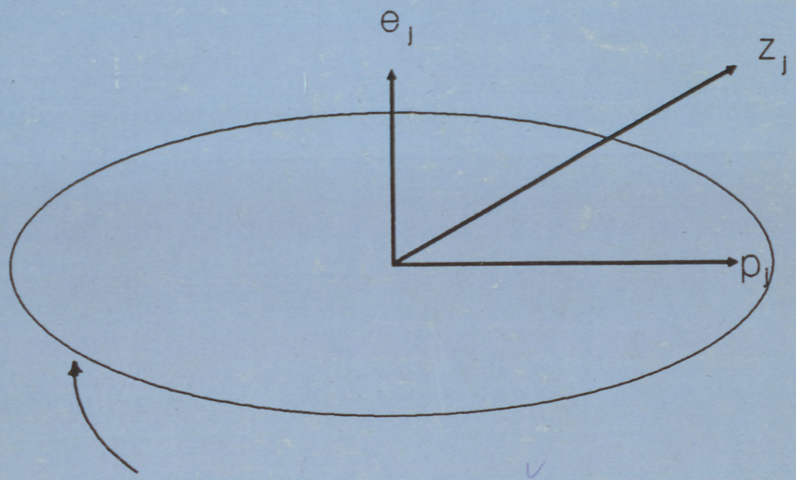
TÁRSADALOMKUTATÁSI MÓDSZERTANI TANULMÁNYOK
III.

Füstös László

Látens változós modellek



OMIKK TÁRKI

TÁRSADALOMKUTATÁSI MÓDSZERTANI TANULMÁNYOK
III.

Sorozatszerkesztő: Kolosi Tamás

TÁRSADALOMKUTATÁSI MÓDSZERTANI TANULMÁNYOK
III.

Füstös László

Látens változós modellek

BUDAPEST, 1991

OMIKK TÁRKI

HU ISSN 0238-4701
ISBN 963 593 120 4
Össz. ISBN 963 593 121 2

Tartalomjegyzék

Előszó

Ezt a munkát Kolosi Tamás kezdeményezésére és ösztönzésére végeztem, hálás köszönettel tartozom érte.

Célom az volt, hogy a szociológiai kutatásokban nagyon gyakran előforduló látens változós modellekről készítsek egy könyvet, amely egyrészt megkísérelné bevezetni ezeket a modelleket — lehetne az is a címe, hogy Bevezetés a látens változós modellekbe — azon kutatók számára, akik érdeklődnek és érdekeltek a kovariancia, vagy korrelációs mátrixok struktúrájának kiértékelésében, másrészt a látens modellek vizsgálatával hozzájárulna ezen modellek adekvát társadalomtudományi alkalmazásaihoz.

A könyv tartalmaz olyan széles körben ismert eljárást, mint pl. a faktorelemzést és tárgyal kevesek által ismert és alkalmazott modelleket is. A könyv feltételezi, hogy az olvasó jártas a statisztika alapvető fejezeteiben, és nem okoz számára gondot a mátrixalgebra sem.

Szeretnék köszönetet mondani tanáraimnak, Meszéna Györgynek és Simonné Mosolygó Nórának, akik nemcsak mint tanárok, hanem mint munkatársak is a legnagyobb befolyással voltak rám, és komoly mértékben hozzájárultak a könyv megírásához.

Szeretném megköszönni Hankiss Elemérnek, hogy éveken keresztül együttdolgozhattunk az Értékszociológiai Mühelyben, és hogy a látens változós modellek értékszociológiai alkalmazásának legnagyobb támogatója és ösztönzője volt.

Köszönettel tartozom Németh Évának a könyv kéziratának nagyon gondos és komoly szakértelmet igénylő (és ezt bizonyító) számítógépes szerkesztéséért.

Köszönöm Reiczigel Jenőnek a gondos korrektúrát, és ha még ezután is maradtak az anyagban hibák, akkor az kizárólagosan a Szerző hibája és felelőssége.

Budapest, 1990.

Füstös László

Látens változós modellek

1. Az általános látens változós modell

Két változóhalmazt különböztetünk meg alapvetően egymástól. Az egyik a közvetlenül megfigyelhető, mérhető, manifeszt változók halmaza: $\underline{x}' = [x_1, x_2, \ldots, \ldots, x_m]$, $\underline{x} \in X$, a másik a közvetlenül nem mérhető, látens változók, faktorok halmaza: $\underline{\xi}' = [\xi_1, \xi_2, \ldots, \xi_r]$, $\underline{\xi} \in \Xi$.

A látens változók száma, r általában lényegesen kisebb, mint a megfigyelt, manifeszt változók száma, m ($r << m$), mivel a látens változós modellek általában adatredukciós módszerek is, melyek a manifeszt változók nagy száma helyett kevés számú látens változóval próbálják a megfigyelt adatokat a lehető legpontosabban reprodukálni.

A megfigyelt változók mérési skáláját, mérési szintjét szokásosan négy kategóriába sorolják, úgymint nominális, ordinális, intervallum és arány mérési skála. Ettől eltérően, a következőkben a feltételezésünkhöz jobban illeszkedő kétosztályos klasszifikációt alkalmazzuk: eszerint a változók vagy metrikusak vagy kategorikusak. A metrikus változó szinonimájaként használjuk a mennyiségi, kvantitatív változó kifejezéseket. A kategorikus változó szinonimájaként használjuk a minőségi, kvalitatív változó kifejezéseket.

A metrikus változók értékei valós számok (vagy a valós számoknak egy részhalmaza), reális, valódi értékek, amelyek lehetnek folytonosak vagy diszkrétek.

A kategórikus változók a megfigyeléseket egymást kizáró, diszjunkt csoportokba, kategóriákba sorolják. A kategóriák lehetnek rendezettek is.

A látens változós módszereket a változók mérési skálája és közvetlen megfigyelhetősége szerint a következőképpen klasszifikálhatjuk:

		Manifeszt változók	
		Metrikus	Kategorikus
Látens változók	Metrikus	Faktorelemzés LISREL LVPLS	Látens tulajdonság-elemzés Faktorelemzés kategorikus adatokkal
	Kategorikus	Látens profil-elemzés	Látens osztályelemzés

Mind a megfigyelt, mind a látens változók valószínűségi változók, így a kapcsolataikat valószínűségeloszlással jellemezhetjük.

A látens változós modelleknél feltételezzük, hogy létezik a megfigyelt változóknak a látens változókra vonatkozó feltételes együttes sűrűségfüggvénye:

$$g(\underline{x}|\underline{\xi})\,. \tag{1}$$

Ha a manifeszt változók diszkrétek, akkor g a feltételes valószínűségek halmaza. Ha a látens változók sűrűségfüggvénye $h(\underline{\xi})$, a megfigyelt változók (feltétel nélküli) együttes sűrűségfüggvényét a következőképpen írhatjuk:

$$f(\underline{x}) = \int h(\underline{\xi})g(\underline{x}|\underline{\xi})d\underline{\xi}\,. \tag{2}$$

A látens változós modellekben az érdeklődés a látens változók manifeszt változókra vonatkozó feltételes eloszlására irányul, vagyis arra, hogy mit tudunk a látens változókról a megfigyelt változók ismeretében:

$$h(\underline{\xi}|\underline{x}) = h(\underline{\xi})g(\underline{x}|\underline{\xi})/f(\underline{x})\,. \tag{3}$$

Ahhoz, hogy a $h(\underline{\xi}|\underline{x})$ feltételes sűrűségfüggvényt megkapjuk, ismernünk kell a h, g és f sűrűségfüggvényeket, de csak az f sűrűségfüggvényt tudjuk becsülni. A (2) egyenletből h és g nem határozható meg egyértelműen, ezért pótlólagos feltételezést kell tennünk. Ez a pótlólagos feltételezés a látens változós modelleknél a *megfigyelt változók látens változókra vonatkozó feltételes függetlensége*:

$$g(\underline{x}|\underline{\xi}) = \prod_{i}^{m} g_i(x_i|\underline{\xi})\,. \tag{4}$$

A manifeszt változók feltételes függetlensége azt fejezi ki, hogy a megfigyelt változók közötti kapcsolatok a megfigyelt változóknak a látens változóktól való függőségéből származnak, vagyis a látens változók idézik elő a megfigyelt változók közötti kapcsolatokat. Így ha a látens változók hatásait kontrolláljuk, akkor a megfigyelt változók függetlenekké válnak. A feltételes függetlenség axiómáját felhasználva a megfigyelt változók együttes sűrűségfüggvényét a következőképpen írhatjuk:

$$f(\underline{x}) = \int h(\underline{\xi}) \prod_{i}^{m} g_i(x_i|\underline{\xi})d\underline{\xi}\,. \tag{5}$$

$f(x)$ függ a h és g_i függvények mellett a látens változók számától, r-től is. A gyakorlatban szeretnénk a lehető legkisebb r érték mellett megkeresni az (5) egyenletet adekvátan reprezentáló h és g_i függvényeket.

