_j^2} b_i' \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} b_j = \frac{1}{\lambda_j^2} b_i' \Sigma_{21} \lambda_j a_j = \frac{1}{\lambda_j} b_i' \Sigma_{21} a_j = \frac{1}{\lambda_j} \varepsilon v_i u_j . \quad (2.27)$$

Itt a kifejezés második átalakításánál a (2.19) összefüggést alkalmaztuk, és mivel már beláttuk, hogy $\varepsilon v_i v_j$ zérus, ezért $\varepsilon v_i u_j$ is az.

Tehát a $\lambda_1 > \lambda_2 > \ldots > \lambda_q > 0$ értékhez tartozó $u_1 = a_1' x$, $v_1 = b_1' y$, $\ldots u_q = a_q' x$, $v_q = b_q' y$ változókra teljesül az $\varepsilon u_i^2 = 1$ és $\varepsilon v_i^2 = 1$ feltétel mellett az $\varepsilon u_i u_j = \varepsilon u_i v_j = \varepsilon v_i v_j = 0$, azaz a nevezett korrelációk eltűnnek.

Ha még bebizonyítjuk, hogy minden egyes $\lambda_1, \ldots, \lambda_q$ korreláció érték az εuv függvénynek lokális maximuma, akkor a tételbe foglalt állítást teljes egészében bebizonyítottuk. Ez a hátralevő bizonyítás elég hosszadalmas, és a kanonikus változók természetéről nem tartalmaz használható tulajdonságokat, azért itt mellőzzük, részletesen megtalálható [1]-ben. Mivel λ (a Lagrange-féle állandó) egyenlőnek bizonyult a kanonikus korrelációs együtthatóval, a jelölésen már nem változtatunk, λ-val jelöljük továbbra is a kanonikus korrelációs együtthatót.

A (2.17) egyenlet gyökeinek száma $p + q$. Ezek közül csak q számú pozitív figyelembevételével képeztünk kanonikus változókat. Mivel a (2.23) egyenletben λ^2 szerepel, ezért λ helyébe a negatív megoldásokat helyettesítve ugyanazt az egyenletrendszert kapjuk, mintha a pozitív megoldásokat helyettesítettük volna be, így lényegileg különböző megoldáshoz nem jutunk.

3. A kanonikus korrelációs együtthatók és a kanonikus változók becslése

Az $\begin{pmatrix} x \\ y \end{pmatrix}$ vektorból vett T elemű minta mátrixa, (X, Y), mérete $T \times (p+q)$, egy sora az (x', y') vektorból vett minta egy eleme. A vektor variancia-kovariancia mátrixának, Σ-nak a becslése:

$$S = \frac{1}{T}(X, Y)' \, (X, Y) = \frac{1}{T} \begin{pmatrix} X'X & X'Y \\ Y'X & Y'Y \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad (3.1)$$

mert az $\varepsilon \begin{pmatrix} x \\ y \end{pmatrix} = 0$ feltételt még megtartjuk.

$\lambda_1, \ldots, \lambda_q$ becsléseit, $\hat{\lambda}_1, \ldots, \hat{\lambda}_q$ értékeket a

$$\begin{vmatrix} -\hat{\lambda} S_{11} & S_{12} \\ S_{21} & -\hat{\lambda} S_{22} \end{vmatrix} = 0 \quad (3.2)$$

egyenlet gyökei adják meg. Ezeket egyszerűbben az

$$|\,(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12} - \hat{\lambda}^2 I)\,| = 0 \quad (3.3)$$

egyenletből számítjuk ki. $\hat{\lambda}^2$ pozitív gyökei adják a q számú megoldást.

A kapott $\hat{\lambda}_1, \ldots, \hat{\lambda}_q$ értékeket egyenként visszahelyettesítjük a

$$\begin{pmatrix} -\hat{\lambda} S_{11} & S_{12} \\ S_{21} & -\hat{\lambda} S_{22} \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = 0 \tag{3.4}$$

egyenletbe. Ebből kiszámíthatjuk az $\begin{pmatrix} \hat{a}_i \\ \hat{b}_i \end{pmatrix}$ vektorokat, egy határozatlan konstans szorzó erejéig. (Az egyenlet csak a vektor állását határozza meg, hosszát nem.) A vektoroknak ki kell még elégíteniük az

$$\hat{a}_i\, S_{11} \hat{a}_i = 1 \tag{3.5}$$

és

$$\hat{b}_i\, S_{22} \hat{b}_i = 1 \tag{3.6}$$

egyenleteket. Ezek már egyértelműen meghatározzák $\begin{pmatrix} \hat{a}_i \\ \hat{b}_i \end{pmatrix}$ számszerű értékét.