Az előzőek bemutatására tekintsünk egy egyszerű számpéldát (forrás D. J. BARTHOLOMEW (1987)). Legyenek A és B dichotom változók, és legyen a 2×2-es gyakoriságtáblázat a következő:

	A	$\bar{A}$	
B	350	200	550
$\bar{B}$	150	300	450
	500	500	1000

Tételezzük fel, hogy az A és B változók között megfigyelt asszociáció egy harmadik dichotom változó (C) kontrollálásával megszűnik, vagyis a C változó eredményezi az A és B közötti kapcsolatot. Osszuk a C változó két kategóriája szerinti két almintára a megfigyeléseket, és nézzük meg az almintákban A és B együttes gyakoriságeloszlását:

C	A	$\bar{A}$	
B	320	80	400
$\bar{B}$	80	20	100
	400	100	500

$\bar{C}$	A	$\bar{A}$	
B	30	120	150
$\bar{B}$	70	280	350
	100	400	500

A két résztáblázatban az A és B változók közötti asszociáció megszűnt, így a teljes mintában megfigyelt kapcsolódás hamis asszociációnak tekinthető, amit a C változó idézett elő.

A következőkben röviden bemutatjuk a két legismertebb látens változós modellt.

1.1 Bináris manifeszt változók és egy bináris látens változó

Tegyük fel, hogy $x_1, x_2, \ldots, x_m$ a bináris megfigyelt változókat jelöli, ahol $x_i = 0$ vagy $x_i = 1$ minden i-re. Tételezzük fel, hogy a megfigyelt változók asszociációit egy ξ bináris változó magyarázza. A bináris látens ξ valószínűségi változó karakterisztikus eloszlása:

$$h(1) = P(\xi = 1) = \eta \quad \text{és} \quad h(0) = P(\xi = 0) = 1 - \eta\,. \tag{6}$$

A megfigyelt változók feltételes eloszlása:

$$g_i(x_i|\xi) = P(x_i|\xi) = \pi_{i\xi}^{x_i}(1 - \pi_{i\xi})^{1-x_i}, \quad (x_i, \xi \in \{0,1\})\,. \tag{7}$$

Miután ebben az egyszerű esetben a h és g_i eloszlások adottak, csak a η, $\{\pi_{i0}\}$, és a $\{\pi_{i1}\}$ paramétereket kell becsülni.

Az $f(\underline{x})$ függvény:

$$f(\underline{x}) = \eta \prod_{i=1}^{m} \pi_{i1}^{x_i}(1-\pi_{i1})^{1-x_i} + (1-\eta)\prod_{i=1}^{m} \pi_{i0}^{x_i}(1-\pi_{i0})^{1-x_i}. \tag{8}$$

A modell paramétereit becsülhetjük a maximum likelihood módszerrel (lásd részletesen később a megfelelő fejezetben), és a modell illeszkedését a megfigyelt együttes gyakoriságokhoz statisztikailag tesztelhetjük. Ha az illeszkedés nem elég jó, akkor próbálkozhatunk három vagy több látens osztállyal vagy folytonos látens változóval.

Elfogadhatóan illeszkedő modell estén a megfigyelési egységeket a manifeszt változók mérési adatai alapján sorolhatjuk be az egyes látens osztályokba az *a posteriori* valószínűségek alapján. Az *a posteriori* valószínűségeket a feltételes sűrűségfüggvény $h(\underline{\xi}|\underline{x})$ segítségével határozzuk meg. A konkrét esetben:

$$\begin{aligned} h(1|\underline{x}) = P(\xi = 1|\underline{x}) &= \eta \prod_{i=1}^{m} \pi_{i1}^{x_i}(1-\pi_{i1})^{1-x_i} / \\ &/ \left\{ \eta \prod_{i=1}^{m} \pi_{i1}^{x_i}(1-\pi_{i1})^{1-x_i} + (1-\eta)\prod_{i=1}^{m} \pi_{i0}^{x_i}(1-\pi_{i0})^{1-x_i} \right\} = \\ &= 1/\left[1 + \left(\frac{1-\eta}{\eta}\right) \exp \sum_{i=1}^{m} \left\{ x_i \log \frac{\pi_{i0}}{\pi_{i1}} + (1-x_i) \log \frac{1-\pi_{i0}}{1-\pi_{i1}} \right\}\right]. \end{aligned} \tag{9}$$

A besorolási szabály szerint az $\underline{x}$ mérési vektorral rendelkező megfigyelési egységet az 1 látens osztályba soroljuk, ha

$$h(1|\underline{x}) > h(0|\underline{x}),$$

vagy ha

$$\begin{aligned} X = \sum_{i=1}^{m} x_i \{\text{logit } \pi_{i0} - \text{logit } \pi_{i1}\} &> \\ > \sum_{i=1}^{m} \log\{(1-\pi_{i1})/(1-\pi_{i0})\} - \text{logit } \eta, \end{aligned}$$

ahol

$$\text{logit } u = \log\{u/(1-u)\}.$$

A fentiekből látható, hogy a besorolási szabály az x_i-k lineáris függvénye.

1.2 Normális eloszlású változók

A másik általános eset, amikor mind a manifeszt, mind a látens változókról feltételezzük, hogy folytonosak és normális eloszlású valószínűségi változók. Tételezzük fel, hogy a megfigyelt változók ($\underline{x}$) együttes eloszlása többváltozós normális eloszlás $\underline{\mu}$ várható értékkel és $\underline{\Sigma}$ variancia–kovariancia mátrixszal:

$$\underline{x} \sim N_m(\underline{\mu}, \underline{\Sigma}) .$$

Tegyük fel, hogy a látens változók standard normális eloszlásúak:

$$\underline{\xi} \sim N_r(\underline{0}, \underline{I})$$

és

$$\underline{x}|\underline{\xi} \sim N_m(\underline{\mu} + \underline{\Lambda}\underline{\xi}, \underline{\Theta}) , \tag{10}$$

ahol

$\underline{\Lambda}$ a faktorsúlyok ($m \times r$) típusú mátrixa,

$\underline{\Theta}$ a hibakomponensek varianciáinak diagonális mátrixa.

A fentiekből következik, hogy

$$\underline{x} \sim N_m(\underline{\mu}, \underline{\Lambda}\underline{\Lambda}' + \underline{\Theta}) . \tag{11}$$

Amennyiben $m = r$, $\underline{\Lambda}\underline{\Lambda}' + \underline{\Theta}$ pontosan egyenlő lesz $\underline{\Sigma}$-val, azonban $r < m$ esetén ez nem feltétlenül van így.

$\underline{\xi}$ *a posteriori* eloszlását a következőképpen határozhatjuk meg:

$$\underline{\xi}|\underline{x} \sim N_r \left\{\underline{\Lambda}'\underline{\Sigma}^{-1}(\underline{x} - \underline{\mu}), (\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda} + \underline{I})^{-1}\right\} . \tag{12}$$

A fentieket részletesebben tárgyaljuk a faktorelemzés fejezetben.

1.3 Látens változó és mérés

Minden empirikus társadalomtudományi kutatás a vizsgálat által körülhatárolt jelenségek megfigyelésén alapul. Az operacionalizálás transzformálja a megfigyeléseket változókká, és a változók kompozíciója definiálja a mérési skálát (itt ne a mérési skála — nominális, ordinális, intervallum, arány — típusaira gondoljunk), amely összefüggésbe hozható az elméleti koncepciókkal. A mérési skála minőségének kiértékelése meg kell, hogy előzze a hipotézisről alkotott végső konklúziónkat, és a mérés minősége korlátozza is a konklúzió súlyát.

A mérés fontosságát túlhangsúlyozni alig lehet. Pl. HAUSER (1969) megállapítása: "I should like to venture the judgement that it is inadequate measurement, more than inadequate concept or hypothesis that has plagued social researchers and prevented fuller explanations of the variances with which they are confounded".