Lényegében egyszerűbb számítási módszer, ha a kapott $\hat{\lambda}_i (i = 1, \ldots, q)$ értékeket az

$$(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12} - \hat{\lambda}_i^2 I)\, \hat{b}_i = 0 \tag{3.7}$$

egyenletbe helyettesítjük vissza. (3.7) ugyanis csak $(q \times q)$ méretű homogén, lineáris egyenletrendszer $\hat{b}_i$ komponenseire nézve, míg (3.4) $(p+q) \times (p+q)$ méretű. (3.7) megoldásait a (3.6) egyenlet teszi egyértelművé. $\hat{a}_i$ értékeit a (2.20) formulából számítjuk ki:

$$\hat{a}_i = \frac{1}{\hat{\lambda}_i} S_{11}^{-1} S_{12} b_i \tag{3.8}$$

Ezzel megkaptuk a kanonikus változók együtthatóit. Az eredményül kapott a_i konstans vektorokra a (3.5) feltételnek azonosan teljesülnie kell (jó módszer az eredmény ellenőrzésére), mert

$$\hat{a}_i S_{11} \hat{a}_i = \left(\frac{1}{\hat{\lambda}_i} S_{11}^{-1} S_{12} \hat{b}_i \right) S_{11} \hat{a}_i = \frac{1}{\hat{\lambda}_i} \hat{b}_i' S_{21} S_{11}^{-1} S_{11} \hat{a}_i =$$

$$= \frac{1}{\hat{\lambda}_i} \hat{b}_i' S_{21} \hat{a}_i = \frac{1}{\hat{\lambda}_i} \hat{\lambda}_i = 1.$$

Az utolsó átalakításban a (2.6) egyenlet megfelelőjét használtuk fel.

4. Megjegyzés a számításhoz

A többváltozós statisztikai analízisben gyakran van szükség a változók transzformálására, pl. indexszámokká alakítás, standardizálás stb. Be fogjuk bizonyítani, hogy a kanoniku korrelációs együtthatók értékei invariánsok a változók ilyen transzformációival szemben.

Tétel: Ha a 2. szakaszban szereplő x és y vektorok koordinátáit a $w_i = e_i x_i + f_i \,(i = 1, \ldots, p)$ és $z_i = g_i y_i + h_i \,(i = 1, \ldots, q)$ transzformációknak vetjük alá, ahol $e_i \neq 0$ és $g_i \neq 0$, a transzformált valószínűségi vektorokhoz akkor, és csak akkor találhatunk egyértelmű kanonikus változókat, ha x és y-hoz is találhatunk, és ha az x és y vektorok kanonikus korrelációs együtthatói $\lambda_1 \ldots, \lambda_q$, akkor ugyanezek lesznek a w és z vektoroké is.

Bizonyítás: A tételben szereplő transzformációk a

$$w = Ex + f \tag{4.1}$$

$$z = Gy + h \tag{4.2}$$

egyenletekbe foglalhatók össze, ahol $E = \langle e_1, \ldots, e_p \rangle$ és $G = \langle g_1, \ldots, g_q \rangle$ diagonális mátrixok,

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_p \end{pmatrix} \quad \text{és} \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_q \end{pmatrix}.$$

A transzformált vektorok várható értéke ($\varepsilon x = 0$ és $\varepsilon y = 0$ következtében) $\varepsilon w = \varepsilon(Ex + f) = E\varepsilon x + f = f$, ugyanígy $\varepsilon z = h$, tehát a konstans vektorokkal egyenlő. A transzformált mátrixok variancia-kovariancia mátrixa:

$$\Sigma = \begin{pmatrix} \varepsilon(w-f)(w-f)' & \varepsilon(w-f)(z-h)' \\ \varepsilon(z-h)(w-f)' & \varepsilon(z-h)(z-h)' \end{pmatrix} = \begin{pmatrix} \varepsilon(Ex)(Ex)' & \varepsilon(Ex)(Gy)' \\ \varepsilon(Gy)(Ex)' & \varepsilon(Gy)(Gy)' \end{pmatrix} =$$

$$= \begin{pmatrix} \varepsilon(Exx'E) & \varepsilon(Exy'G) \\ \varepsilon(Gyx'E) & \varepsilon(Gyy'G) \end{pmatrix} = \begin{pmatrix} E\Sigma_{11}E & E\Sigma_{12}G \\ G\Sigma_{21}E & G\Sigma_{22}G \end{pmatrix}. \tag{4.3}$$

A kanonikus korrelációs együtthatók és a kanonikus változók keresésekor most az $a'E\Sigma_{12}Gb$ függvény értékének a maximumát kell keresni, $a'E\Sigma_{11}Ea = 1$ és $b'G\Sigma_{22}Gb = 1$ feltétel mellett. Ez a számítás pontosan ugyanúgy történik, mint azt a 2. pontban bebizonyított tételben láttuk. A különbség csak az, hogy Σ_{11} helyén $E\Sigma_{11}E$ stb. áll. Ennek eredményeképpen kapjuk, hogy a kanonikus korrelációs együtthatók a

$$\begin{vmatrix} -\lambda E\Sigma_{11}E & E\Sigma_{12}G \\ G\Sigma_{21}E & -\lambda G\Sigma_{22}G \end{vmatrix} = 0 \tag{4.4}$$

egyenlet gyökei. E determináns mátrixát átalakíthatjuk:

$$\begin{pmatrix} -\lambda E\Sigma_{11}E & E\Sigma_{12}G \\ G\Sigma_{21}E & -\lambda G\Sigma_{22}G \end{pmatrix} = \begin{pmatrix} E & 0 \\ 0 & G \end{pmatrix} \begin{pmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{pmatrix} \begin{pmatrix} E & 0 \\ 0 & G \end{pmatrix}. \tag{4.5}$$

Ebből következik, hogy a (4.4) egyenlet és a

$$\begin{vmatrix} -\lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\lambda\Sigma_{22} \end{vmatrix} = 0$$

egyenletek ekvivalensek. Itt felhasználtuk azt, hogy $|AB| = |A|\,|B|$. Ezzel a tételt bebizonyítottuk.

A tétel értelmében az idősorok adatait számolhatjuk tényleges számokkal, tetszés szerinti bázisú indexszámokkal, tetszés szerinti egységeket választhatunk, vagy standardizálhatunk. Ezek a transzformációk mind a tételbeli transzformációval egyező alakúak.

Vigyázat! Az első differenciákkal való számolás nem ad helyes eredményt. Valójában a tételbe foglaltaknál jóval több is igaz az F és G mátrixoknak nem szükséges diagonálisaknak lenniök, elég ha négyzetes mátrixok, továbbá nem szingulárisak.

5. Regressziós egyenletek a kanonikus változók segítségével

A következőkben két változó halmaz közötti sztochasztikus kapcsolatot akarjuk vizsgálni. Az x vektor komponenseit tekintjük független változóknak, az y-éit pedig függő változóknak. (Ökonometriai terminológiával: predeterminált és endogén változók.) Abból az alapfeltevésből indulunk ki, hogy a független változók nem hatnak egymásra, értéküket egymástól függetlenül változtatják, ellenben az összes független változó hatást gyakorol az összes függő változóra. Ezekkel összhangban fel kell tennünk azt is, hogy a függő változók között is kölcsönös egymásrahatások vannak (interdependencia). Ezeket a kapcsolatokat jellemző modellt kapunk, ha kétváltozós regressziós egyenleteket írunk fel a kanonikus változók között, (1.3) alakjában. Ehhez a becsléshez az $u' = (u_1, \ldots, u_q)$ és $v' = (v_1, \ldots, v_q)$ vektorokból vett mintára (idősorra) van szükség. Legyen ezek mátrixa:

$$U = \begin{pmatrix} u_{11} & u_{12}, & \ldots, & u_{1q} \\ u_{21} & u_{22}, & \ldots, & u_{2q} \\ \vdots & & & \\ u_{T1} & u_{T2}, & \ldots, & u_{Tq} \end{pmatrix} \tag{5.1}$$

és

$$V = \begin{pmatrix} v_{22} & v_{12} & \ldots, & v_1 \\ v_{21} & v_{22} & \ldots, & v_2 \\ \vdots & & & \\ v_{T1} & v_{T2}, & \ldots, & v_{Tq} \end{pmatrix} \tag{5.2}$$

A mátrixok i-edik sora az u, illetve v vektorokból vett minta (idősor) i-edik eleme. Mivel az u és v komponensei elvont változók, közgazdasági jelentés nélkül, ezért azokból a szokásos módon nem is tudunk ,,mintát venni''. Az i-edik komponensre vett T elemű minta itt azt jelenti, hogy az $u_i = x' a_i$ skaláris szorzatban x' helyébe rendre behelyettesíthetjük az x' vektorra vett T elemű minta elemeit. Ez mátrix alakban:

$$\begin{pmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{Ti} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12}, & \ldots, & x_{1p} \\ x_{21} & x_{22}, & \ldots, & x_{2p} \\ \vdots & & & \\ x_{T1} & x_{T2}, & \ldots, & x_{Tp} \end{pmatrix} \cdot a_i \tag{5.3}$$