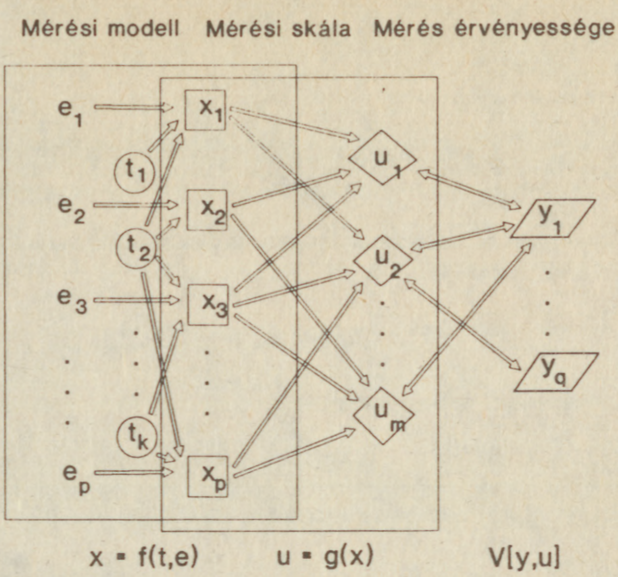
A mérés minőségének két legfontosabb eleme a mérés érvényessége (validity) és a mérés megbízhatósága (reliability). A két komponens nem független egymástól, mivel ha egyszer a mérés nem megbízható, nem lehet érvényes sem.

A megbízhatóság mérése a pszichometriai kutatásokban központi kérdés, ezért nem meglepő, hogy számtalan mutatót dolgoztak ki: split–half, Spearman–Brown–féle, Cronbach–féle, Kuder–Richardson–féle, Armor–féle, Heise– és Bohrnstedt–féle, Tucker–Lewis–féle stb. Keveset tudunk ezeknek a megbízhatósági mutatóknak a tulajdonságairól, alkalmazásaikról, hiszen leginkább a számítógépes program elérhetősége dönti el, hogy melyiket alkalmazzák.

A társadalomtudományokban a legritkább esetben lehetséges az elméleti fogalmakat közvetlenül mérni, így a mérhető indikátorokból kell a mérési skálát összeállítani, mint a manifeszt változók valamilyen függvényét. Keveset tudunk az így képzett mérési skála megbízhatóságáról. Be lehet látni, hogy általában a képzett változó megbízhatósága jobb, mint az item–ek megbízhatósága, de számtalan kérdés merül fel ezzel kapcsolatban.

Tételezzük fel, hogy a megfigyelt változót (x) kifejezhetjük a valódi érték (t) és a mérési hiba (e) függvényeként: $x = f(t, e)$. A mérési skála (u) függvénye a megfigyelt változóknak: $u = g(x)$. Az érdekel bennünket, hogy a valódi érték és a mérési hiba hogyan befolyásolja a mérési skálát (u) a megfigyelt változókon (x) keresztül.

A mérés és a skálázás szkémáját a következő ábrával szemléltetjük:



A két átfedő rész mérési modellnek és mérési skálának nevezhető. Fontos különbséget tenni a kettő között, mivel különböző a mögöttes struktúrájuk. A mérési modellben feltételezzük, hogy a manifeszt változó két látens hatás eredője. Feltételezzük, hogy a látens valódi értékeknek (true scores) létezik valamilyen struktúrája, és hogy befolyásolják a megfigyelt változókat. Ezt a kapcsolatot az $x = f(t, e)$ függvénykapcsolattal fejezhetjük ki. A mérési skála, a jobboldali keret, a megfigyelt változók és a mérési skála függvénykapcsolatát fejezei ki: $u = g(x)$. A g függvényt a skála vagy az operacionalizálás definiálásával adhatjuk meg. Ha létezik a mérés külső kritériuma (y), akkor a validitást a kritérium (y) és a mérési skála (u) közötti korrelációval mérhetjük.

A congeneric mérésnél azt tételezzük fel, hogy az egyes megfigyelt változók ugyanazt a látens tulajdonságot (trait) mérik, csupán a hibatagban különböznek, és bármelyik két változó valódi értéke (látens tulajdonságértéke, szisztematikus komponense) lineáris kapcsolatban van egymással. Jelölje x_i az i-edik változót. A klasszikus mérési modell a következő:

$$x_i = t_i + e_i \qquad i = 1, \ldots, m, \tag{13}$$

ahol

t_i a valódi érték, szisztematikus látens komponens,

e_i a hiba érték, hiba komponens.

A klasszikus mérési modellnél feltételezzük, hogy a valódi és a hiba komponens korrelálatlan

$$\text{cov}(t_i, e_i) = 0,$$

és feltételezzük, hogy a különböző változók hiba komponensei is korrelálatlanok:

$$\mathrm{cov}(e_i, e_j) = 0 .$$

Általában feltételezzük továbbá, hogy

$$E(e_i) = 0 .$$

A congeneric mérésnél feltételezzük, hogy minden valódi érték lineárisan függ egy látens valószínűségi változótól, τ-tól:

$$t_i = \mu_i + \beta_i \tau , \tag{14}$$

ahol feltételezzük, hogy $E(\tau) = 0$ és $\mathrm{var}(\tau) = 1$.

A congeneric mérési modell:

$$x_i = \mu_i + \beta_i \tau + e_i , \tag{15}$$

ahol

$E(x_i) = \mu_i$.

Az i-edik változó megbízhatósági együtthatója (reliability):

$$\rho_i = \frac{\beta_i^2}{\beta_i^2 + \theta_i^2} , \tag{16}$$

ahol

$\mathrm{var}(t_i) = \beta_i^2$,

$\mathrm{var}(e_i) = \theta_i^2$.

Ha m változóra írjuk fel a congeneric mérési modellt, akkor a (15) egyenlet a következőképpen írható:

$$\underline{x} = \underline{\mu} + \underline{\beta} \tau + \underline{e} , \tag{17}$$

ahol

$\underline{x}, \underline{\mu}, \underline{e}$ m elemű oszlopvektorok.

A változók variancia–kovariancia mátrixa:

$$\underline{\Sigma} = \underline{\beta \beta'} + \underline{\Theta} , \tag{18}$$

ahol

$\underline{\Theta}$ a hiba varianciák diagonális mátrixa.

A (17) és (18) egyenletek a faktorelemzés alapegyenletei egy közös faktor esetére. A faktorelemzés eljárását alkalmazhatjuk a $\underline{\beta}$ és $\underline{\Theta}$ paraméterek becslésére. A congeneric mérés speciális esetei a párhuzamos és a tau–ekvivalens mérések.

A párhuzamos mérésnél feltételezzük, hogy

$$\beta_1 = \beta_2 = \ldots = \beta_m$$

és

$$\theta_1 = \theta_2 = \ldots = \theta_m ,$$

vagyis a párhuzamos mérésnél azonosak a szisztematikus komponensek varianciái és a hiba komponens varianciái is.

A tau–ekvivalens mérésnél csak a szisztematikus komponensek varianciáinak az egyenlőségét tételezzük fel, a hiba tagok varianciái különbözhetnek:

$$\beta_1 = \beta_1 = \ldots = \beta_m .$$

A tau–ekvivalens mérésnél a változók kovarianciái megegyeznek, de a varianciák különbözhetnek.

Általánosan, ha s különböző congeneric tulajdonságú változóhalmazunk van és $\underline{x}_g$ jelöli a g-edik csoport vektorváltozóját ($\underline{x}_g$ m_g elemű vektor), akkor $\underline{x}_g$ mérési modellje:

$$\underline{x}_g = \underline{\mu}_g + \underline{\beta}_g \tau_g + \underline{e}_g . \tag{19}$$

Ha a látens szisztematikus komponensek korrelálnak egymással, akkor az s változóhalmazt együtt elemezve a következő modellt írhatjuk fel:

$$\underline{x} = \underline{\mu} + \underline{\beta}\underline{\tau} + \underline{e} , \tag{20}$$

ahol

$\underline{x}$ $p = (m_1 + \ldots + m_s)$ elemű vektor,

$\underline{\beta}$ $(p \times s)$ típusú kvázi– diagonális mátrix, ahol a diagonális „elemek" a $\underline{\beta}_g$-k,

$\underline{\tau}$ s elemű vektor, elemei a τ_s látens komponensek,

$\underline{\mu}$ a változók várható értékeit tartalmazza.