és

$$\begin{pmatrix} v_{1i} \\ v_{2i} \\ \vdots \\ v_{Ti} \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12}, & \ldots, & y_{1q} \\ y_{21} & y_{22}, & \ldots, & y_{2q} \\ \vdots & & & \\ y_{T1} & y_{T2}, & \ldots, & y_{Tq} \end{pmatrix} b_i \tag{5.4}$$

$i = 1, \ldots, q$-ra. Így az U és V mátrixok elemeit az X és X mátrixokból, valamint a_i és b_i értékeiből számítjuk ki. Szükség lesz még $\hat{\lambda}_i$ értékére is, a regressziós egyenletek becslésénél. Mivel ettől a paragrafustól kezdve elméleti értékekről nem lesz szó, az egyszerűség kedvéért a becslésre vonatkozó $\wedge$ jelet elhagyjuk, $\hat{a}_i$, $\hat{b}_i$, $\hat{\lambda}_i$ helyett csak a_i, b_i, λ_i-t fogunk írni. A $\wedge$ jelet, ha használjuk, az a függő változók becslésekor kapott regresszió értékeket fogja jelölni.

Jelöljük röviden $A = (a_1, \ldots, a_q)$ és $B = (b_1, \ldots, b_q)$. Ennek segítségével az (5.3) és (5.4) egyenlőségek $i = 1, \ldots, q$-ra az

$$U = XA \tag{5.5}$$

$$V = YB \tag{5.6}$$

alakba írhatók röviden. A variancia-kovariancia mátrixok jelölése (3.1)-gyel megegyezően:

$$S_{11} = \frac{1}{T} X' X \tag{5.7}$$

$$S_{12} = \frac{1}{T} X' Y \tag{5.8}$$

$$S_{22} = \frac{1}{T} Y' Y. \tag{5.9}$$

Felírjuk a kanonikus változók variancia-kovariancia mátrixait is:

$$\frac{1}{T} U' U = \frac{1}{T} A' X' XA = A' \frac{1}{T} X' XA = A' S_{11} A.$$

Az $A' S_{11} A$ mátrix i-edik sorának j-edik eleme $a_i' S_{11} a_j$. Ennek értéke 1, ha $j = i$, — azaz a főátlóban — és zérus, ha $i \neq j$, tehát $\frac{1}{T} U'U$ egyegységmátrix. Az

$$\frac{1}{T} U' U = A' S_{11} A = I \tag{5.10}$$

egyenlőség tartalmazza azt a feltételt, hogy a kanonikus változók egységnyi varianciájúak, és azt a tételt, hogy a különböző kanonikus változók korrelálatlanok. Hasonlóképpen írhatjuk fel az

$$\frac{1}{T} V' V = B' S_{22} B = I \tag{5.11}$$

egyenlőséget is.

Bevezetjük a

$$\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_q \end{pmatrix}$$

diagonális mátrixot, és ennek segítségével az

$$\frac{1}{T} U' V = A' S_{12} B = \Lambda \tag{5.12}$$

mátrixegyenlőséget írhatjuk fel. E mátrix i-edik sorának j-edik eleme $a_i'S_{12}b_j$, mely $i = j$ esetén λ_i (az i-edik kanonikus korrelációs együttható), $i \neq j$ esetén pedig zérus.

Végül megemlítjük még azt, hogy a (3.8) egyenleteket:

$$a_i = \frac{1}{\lambda_i} S_{11}^{-1} S_{12} b_i \quad i = 1, \ldots, q$$

röviden az

$$A = S_{11}^{-1} S_{12} B \Lambda^{-1} \tag{5.13}$$

alakban írhatjuk. Ugyanígy levezethető a

$$B = S_{22}^{-1} S_{21} A \Lambda^{-1} \tag{5.14}$$

mátrixegyenlőség.

Ezekután elkezdhetjük a tulajdonképpeni feladatot, a „kanonikus függő" változók becslését a „kanonikus független" változók segítségével. A becslést a legkisebb négyzetek módszerével hajtjuk végre:

$$\hat{V} = U(U'U)^{-1}U'V = U\frac{1}{T} \cdot I \cdot T \cdot \Lambda = U\Lambda. \tag{5.15}$$