Legyen $\underline{\Gamma}$ $\underline{\tau}$ variancia–kovariancia mátrixa. A megfigyelt változók variancia–kovariancia mátrixa:

$$\underline{\Sigma} = \underline{\beta}\underline{\Gamma}\underline{\beta}' + \underline{\Theta}_\varepsilon , \tag{21}$$

ahol

$\underline{\Theta}_\varepsilon$ a hiba varianciák p-edrendű mátrixa.

A (20) és (21) egyenletek formálisan megegyeznek a későbbi fejezetekben részletesen tárgyalt faktorelemzés modelljével, feltételezve q korrelált faktort, amelyek a diszjunkt változóhalmazokhoz tartoznak. A modellt a faktorelemzés becslési eljárásaival illeszthetjük az adatokhoz.

A szisztematikus látens komponensek korrelálnak egymással, (feltételezésünk szerint) így kifejezhetjük őket egy közös szisztematikus faktormodellel:

$$\underline{\tau} = \underline{\Lambda}\underline{\xi} + \underline{u}, \tag{22}$$

ahol

$\underline{\xi}$ a közös faktorok r-elemű vektora,

$\underline{u}$ az egyedi faktorok vektora (q–elemű).

Legyen $\underline{\Phi}$ $\underline{\xi}$ variancia–kovariancia mátrixa, és $\underline{\Psi}$ az egyedi faktorok diagonális varianciamátrixa. Ekkor a $\underline{\tau}$ szisztematikus látens komponensek variancia–kovariancia mátrixát a következőképpen írhatjuk:

$$\underline{\Gamma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Psi}. \tag{23}$$

A (23) egyenletet behelyettesítve a (21) egyenletbe:

$$\underline{\Sigma} = \underline{\beta}(\underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Psi})\underline{\beta}' + \underline{\Theta}_{\epsilon}. \tag{24}$$

A (24) egyenlet a másodrendű faktorelemzés alapegyenlete, ahol $\underline{\tau}$ az elsőrendű faktorokat, $\underline{\xi}$ a másodrendű faktorokat tartalmazza.

2. Az exploratív faktorelemzés módszerei

A faktorelemzés abból a klasszikus méréselméleti feltételezésből indul ki, hogy a megfigyelt változók kifejezhetők két komponens, a szisztematikus vagy közös és a hiba komponens lineáris függvényével. Szimbolikusan:

$$\underline{Z} = \underline{C} + \underline{E}, \tag{1}$$

ahol

$\underline{Z}$ a megfigyelt változók standardizált mátrixa ($(n \times m)$ típusú, ahol n a megfigyelési egységek, m a változók száma),

$\underline{C}$ a közös komponensek mátrixa ($n \times m$ típusú),

$\underline{E}$ a hiba komponensek mátrixa ($n \times m$ típusú).

Feltételezzük, hogy a két komponens korrelálatlan egymással ($\underline{C}'\underline{E} = \underline{E}'\underline{C} = \underline{0}$), valamint hogy a hiba komponensek függetlenek. Ez utóbbi feltételezésből következik, hogy a hibák variancia–kovariancia mátrixa diagonális ($\underline{E}'\underline{E}/n = \underline{U}^2$).

A közös komponensek $\underline{C}$ mátrixát a faktorok lineáris kombinációjával fejezzük ki:

$$\underline{C} = \underline{F}\,\underline{\Lambda}', \tag{2}$$

ahol $\underline{F}$ a faktortérértékek mátrixa ($n \times r$ típusú, ahol n a megfigyelések száma, r a közös faktorok száma),

$\underline{\Lambda}$ a faktorsúlyok mátrixa ($m \times r$ típusú).

A harmadik feltételként általában előírjuk, hogy a faktorok legyenek függetlenek ($\underline{F}'\underline{F}/n = \underline{I}$). Ez a feltétel azonban nem tartozik a modell alapfeltételeihez, így sokszor megengedjük, hogy a faktorok korreláljanak egymással. Ekkor $\underline{F}'\,\underline{F}/n = \underline{\Phi}$. A továbbiakban azonban az $\underline{F}'\,\underline{F}/n = \underline{I}$ feltételt tartjuk érvényesnek. A faktormodell három feltétele alapján a korrelációs mátrixot a következőképpen bonthatjuk fel:

$$\begin{aligned} \underline{R} &= \underline{Z}'\underline{Z}/n \\ &= \underline{C}'\underline{C}/n + \underline{U}^2 \\ &= (\underline{F}\,\underline{\Lambda}')'(\underline{F}\,\underline{\Lambda}')/n + \underline{U}^2 \\ &= \underline{\Lambda}\,\underline{\Lambda}' + \underline{U}^2, \end{aligned} \tag{3}$$

ahol

$\underline{R}$ a megfigyelt változók páronkénti korrelációs együtthatóit tartalmazó mátrix ($m \times m$ típusú).

THURSTONE (1947) a (3) egyenletet a faktorelemzés alapegyenletének nevezte.

Láthatjuk, hogy a megfigyelt változók közötti korrelációkat reprodukálni tudjuk a faktorsúlyokkal, és a változók varianciáinak (a diagonális elemeknek) a faktorokkal nem magyarázott része pedig a hiba varianciákkal egyenlő. Az egyes változók varianciájának a közös komponensekkel (faktorokkal) megmagyarázott részét kommunalitásnak nevezik. Rendezzük a kommunalitásokat a $\underline{H}^2$ diagonális mátrix megfelelő elemeibe, ekkor

$$\underline{H}^2 = \underline{I} - \underline{U}^2 , \tag{4}$$

vagy

$$\underline{U}^2 = \underline{I} - \underline{H}^2 .$$

A faktorelemzésnél a korrelációs mátrixból indulunk ki, és keressük a faktorsúlyok mátrixát. Ha az $\underline{U}^2$ mátrix ismert, akkor a $\underline{\Lambda}$ mátrixot az $\underline{R} - \underline{U}^2 = \underline{\Lambda\Lambda}'$ egyenletből kiindulva határozzuk meg. Matematikailag a feladat egy sajátérték–sajátvektor problémára vezethető vissza (RUMMEL, 1970). Az egyenlet:

$$(\underline{R} - \underline{U}^2)\underline{Q} = \underline{A}\,\underline{Q} . \tag{5}$$

A $\underline{Q}$ jelöli $\underline{R} - \underline{U}^2$ sajátvektorát, s $\underline{A}$ a megfelelő sajátértékeket, ezek alapján a faktorsúlyok mátrixa: $\underline{\Lambda} = \underline{Q}\underline{A}^{1/2}$. A faktorsúlyokat mint a megfigyelt változók ($\underline{Z}$) és a nem megfigyelt faktorok ($\underline{F}$) közötti korrelációs együtthatókat értelmezhetjük, ha feltesszük a faktorok korrelálatlanságát, továbbá igaz, hogy $\underline{\Lambda}'\underline{\Lambda}$ egyenlő a sajátértékek diagonális mátrixával. Ez azt jelenti, hogy egy adott faktor súlyainak a négyzetösszege (a sajátérték) az adott faktor hozzájárulását adja az összvarianciához (a sajátértékek összege egyenlő a megfigyelt változók korrelációs mátrixának a nyomával — trace ($\underline{R}$) –, és ez egyenlő a változók számával). A faktorok fontosságát szokás ezért a sajátértékek relatív, az összvarianciához viszonyított mértékével is kifejezni.

A (4) azonosságra visszatérve láthatjuk, hogy

$$\underline{R} - \underline{U}^2 = \underline{R} - \underline{I} + \underline{H}^2 = \underline{\Lambda}\,\underline{\Lambda}' ,$$

vagyis a faktorsúlyok mátrixának és saját transzponáltjának szorzata egyenlő a korrelációs mátrixszal, ha a diagonális elemeket kicseréljük a kommunalitásokkal.

Az előzőekben feltételeztük, hogy az $\underline{U}^2$ mátrix ismert. A gyakorlatban a hiba varianciákat nem ismerjük, hanem becsüljük őket. Attól függően, hogy a kommunalitásokat milyen módon becsüljük, különböző faktorelemző eljárásokat különböztetünk meg.

2.1 A főkomponenselemzés

A főkomponenselemzésnél elkerüljük a kommunalitások problémáját, és $\underline{R} - \underline{U}^2$ dekompozíciója helyett az $\underline{R}$ mátrixot magyarázzuk a faktorsúlyokkal. A legfőbb jellemzője a főkomponenseknek, hogy mindegyik komponens a lehető legnagyobb mértékben járul hozzá a megfigyelt változók varianciájához. A faktorsúlyokat a következő homogén egyenlet megoldása alapján kapjuk:

$$(\underline{R} - \alpha \underline{I})\underline{q} = \underline{0}, \tag{6}$$

ahol α az $\underline{R}$ mátrix sajátértéke, az $\underline{R}$ mátrix sajátvektorát pedig $\underline{q}$ jelöli.

Az első sajátértékhez tartozó faktor magyarázza a legnagyobb részét a megfigyelt változók varianciájának, a sorrendben következő faktorok csökkenő mértékben járulnak hozzá az összvarianciához.

A faktorsúlyokat a sajátértékek és a sajátvektorok alapján számítjuk:

$$\underline{\Lambda} = \underline{Q}\underline{A}^{1/2}. \tag{7}$$

A változók számával megegyező számú főkomponens pontosan reprodukálja a korrelációs mátrixot (a diagonális elemeket is). A gyakorlatban azonban az 1-nél kisebb sajátértékekhez tartozó főkomponenseket elhagyjuk, így a reprodukált korrelációs mátrix általában csak közelíti a megfigyelt korrelációs mátrixot, de nem lesz egyenlő vele. Az r számú főkomponenshez ($r < m$) az első r legnagyobb sajátérték tartozik, így az első r számú faktor a lehető legnagyobb, de csökkenő mértékben járul hozzá a változók kommunalitásához.

A főkomponenselemzés alkalmazása akkor indokolt, ha a megfigyelt változók száma nagy, és az első főkomponensekhez tartozó sajátértékek kiugróan magasak.

2.2 A főfaktorok módszere

A faktorelemzésnél széles körben használatos az a megoldás, amelyet a főfaktorok módszerének neveznek. Ez abban különbözik a klasszikus főkomponenselemzéstől, hogy a korrelációs mátrix diagonális elemeit kicseréljük a kommunalitások becsléseivel (ezek legtöbbször az egyes változók többi változóra vonatkozó többszörös korrelációs együtthatójának négyzetei, de lehetnek a többi változókkal mért korrelációs együtthatók közül a maximálisaknak az abszolút értékei is), és az így redukált korrelációs mátrixra alkalmazzuk a főkomponenselemzés módszerét. Ennek a módszernek van iteratív változata, amikor a kommunalitásokat az iteráció egyes lépéseiben kicseréljük az adott lépésbeli becslésekkel:

$$\underline{U}^2 = \mathrm{diag}\,(\underline{R} - \underline{\Lambda}\,\underline{\Lambda}'),$$

és addig folytatjuk az eljárást, amíg a kommunalitások becslései az egyes lépések során már csak elhanyagolható mértékben változnak. Könnyen megmutatható, hogy ez az eljárás ugyanahhoz a faktorsúly mátrixhoz vezet, amelyet az $(\underline{R} - \underline{\Lambda}\,\underline{\Lambda}')$ reziduális mátrix diagonálison kívüli elemei négyzetösszegének minimalizálásával kapnánk.

2.3 Image–elemzés

Tekintsünk m megfigyelt változót, és jelölje z_j $(j = 1, 2, \ldots, m)$ a j-edik megfigyelt változót. Tételezzük fel, hogy a megfigyelt változók standardizáltak, vagyis mindegyik várható értéke 0 és szórása 1.

Becsüljük a megfigyelt változók mindegyikét a többi $(m-1)$ változóval (a legkisebb négyzetek módszerével):

$$p_j = \sum_{k=1}^{m} w_{jk} z_k \,, \tag{8}$$

ahol a w_{jk} standardizált regressziós együttható a többváltozós regressziós egyenletben a k-adik változó hatását mutatja a j-edik változóra. Definíció szerint $w_{jj} = 0$, mivel egy változó önmagára vonatkozó regressziójától eltekintünk.

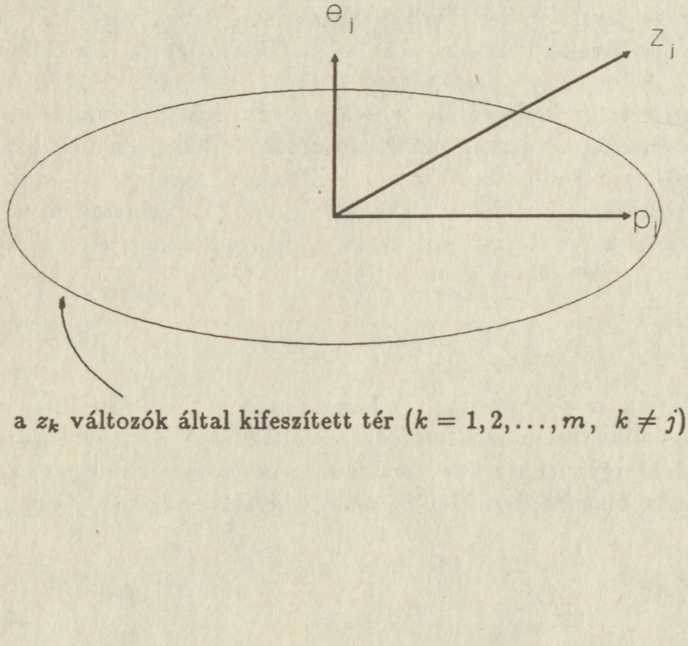
A p_j valószínűségi változót a j-edik megfigyelt változó image-ének (képének) nevezzük. Geometriailag ez a j-edik változónak a többi $(m-1)$ változó terére vonatkozó vetülete. *A* j-edik változó anti–image–e:

$$e_j = z_j - p_j \,, \tag{9}$$

ami a j-edik változónak azon része, amit a többi $(m-1)$ változó nem magyaráz. Az anti–image a többi változó terére ortogonális, vagyis az anti–image független a többi változótól.

A fentieket szemlélteti a következő ábra:

Az image és anti–image geometriai reprezentációja



Általánosságban a (9) egyenletet a következőképpen írhatjuk:

$$\underline{z} = \underline{p} + \underline{e}, \tag{10}$$

ahol

$\underline{z}$ a megfigyelt változók m–elemű oszlopvektora, $\underline{p}$ a megfelelő image-ek vektora, $\underline{e}$ pedig az anti–image-eket tartalmazza. A (10) egyenletet nevezzük az image–elemzés kiindulási feltételének.

Az m számú megfigyelt változó image-einek vektorát az (8) egyenlet alapján a következőképpen határozzuk meg:

$$\underline{p} = \mathbf{Wz}, \tag{11}$$

ahol $\mathbf{W}$ a többváltozós regressziós együtthatók mátrixa, amelynek minden sora a megfelelő változó regressziós becsléseinek együtthatóit tartalmazza. A regresszióelmélet szerint az együtthatók mátrixát a következő képlet alapján számítjuk:

$$\mathbf{W} = \underline{I} - \underline{S}^2 \underline{R}^{-1}, \tag{12}$$

ahol

$\underline{R}$ a megfigyelt változók korrelációs mátrixa, és

$$\underline{S}^2 = [\mathrm{diag}\underline{R}^{-1}]^{-1} \tag{13}$$

a becslések hibájának variancia mátrixa (diagonális mátrix), vagy másképpen az anti–image varianciák diagonális mátrixa.

A fentiekben feltételeztük, hogy a megfigyelt változók korrelációs mátrixa ($\underline{R}$) nemszinguláris.

A megfigyelt változók image-einek variancia–kovariancia mátrixát a (11) és (12) egyenletek alapján a következőképpen írhatjuk fel:

$$\begin{aligned} \underline{G} = E(\underline{pp}') &= (\underline{I} - \underline{S}^2\underline{R}^{-1})\underline{R}\,(\underline{I} - \underline{R}^{-1}\underline{S}^2) \\ &= \underline{R} + \underline{S}^2\underline{R}^{-1}\underline{S}^2 - 2\underline{S}^2 \end{aligned} \tag{14}$$

Hasonlóan határozhatjuk meg az anti–image-ek variancia–kovariancia mátrixát ($\underline{Q}$) a fenti egyenletek alapján:

$$\underline{Q} = E(\underline{ee}') = \underline{S}^2\underline{R}^{-1}\underline{S}^2. \tag{15}$$

Vegyük észre a (13) egyenlet alapján, hogy

$$\mathrm{diag}\underline{Q} = \underline{S}^2, \tag{16}$$

vagyis az anti–image varianciák az $\underline{S}^2$ diagonális mátrix diagonálelemei. A (14) és (15) egyenleteket kombinálva jutunk az image–elemzés alapegyenletéhez:

$$\underline{R} = \underline{G} - \underline{Q} + 2\underline{S}^2 , \tag{17}$$

ami azt fejezi ki, hogy a megfigyelt változók korrelációs mátrixát felbonthatjuk az image és anti–image kovarianciáival összefüggő részekre.