Az i-edik egyenletben a regressziós együttható az i-edik kanonikus korrelációs együttható:

$$\hat{v}_i = u_i \lambda_i$$

U és V értékét az (5.5) és (5.6) egyenletek segítségével átalakítva:

$$\widehat{YB} = XA\Lambda \tag{5.16}$$

Vizsgáljuk meg, mit kapunk, ha közvetlenül Y-t becsüljük meg az U függvényében:

$$\hat{Y} = XA(A'X'XA)^{-1}A'XY \tag{5.17}$$

(5.7), (5.8) és 5.10) segítségével:

$$\hat{Y} = XA(TI)^{-1}A'TS_{12}$$

$$\hat{Y} = XAA'S_{12} \tag{5.18}$$

(5.14)-ből kapjuk:

$$S_{22}B'\Lambda = S_{21}A,$$

ennek transzponáltja:

$$A'S_{12} = \Lambda B'S_{22}.$$

Ennek jobboldalát (5.18)-ban $A'S_{12}$ helyébe írva

$$\hat{Y} = XA\Lambda B'S_{22} \tag{5.19}$$

egyenletet kapjuk. Ezt a képletet később fel fogjuk használni, előbb azonban egy ebből következő fontos tényt kell megvizsgálnunk. (5.19)-ben A helyébe (5.13) jobb oldalát helyettesítjük.

$$\hat{Y} = XS_{11}^{-1}S_{12}B\Lambda^{-1}\Lambda B'S_{22}$$

$$Y = XS_{11}^{-1}S_{12}BB'S_{22} \tag{5.20}$$

$BB'S_{22}$-ről megmutatjuk, hogy $(q \times q)$ méretű egységmátrix: az (5.11) egyenlőséget balról és jobbról beszorozzuk B-vel, illetve B^{-1}-vel

$$BB' \cdot S_{22} BB^{-1} = BB' S_{22} = I.$$

Ezt megtehettük, mert $B(q \times q)$ méretű nem szinguláris mátrix. Így (5.19)-ből az

$$\hat{Y} = X S_{11}^{-1} S_{12} \tag{5.21}$$

egyenlőséget kaptuk, ami nem más, mint az Y-nak a legkisebb négyzetek módszerével történő becslése X függvényében:

$$\hat{Y} = X(X'X)^{-1} XY$$

$$\left(\text{beszorozva } T \text{ és } \frac{1}{T}\text{-vel}\right).$$

Ebből az eredményből azt a következtetést vonhatjuk le, hogy az (5.19) vagy az (5.17) formulával végrehajtott becslés ugyanazt az $\hat{Y}$ regressziós értékeket adja, mint a legkisebb négyzetek módszerével történő becslés.

Most térjünk vissza az (5.15) egyenlettel kapcsolatos eljáráshoz: az elvont „kanonikus változókkal" hajtjuk végre a becslést. A becsült egyenletek alakja a következő:

$$\hat{v}_i = u_i \lambda_i \tag{5.22}$$

Mivel a λ_i együtthatók az εuv függvény lokális maximumai, ezért azt várjuk, hogy jó részük nagy lesz. Mégis az fog bekövetkezni, hogy az utolsó egy-két konstans nem különbözik szignifikánsan zérustól. Ezt a korrelációs együtthatókra vonatkozó szignifikancia vizsgálattal dönthetjük el, mivel ezek a kanonikus változók közönséges korrelációs együtthatói. (A szignifikancia szint függ a mintavétel számától.) Azokhoz a kanonikus korrelációs együtthatókhoz tartozó kanonikus változókra nem is írhatjuk fel az (5.22) egyenletet, ezek részvétele a becslésben nem „jogos" (nem írhatunk fel sztochasztikus összefüggést korrelálatlan változók között).

Tegyük fel, hogy az első $r(r < q)$ számú kanonikus korrelációs együttható szignifikánsan különbözik zérustól. Az a számítás, hogy ezekkel (5.22) alakú egyenleteket írunk fel — a függő változók szempontjából —, ekvivalens az (5.18) vagy (5.19) egyenletekkel végrehajtott becslésekkel, ahol most már csak az első r számú kanonikus változó vesz részt a becslésben. A két egyenlet formailag változatlan marad, csupán a mátrixok mérete változik meg:

$$\Lambda(r \times r),\ A(p \times r) \text{ és } B(q \times r)$$

méretű.

Az eredményül kapott $\hat{Y}$ mátrix rangja r lesz, ezért $q-r$ számú változó kifejezhető r számú lineárisan független lineáris kombinációjaként. Így a vizsgálatban szereplő változók nem alkothatnak szimultán egyenletrendszert.