GUTTMAN (1956) a (8) és (9) egyenleteket parciális image-nek és parciális anti–image-nek nevezte, megkülönböztetve azon elméleti hipotetikus változóktól, amelyek a változók általános terében léteznek, ha $m \to \infty$.

Így a j-edik változó általános image–e:

$$\pi_j = \lim_{m\to\infty} \sum_{k=1}^{m} \omega_{jk} z_k \qquad (\omega_{jj} = 0) , \tag{18}$$

ahol

ω_{jk}-k az általános regressziós együtthatók.

Hasonlóan az általános anti–image definíciója:

$$\varepsilon_j = z_j - \pi_j . \tag{19}$$

GUTTMAN (1956) bebizonyította, hogy ha

$$\lim_{m\to\infty} \frac{r}{m} = 0 , \tag{20}$$

akkor az image–elemzés $m \to \infty$ esetén (az általános image–elemzés) és a faktor–elemzés megegyeznek.

Részletezve:

$$p_j \to \pi_j = c_j , \tag{21}$$

ahol

c_j a z_j változó közös faktorok által megmagyarázott része.

Hasonlóan az anti–image faktoranalitikus megfelelője:

$$e_j \to \varepsilon_j = y_j , \tag{22}$$

ahol

y_j a z_j változó egyedi része, az a rész, amit a közös faktorok nem magyaráznak ($z_j = c_j + y_j$).

A faktorelemzés feltételezi, hogy a megfigyelt változók közös faktorai által nem magyarázott egyedi részek közötti, valamint a közös faktorok és az egyedi részek közötti kovarianciák nullák:

$$q_{jk} = \mathrm{Cov}(e_j, e_k) \to \mathrm{Cov}(\varepsilon_j, \varepsilon_k) = \mathrm{Cov}(y_j, y_k) = 0 \qquad (j \neq k)\,, \tag{23}$$

$$\mathrm{Cov}(p_j, e_k) \to \mathrm{Cov}(\pi_j, \varepsilon_k) = \mathrm{Cov}(c_j, y_k) = 0\,. \tag{24}$$

A faktorelemzés modellje szerint érvényes még további három azonosság:

$$g_{jj} = \mathrm{Var}(p_j) \to \mathrm{Var}(\pi_j) = \mathrm{Var}(c_j) = h_j^2\,, \tag{25}$$

$$q_{jj} = \mathrm{Var}(e_j) \to \mathrm{Var}(\varepsilon_j) = \mathrm{Var}(y_j) = u_j^2\,, \tag{26}$$

$$g_{jk} = \mathrm{Cov}(p_j, p_k) \to \mathrm{Cov}(\pi_j, \pi_k) = \mathrm{Cov}(c_j, c_k) = r_{jk} \qquad (j \neq k)\,, \tag{27}$$

ahol h_j^2 a j-edik változó kommunalitása, u_j^2 a j-edik változó egyedisége ($h_j^2 + u_j^2 = 1$) és r_{jk} a j-edik és a k-adik változó közötti korrelációs együttható.

Általánosságban azt mondhatjuk a (19) azonosság feltételezésével, hogy

$$\underline{G} \to \underline{R} - \underline{U}^2\,, \tag{28}$$

valamint

$$\mathrm{diag}\ \underline{Q} = \underline{S}^2 \to \underline{U}^2 \tag{29}$$

$$\underline{Q} \to \underline{U}^2\,, \tag{30}$$

ahol

$\underline{U}^2$ az egyediségek diagonális mátrixa.

A fenti egyenletek alapján azt mondhatjuk, hogy az image kovariancia mátrix jól közelíti a redukált korrelációs mátrixot ($\underline{R} - \underline{U}^2$). A redukált korrelációs mátrix diagonális elemei a megfigyelt változók kommunalitásai, vagyis a változók varianciáinak a közös faktorok által magyarázott része. Az image kovariancia mátrix diagonális elemei az egyes változóknak a többi változóra vonatkozó többszörös korrelációs együtthatójának a négyzetei (ezeket gyakran tekinthetjük a kommunalitások legjobb becsléseinek), a diagonálison kívüli elemek az image kovarianciák, amelyek a megfelelő korrelációktól csak 0-hoz közelítő mértékben különböznek, véges n esetén a különbség egyenlő az anti–image kovarianciákkal.

A közelítést akkor mondhatjuk jónak, amikor az anti–image kovariancia mátrix közel diagonális, a diagonális elemein kívüli elemek a 0-hoz közeliek.

A faktormátrixot (RAO, 1955) kanonikus faktorelemzése alapján határozzuk meg. Rao eljárása, amely azonos megoldást ad LAWLEY (1940) maximum likelihood módszerével, a következő sajátérték–sajátvektor feladatra vezethető vissza:

$$(\underline{U}^{-1}\underline{R}\ \underline{U}^{-1} - \beta\underline{I})\underline{\xi} = \underline{0}\,. \tag{31}$$

Legyen az $\underline{U}^2$ mátrix kezdeti becslése $\underline{S}^2$, ekkor a (31) egyenlet a következőképpen írható:

$$[\underline{S}^{-1}\underline{R}\,\underline{S}^{-1} - b\underline{I}]\underline{x} = \underline{0}. \tag{32}$$

Jelölje $\underline{B}$ az $\underline{R}^* = \underline{S}^{-1}\underline{R}\,\underline{S}^{-1}$ mátrix sajátértékeinek diagonális mátrixát, akkor $\underline{B}$ diagonális elemei (b) a β becslései, $\underline{X}$ pedig az $\underline{R}^*$ egységnyi hosszúságú sajátvektorait tartalmazza (oszlopvektorai a $\underline{\xi}$ becslései).

A főkomponens–elemzés eredményeit felhasználva a (31) egyenletből az $\underline{R}^*$ mátrix faktormátrixa $\underline{XB}^{1/2}$. Ennek alapján magának az $\underline{R}$ mátrixnak a faktormátrixa:

$$\underline{\Lambda}_r = \underline{S}\,\underline{X}\,\underline{B}^{1/2}. \tag{33}$$

Az image kovarianciamátrix $\underline{G}$ faktormátrixát a $\underline{G}^* = \underline{S}^{-1}\underline{G}\,\underline{S}^{-1}$ mátrix sajátértékei és sajátvektorai alapján számítjuk. HARRIS (1962) mutatta meg a (15) és a (17) egyenletek alapján, hogy $\underline{G}^*$ sajátvektorai egyenlőek $\underline{R}^*$ sajátvektoraival, de a sajátértékek különböznek, $\underline{B}$ helyett $(\underline{B} - \underline{I})^2\underline{B}^{-1}$-vel egyenlőek. Így a $\underline{G}$ mátrix faktormátrixa:

$$\underline{\Lambda}_g = \underline{S}\,\underline{X}[(\underline{B} - \underline{I})^2\underline{B}^{-1}]^{1/2}. \tag{34}$$

A két faktormátrixot összehasonlítva látható, hogy közöttük az oszlopvektoraik skálaterjedelmében van különbség, mégpedig $\underline{\Lambda}_g = \underline{\Lambda}_r\underline{D}$, ahol

$$\underline{D} = \underline{I} - \underline{B}^{-1}. \tag{35}$$

Harris mutatott rá arra is, hogy a két faktormátrix megfelelő oszlopai közötti korrelációk megegyeznek. Ezenkívül belátható, hogy az image–faktorok ($\underline{\Lambda}_g$), hasonlóan a kanonikus faktorokhoz, invariánsak az eredeti mértékegységek megváltoztatásával szemben. Fontos rámutatni a $\underline{G}^*$ és az $\underline{R}^*$ mátrixok sajátértékei közötti eltérésre. A $\underline{G}^*$ mátrix $(b-1)^2/b$ sajátértéke és az $\underline{R}^*$ mátrix b sajátértéke közötti függvény monoton növekvő $b > 1$ esetén, azonban ha $b < 1$, akkor a függvény pontosan ellentétesen mozog, monoton csökkenővé válik ($b = 1$ esetén a függvény értéke 0). A sajátértékek ezen ellentétes mozgása miatt az 1-nél kisebb sajátértékű faktorokat ki kell hagynunk az értelmezésből.