Most úgy tűnik, hogy egy „keserű lemondást" kellett végrehajtanunk. El kellett ejtenünk olyan információk felhasználását, amelyet a korábbi számítási módszerben figyelembe vettünk. Ez azonban nagyon leegyszerűsített értékelés. Sztochasztikus egyenleteinkkel történő előrebecslésnél nagyon fontos, hogy azt kizárólag megbízható információ felhasználásával végezzük, különben előrebecslésünk sem lesz megbízható. Másrészt a két módszer összehasonlítása nem is mutatja meg az eljárás igazi előnyeit. Gondoljunk el ugyanis egy olyan példát, amelyben 2 függő változó és 3 független változó közötti összefüggést közelítjük a hagyományos módszerrel. Legyen a kapott egyenlet:

$$y_1 = x_1\alpha_1 + x_2\alpha_2 + x_3\alpha_3$$

$$y_2 = x_1\beta_1 + x_2\beta_2 + x_3\beta_3.$$

Ha most a függő változókhoz hozzávesszünk egy újabb változót, y_3-t, a független változókhoz meg mondjuk újabb kettőt, x_4 és x_5-t és ezeken már a kanonikus korrelációs módszerrel hajtunk végre becslést, akkor ez a módszer figyelmen kívül hagyja az x_1, x_2, x_3-ban rejlő információ értéktelen részét, és felcseréli azt az x_4, x_5 hasznos információjával. Így y_1, y_2 és y_3 elörejelzése megbízhatóbb lesz.

Amennyiben lehetőségünk van igen nagy (100—1000) elemű minta vételére, a változók számát is jelentősen növelhetjük. A változók számára az szab határt, hogy kovariancia mátrixuknak, S-nek a gyakorlatban zérustól jelentősen el kell térnie.

6. A korrelációs együttható általánosítása a kanonikus korreláció elméletének segítségével

A p komponensű x és a q komponensű y vektorok közötti sztochasztikus kapcsolat mérésére alkalmas mutatót keresünk. Megkívánjuk tőle, hogy 0 és 1 közé eső valós szám legyen. Azonkívül a korrelációs együttható általánosításaként legyen felfogható, így a $q = 1$ esetben egyezzen meg a többszörös korrelációs együtthatóval.

Induljunk ki megint abból a problémából, hogy az y vektort a legkisebb négyzetek módszerével becsülni akarjuk az x lineáris függvényeként. (1.1) és (1.2) a mintavételi mátrixok. Az egyenlet:

$$Y = XP + W \tag{6.1}$$

Itt $P = (X'X)^{-1}X'Y$. Y és W mérete $(T\times q)$, X-é $(T\times p)$ és P együttható mátrixé $(p\times q)$. Most nem tesszük fel, hogy $q \leq p$. A W mátrix tartalmazza az úgynevezett zavaró tényezők hatását, amely független az x és y vektortól. A mintavételi mátrixok esetében ez abban nyilvánul meg, hogy W korrelálatlan X-szel és Y-nal:

a) $X'W = X'(Y-XP) = X'Y - X'XP = X'Y - X'X(X'X)^{-1}X'Y =$

$= X'Y - X'Y = 0,$

tehát

$$X'W = 0, \tag{6.2}$$

b) $\hat{Y}'W = P'X'W = P'0 = 0,$ (6.3)

tehát

$$\hat{Y}'W = 0.$$

A most levezetett összefüggések segítségével belátható, hogy Y kovariancia mátrixa $\hat{Y}$ és W kovariancia mátrixainak összege:

$$Y'Y = P'X'XP + W'W \tag{6.4}$$

mert $Y'Y = (XP + W)'(XP + W) = P'X'XP + W'W + W'XP + P'X'W$ és itt a 3. és 4. tag (6.2) szerint zérusmátrix.

(6.4) $q \times q$ méretű mátrix. Az egyenletet komponensekként kiírva:

$$\operatorname{cov}(y_i y_j) = \operatorname{cov}(\hat{y}_i \hat{y}_j) + \operatorname{cov}(w_i w_j) \tag{6.5}$$

$i = 1, \ldots, q$ és $j = 1, \ldots, q$ összes kombinációira. Az egyenlet $i = j$ esetén az i-edik változó varianciáját tartalmazza.

(6.4)-et szorozzunk be balról $(Y'Y)^{-1}$-vel:

$$I = (Y'Y)^{-1}P'X'XP + (Y'Y)^{-1}W'W \tag{6.6}$$

Ebből

$$P'X'XP = Y'X(X'X)^{-1}X'X(X'X)^{-1}X'Y = Y'X(X'X)^{-1}X'Y.$$

Ezt (6.6)-ba helyettesítve:

$$I = (Y'Y)^{-1}Y'X(X'X)^{-1}X'Y + (Y'Y)^{-1}W'W. \tag{6.7}$$

Jelöljük $(Y'Y)^{-1}W'W = D$, így a kanonikus korrelációs együtthatókat meghatározó sajátértékegyenlet:

$$|(I - D) - \lambda^2 I| = 0 \tag{6.8}$$

alakba írható, mert (6.7)-ből $I - D = S_{22}^{-1}S_{21}S_{11}^{-1}S_{12}$.

$Y'Y$, $\hat{Y}'\hat{Y}(= |Y'X(X'X)^{-1}X'Y|)$ és $W'W$ mátrixok felfoghatók úgy, mint az egyváltozós variancia általánosításai. Ugyanígy az $I - D$ és D mátrixok pedig úgy tekinthetők, mint a megmagyarázott, illetve a meg nem magyarázott varianciának a függő változó varianciájához való aránya. A probléma csak az, hogy az $I - D$ és D mátrixokhoz hogyan rendeljünk számértéket. Megfelelő mértékszám lesz a zérustól különböző sajátértékek számtani közepe:

$$(\bar{\lambda})^2 = \frac{1}{q} \sum_{i=1}^{q} \lambda_i^2 \tag{6.9}$$

$$1 - (\bar{\lambda})^2 = \frac{1}{q} \sum_{i=1}^{q} (1 - \lambda_i^2). \tag{6.10}$$

Az így kapott két mértékszám összege 1, értékük 0 és 1 közé esik. $(\bar{\lambda})^2$ csak akkor 0 vagy 1, ha minden sajátérték is az. Amennyiben csak ezt a két mutatót számoljuk, a bonyolult sajátértékszámításra sincs szükség. Egy négyzetes mátrix sajátértékeinek összege ugyanis mindig egyenlő átlós elemeinek összegével, és ez az összeg változatlan marad tetszőleges bázistranszformáció alkalmazása esetén:

$$tr(A) = \sum_{i=1}^{n} a_{ii} \text{ jelölés mellett } tr(A) = tr(C^{-1}AC)$$

ahol C az átmenet mátrixa, mert $tr(AB) = tr(BA)$. Így a kívánt mutatót a következő alakba írhatjuk:

$$(\bar{\lambda})^2 = tr(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}) \tag{6.11}$$

Ha $q = 1$, az Y mátrix mérete $T \times 1$, azaz Y egy T elemű vektor. Az

$$(Y'Y)^{-1}Y'X(X'X)^{-1}X'Y$$

mátrixszorzat egy számkonstanssá válik, amelynek egyetlen sajátértéke van: önmaga.

A $(\bar{\lambda})^2 = \dfrac{Y'X(X'X)^{-1}X'Y}{Y'Y}$ alakba írható, ami éppen y és x többszörös korrelációs együttható négyzete. Így beláttuk azt is, hogy $\bar{\lambda}$ megfelel, mint a korrelációs együttható általánosítása. Könnyen beláthatjuk, hogy a trace-korrelációs együttható számítható akkor is, ha az egyes nullától különböző kanonikus korrelációs együtthatók között egyenlők is vannak.

A paragrafus elején az $Y = XP + W$ mátrix egyenletből indultunk ki. Ökonometriai elemzésekben az ilyen egyenletekben feltételezzük, hogy x nem sztochasztikus változó, y pedig az. Ebből arra következtethetnénk, hogy a (6.11) formula más eredményt ad, ha a két változó szerepét megcseréljük. Vegyük észre azonban, hogy a (3.3) és (3.4) egyenletek, amelyekből a kanonikus korrelációs együtthatókat számoljuk $p-q$ számú zérus gyököt kivéve ekvivalensek: a pozitív gyökök azonosak. Így azok összege is megegyezik:

$$tr(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}) = tr(S_{11}^{-1} S_{12} S_{22}^{-1} S_{21}) \tag{6.12}$$

okoskodásunk tehát független a (6.1) modelltől, csupán az összefüggések jobb megértése kedvéért használtuk fel a gondolatmenetben.

Függelék

A számítás menete

a) A változókat standardizáljuk. Ez az

$$x_i' = \frac{x_i - \bar{x}}{\sigma_x}$$

transzformációit jelenti $i = 1, \ldots, T$-re, és itt

$$\bar{x} = \frac{\Sigma x_i}{T} \quad \text{és} \quad \sigma_x = \sqrt{\frac{1}{T}\Sigma(x_i - \bar{x})^2}.$$

Erre azért van szükség, hogy a varianciák számításánál ne jöjjenek ki túl nagy számok. Amint azt a 4. pontban beláttuk, a kanonikus korrelációs együtthatók értékén ez nem változtat.

b) Kiszámítjuk $|S|$ értékét. Amennyiben $|S| < \delta$, ahol δ a kerekítési hibák nagyságrendje, a számítást nem folytathatjuk a kiválasztott változókkal, mert azok gyakorlatilag lineárisan összefüggőek.

c) Kiszámítjuk az $S_{11}^{-1} S_{12}$, majd az $S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}$ mátrixszorzatot.