2.4 Rao–féle kanonikus faktorelemzés

A Rao–féle kanonikus faktorelemzés során keressük azokat a faktorokat, amelyek maximálisan korrelálnak a megfigyelt változókkal és korrelálatlanok egymással.

Induljunk ki az alapmodellből, abból, hogy a megfigyelt változókat kifejezhetjük a szisztematikus vagy közös komponensek és a hiba vagy egyedi komponensek lineáris függvényeként:

$$\underline{Z} = \underline{C} + \underline{E}, \tag{36}$$

ahol

$\underline{Z}$ a standardizált megfigyelt változók mátrixa ($n \times m$ típusú, ahol n a megfigyelési egységek száma, m a változók száma),

$\underline{C}$ a közös komponensek mátrixa (szintén $n \times m$ típusú),

$\underline{E}$ a hiba komponensek mátrixa (szintén $n \times m$ típusú).

Feltételezzük, hogy a közös és az egyedi komponensek korrelálatlanok, és hogy az egyedi komponensek függetlenek egymástól.

A közös komponenseket a következőképpen fejezzük ki:

$$\underline{C} = \underline{F}\,\underline{\Lambda}', \tag{37}$$

ahol

$\underline{F}$ a nem standardizált faktorértékek mátrixa ($n \times r$ típusú, ahol n a megfigyelési egységek száma, r a faktorok száma),

$\underline{\Lambda}$ a faktorsúlyok mátrixa ($m \times r$ típusú, ahol m a megfigyelt változók száma, r a faktorok száma).

A kanonikus faktor–modellt elsősorban HARRIS (1956, 1962, 1967, 1968) anyagai alapján tárgyaljuk.

Tartalmazza a $\underline{Z}$ mátrix a megfigyelt változók standardizált értékeit súlyozva oly módon, hogy $\underline{Z}\,\underline{Z}'$ a megfigyelt változók közötti páronkénti korrelációs együtthatókat adja! Ha a megfigyelt z változókat és a nem megfigyelt $\underline{F}$ faktorokat egy hipermátrixba foglaljuk, akkor a

$$\begin{pmatrix} \underline{Z}' \\ \underline{F}' \end{pmatrix} [\underline{Z}, \underline{F}] = \begin{pmatrix} \underline{R} & \underline{\Lambda} \\ \underline{\Lambda}' & \underline{I} \end{pmatrix} \tag{38}$$

úgynevezett szupermátrixhoz jutunk.

ANDERSON (1958) adott kanonikus elemzést az ilyen mátrixra. A fenti mátrix esetén a kanonikus korrelációk négyzeteit a következő egyenlet α_i paraméterei adják:

$$|\underline{\Lambda}\,\underline{\Lambda}' - \alpha \underline{R}| = 0\,. \tag{39}$$

Az egyenlet α_i gyökei a megfigyelt változóhalmaz és a nem megfigyelt faktorok halmaza közötti kanonikus korrelációk négyzetei.

Ha $r = m$, vagyis a faktorok száma megegyezik a változók számával, akkor $\underline{\Lambda}\,\underline{\Lambda}' = \underline{R}$, és a determináns minden gyöke egyenlő 1-gyel, ha $\underline{R}$ nemszinguláris. (Ha $\underline{R}$ szinguláris, a rangja kisebb a rendjénél, akkor egy vagy több α_i egyenlő lesz 0-val.)

Ha $r < m$, akkor két lehetőség adódhat. Az első lehetőség az $\underline{R}$ mátrix hiányos faktorelemzése, ebben az esetben a determináns gyökei közül r számú 1-gyel, $(m-1)$ számú pedig 0-val egyenlő. A második esetet kommunalitás típusú megoldásnak nevezik az irodalomban. Ekkor $\underline{\Lambda}\underline{\Lambda}'$ reprodukálja (közelítőleg) $\underline{R}$ diagonálison kívüli elemeit (a korrelációkat), de a diagonális elemeket nem, vagyis $\underline{\Lambda}\,\underline{\Lambda}' = \underline{R} - \underline{U}^2$, ahol $\underline{U}^2$ az egyedi varianciák becsléseinek diagonális mátrixa. Ebben az esetben módosíthatjuk a determináns egyenletet:

$$|\underline{R} - \underline{U}^2 - \alpha\underline{R}| = 0,$$

vagy

$$|\underline{R} - \beta\underline{U}^2| = 0, \tag{40}$$

ahol

$$\beta_i = 1/(1-\alpha_i), \text{ és } \alpha_i = (\beta_i - 1)/\beta_i.$$

Az α_i értékekről feltételezzük, hogy nemnegatív valós számok (kanonikus korrelációk négyzetei), így a β_i-k nagyobbak 1-nél.

Számítási megfontolások miatt Rao transzformálta ezt az egyenletet a következő egyenletté, amelynek gyökei — ahogyan ezt ANDERSON (1958) megmutatta — egyenlőek a transzformált egyenlet gyökeivel:

$$|\underline{U}^{-1}\underline{R}\,\underline{U}{-1} - \beta\underline{I}| = 0. \tag{41}$$

A fenti egyenletben az $\underline{R}$ mátrixról feltételezzük, hogy pozitív definit, $\underline{U}^2$ pedig, mivel diagonális mátrix, melynek elemei 0 és +1 közé esnek, szintén pozitív definit.

Jelölje $\underline{Q}$ az $\underline{U}^{-1}\underline{R}\,\underline{U}^{-1}$ mátrix normalizált sajátvektorait. Ekkor

$$\underline{U}^{-1}\underline{R}\,\underline{U}^{-1} = \underline{Q} < \beta_i > \underline{Q}', \tag{42}$$

amiből

$$\underline{R} - \underline{U}^2 = \underline{U}\,\underline{Q} < \beta_i - 1 > \underline{Q}'\underline{U}. \tag{43}$$

Olyan $\underline{U}^2$ mátrixot kell választanunk, hogy $\underline{R} - \underline{U}^2$ Gram–féle mátrix legyen, akkor

$$\underline{\Lambda} = \underline{U}\,\underline{Q} < \beta_i - 1 >^{1/2}. \tag{44}$$

Az $\underline{U}^2$ mátrix diagonális elemeinek kezdeti becslésére HARRIS (1963) ad eljárást, mely szerint:

$$u_j^2 < 1/r^{jj}, \tag{45}$$

ahol r^{jj} az $\underline{R}^{-1}$ mátrix megfelelő diagonális eleme.

Guttman javaslata, hogy u_j^2 kezdeti értéke $s_j^2 = 1/r^{jj}$ legyen (az egyes változóknak a többi változóra vonatkozó többszörös korreláció négyzete: $1 - 1/r^{jj}$).

HARRIS (1967) javasolta a kezdeti becslésnek a következő módosítását:

$$[v_j^2]' = [r^{jj}]'[(r^{jk})^2]^{-1}, \tag{46}$$

ahol $[(r^{jk})^2]^{-1}$ kvadratikus mátrix, melynek elemei $\underline{R}^{-1}$ megfelelő elemeinek a négyzetei,

$[r^{jj}]'$ az $\underline{R}^{-1}$ mátrix diagonális elemeiből álló sorvektor.

Megmutatható, hogy az egyedi varianciák és fenti kezdeti becsléseik között a következő egyenlőtlenség írható fel:

$$0 < u_j^2 < v_j^2 < s_j^2 < 1\,,$$

vagyis v_j^2 jobb kezdeti becslése az egyedi varianciának, mint a hagyományosan alkalmazott s_j^2.

HARRIS (1963) idézett tanulmányában ismerteti Rao eljárását a kezdeti értékek módosítására és így a kanonikus faktorok becslésére.

A kezdeti értékeket a (41) egyenletbe helyettesítjük, majd megoldjuk a saját-érték–sajátvektor feladatot. A kezdeti értéket a sajátértékek és sajátvektorok felhasználásával a következőképpen módosítjuk :

$$2_j^{u^{-2}} = \sum_{i=1}^{r} ({}_1\beta_i - 1)\,{}_1q_{ji}^2 + 1\,. \tag{47}$$

A (47) egyenlettel módosított kezdeti becsléseket visszahelyettesítjük a (41) egyenletbe, majd megoldva az egyediségek újabb becsléséhez jutunk. Addig folytatjuk az iteratív eljárást rögzített m érték mellett, ameddig a megoldás nem konvergál (.005 tolerancia szinten). Ezután számíthatjuk a (44) egyenlet alapján a faktormátrixot.