d) Kiszámoljuk a mátrix átlós elemeinek összegét:

$$(\bar{\lambda})^2 = tr(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12})$$

e) Kiszámítjuk az $S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}$ nem szimmetrikus mátrix sajátértékeit. Ez eléggé bonyolult számítás, részletes leírását [2]-ben megtaláljuk (275—278. o.).

f) Megoldjuk az $(S_{22}^{-1} S_{21} S_{11}^{-1} S_{12} - \lambda_i^2 I) \beta_i = 0$ homogén lineáris egyenletrendszereket $i = 1, \ldots, r$-re (r az indexe az utolsó zérustól szignifikánsan különböző λ_r-nek).

g) A homogén lineáris egyenletrendszer megoldását a $b_i' S_{22} b_i = 1$, $i = 1, \ldots, q$ feltételek teszik egyértelművé, ezért az f) pontban kiszámított sajátvektorokat beszorozzuk az

$$\frac{1}{\sqrt{\beta_i' S_{22} \beta_i}}$$

konstanssal. A végső megoldás tehát:

$$b_i = \frac{1}{\sqrt{\beta_i' S_{22} \beta_i}} \beta_i \qquad i = 1, \ldots, r\text{-re.}$$

h) Az

$$a_i = \frac{1}{\lambda_i} S_{11}^{-1} S_{12} b_i \qquad i = 1, \ldots, r$$

képletből kiszámoljuk az a_i vektorokat, a_i és b_i vektorok a kanonikus változók konstans együtthatói.

i) Megbecsüljük a függő változókat az $Y = XAA'S_{12}$ mátrix egyenletből.

(Beérkezett: 1972. április 20.)

IRODALOM

1. ANDERSON, J. W.: An introduction to multivariate statistical analysis. New York, 1958. John Wiley and Sons.
2. KREKÓ B.: Mátrixszámítás. Budapest, 1964. Közgazdasági és Jogi Könyvkiadó.
3. JOHNSTON, J.: Economic methods. New York, 1963. McGraw-Hill Book Co.
4. DHRYMES, P. J.: Econometrics: statistical foundations and application. New York, 1970. Harper and Row Publishers.
5. MORRISON, D. F.: Multivariate statistical methods. New York, 1967. McGraw-Hill Book Co.
6. HOOPER, J. W.: Simultaneous equations and canonical correlation theory. Econometrica, Vol. 37. (1969).
7. GLAHN, H. R.: Some relationship derived from canonical correlation theory. Econometrica, Vol. 37. (1969).

ON CANONICAL CORRELATION

Canonical correlation belongs to the multivariate statistical analysis. The closeness of the connection between two probability vector variables can be measured by its help.

Let the two variable be vectors x having p components and y having q components. Suppose $q \leq p$. The linear transforms of the two vectors are as follows: $u = A' x$, $v = B' y$. From among the possible linear transformations those are sought, where the transformed

are uncorrelated and besides the corresponding components of the two vector have maximum correlation, each. According to the theory there is a single solution for the q pairs of variables which meets the above requirements.

As a result the vectors x and y can be transformed into closely correlated pairs of q variables. These pairs of variables, the canonical variables give useful information on the interrelations between the components of y and x. The trace of a specific matrix is an appropriate index number for measuring the closeness of the connection between the vectors x and y, and it can be considered as a generalization of the correlation coefficient, too.

Also a process for estimating the parameters of a multivariate regression model can be obtained from the theory. Practically the dependent variables are estimated here as functions of those canonical variables whose canonical correlation coefficient proved significant.

О КАЛЬКУЛЯЦИИ КАНОНИЧЕСКОЙ КОРРЕЛЯЦИИ

Калькуляция канонической корреляции принадлежит статистическому анализу с многими переменными. При помощи этого метода тесноту связи между двумя вектор-переменными можно измерять.

Пусть будет два вектор-переменного вектор x, с компонентом р и вектор y, с компонентом q. Предполагаем, что $q \leq p$. Линейные трансформации двух векторов: $u = A'x$, $v = B'y$. Среди всех возможных линейных трансформаций мы ищем те, в которых видно, что компоненты новых вектор-переменных, полученные при их помощи не в корреляци друг с другом и кроме этого видно, что корреляции другу другу компонентов векторов имеют максимумы. Согласно теории для пара переменных в номере q имеется точно одно решение, которое удовлетворят вышеупомянутые условия.

В результате векторы x и y можно трансформировать в пара переменных в номер q, которые в тесной корреляции друг с другом. Эти пары переменных — канонических переменных — можно представить полезную информацию о взаимозависимостях компонентов x и y.

Из теории можно получить и процесс для оценки параметров регрессивной модели со многими переменными. Здесь практически зависимые переменные оцениваются как функции тех канонических переменных, коэффициенты канонической корреляции доказываются сигнифактными.