2.5 Alfa–faktorelemzés

Az alfa–faktorelemzés a pszichometrikus faktorelemzés kategóriába tartozik abban az értelemben, ahogyan KAISER és CAFFREY (1965) különbséget tett a faktorelemzés statisztikus és pszichometrikus alkalmazása között. (A statisztikus megközelítés a populáció jellemzőire következtet a minta megfigyelési egységei alapján, míg a pszichometrikus megközelítés a változóknak egy mintájából a változók univerzumának jellemzőire következtet.)

Az alfa–faktorelemzésnél feltételezzük, hogy a változóknak létezik egy általános tere, és a vizsgálatba bevont változóhalmaz annak csupán egy véletlen jellegű mintája. Keressük a megfigyelt változóknak azokat a közös faktorait, amelyek a legmagasabb korrelációt mutatják az általános tér megfelelő faktoraival. A megfigyelt változók faktorainak megbízhatóságát (reliability) a Cronbach–féle alfa együtthatóval mérjük, ami a Kuder–Richardson megbízhatósági együttható általánosítása.

A Cronbach–féle alfa együttható a megbízhatósági együtthatók un. felező (split–half) típusához sorolható. Ennél az eljárásnál a vizsgált problémára összegyűjtött változókat két részre bontjuk, így két becsléshez juthatunk. Ezután a

Spearman–Brown formulát alkalmazva megbízhatósági együtthatót számítunk, ami hasonlít a két mérés között számított korrelációhoz. CRONBACH (1951) megmutatta, hogy az alfa együttható egyenlő a felező együtthatók várható értékével.

BENTLER (1968), valamint KAISER és CAFFREY (1965) is definiált egy alfa együtthatót, amelyek különböző eredményekre vezettek. Ezeket a későbbiekben ismertetjük.

Az eljárás bemutatásakor most is a faktorelemzés alapmodelljéből, a $\underline{Z} = \underline{C} + \underline{E}$ egyenletből indulunk ki. A klasszikus méréselmélet a véletlenszerű hiba mellett megkülönböztet szisztematikus hibát is, így a megfigyelt adatokat három hatás eredőjeként értelmezzük: $\underline{Z} = \underline{C} + \underline{E}_s + \underline{E}_e$. Ennek megfelelően a hiba varianciát két részre bontjuk: $\underline{U}^2 = \underline{E}_s^2 + \underline{U}2_e = \underline{I} - \underline{H}^2$, ahol $\underline{H}^2$ a közös komponensek varianciáit, vagy másnéven a kommunalitásokat tartalmazza.

A megbízhatósági együttható a nem véletlenszerű komponensek varianciáinak és a teljes varianciának az aránya, vagy másképpen fogalmazva a szisztematikus komponensek (a valódi — elméleti — értékek) varianciájának és a megfigyelt értékek varianciájának az aránya:

$$r_{zz} = \frac{\underline{1}'(\underline{R} - \underline{U}_e^2)\underline{1}}{\underline{1}'\underline{R}\,\underline{1}}, \tag{48}$$

ahol

$\underline{1}$ az összegzővektort jelöli.

Ezt az együtthatót nevezik Kuder–Richardson–féle együtthatónak. THOMSON (1940) és PEEL (1948) oldották meg először azt a problémát, hogy a változókhoz úgy rendeljenek súlyokat, hogy a megbízhatósági együttható maximális legyen. A megoldásra BENTLER (1968) adott korszerű leírást. Kereste azt a $\underline{w}$ súlyvektort, amely mellett a következő függvény felveszi maximumát:

$$r_M = \frac{\underline{w}'(\underline{R} - \underline{U}_e^2)\underline{w}}{\underline{w}'\underline{R}\,\underline{w}}. \tag{49}$$

A függvénynek $\underline{w}$ szerinti parciális deriváltját véve, azt egyenlővé téve nullával, és némi algebrai átalakítás után (lásd BENTLER (1968), 336–337) a következő sajátérték–sajátvektor egyenlethez jutunk:

$$[\underline{U}_e^{-1}(\underline{R} - \underline{U}_e^2)\underline{E}_e^{-1} - \beta^2\underline{I}]\underline{v} = \underline{0}, \tag{50}$$

ahol

$\underline{v} = \underline{U}_e\,\underline{w}$,

$\beta = r_M/(1 - r_M)$.

A fenti egyenlet megoldása adja azokat a súlyokat, amelyek maximálják a megbízhatósági együtthatót, és ezen egyenlet, amit a főkomponens és faktorelemzésnél jól ismerünk, megoldásainak sajátértékei és sajátvektorai alapján számítjuk a faktorelemzés mátrixait.

Cronbach definiált a megbízhatóság mérésére egy másik együtthatót (CRONBACH, 1951):

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\Sigma V_i}{V_t}\right), \tag{51}$$

ahol

m a változók száma,

V_t a varianciák és kovarianciák összege,

V_i az i-edik változó varianciája.

Bentler bebizonyította, hogy ha érvényes a belső konzisztenciának $\underline{U}_e^2$-re vonatkozó feltétele: $\underline{1}'\underline{U}_e^2\underline{1} = m(1 - r_{ij}^a)$, ahol r_{ij}^a a változók közötti átlagos korrelációt jelöli, akkor az α együttható megegyezik az (48) egyenletben definiált megbízhatósági együtthatóval (a bizonyítást lásd BENTLER (1968), 337–338):

$$r_{zz} = \alpha = \frac{m}{m-1}\left(1 - \frac{\underline{1}'\underline{I}\underline{1}}{\underline{1}'\underline{R}\underline{1}}\right). \tag{52}$$

Maximalizáljuk a Cronbach–féle α-t az (52) egyenletből kiindulva.

$$\alpha = \frac{m}{m-1}\left(1 - \frac{\underline{p}'\underline{I}\underline{p}}{\underline{p}'\underline{R}\underline{p}}\right). \tag{53}$$

Az α együtthatót maximalizálhatjuk

$$\gamma^2 = \frac{\underline{p}'\underline{R}\,\underline{p}}{\underline{p}'\underline{I}\,\underline{p}} \tag{54}$$

maximalizálásával. A szélsőértékszámítást elvégezve a következő egyenlethez jutunk:

$$(\underline{R} - \gamma^2\underline{I})\underline{p} = \underline{0}. \tag{55}$$

Láthatjuk, hogy a Cronbach–féle alfa maximálizálásával a főkomponens elemzéssel megegyező eredményre jutunk.

A főkomponensek számának meghatározásához helyettesítsük be az (54) egyenletet (53)-ba:

$$\alpha = \frac{m}{m-1}\left(1 - \frac{1}{\gamma^2}\right). \tag{56}$$

Ebből láthatjuk, hogy a megbízhatósági együttható akkor lesz 0-nál nagyobb, ha $\gamma^2 > 1$ (a sajátérték nagyobb 1-nél), és ez pedig jól ismert kritérium a főkomponensek számának meghatározásához.

Bentler mutatott rá, hogy a faktorelemzésnél az $(\underline{R} - \underline{U}_e^2)$ mátrix helyett az $(\underline{R} - \underline{U}^2)$ mátrixot vizsgáljuk, és mivel $\underline{R} - \underline{U}^2 = (\underline{R} - \underline{U}_e^2) - \underline{U}_s^2$, a faktorelemzésnél a szisztematikus komponenshez nem vesszük hozzá a specifikus részt, így a közös varianciából kihagyjuk a specifikus varianciát. Ennek megfelelően Bentler javasolt egy mutatót, amely a közös komponens varianciájának és a teljes varianciának az arányát méri:

$$\alpha_0 = \frac{\underline{1}'(\underline{R} - \underline{U}^2)\underline{1}}{\underline{1}'\underline{R}\,\underline{1}}. \tag{57}$$

Ezt a mutatót belső konzisztencia együtthatónak nevezte, és kereste azt a $\underline{w}$ súlyvektort, amely az

$$\alpha_0 = \frac{\underline{w}'(\underline{R} - \underline{U}^2)\underline{w}}{\underline{w}'\underline{R}\,\underline{w}} \tag{58}$$

együtthatót maximalizálja. A maximalizálási feladatot az előzőekhez hasonlóan elvégezve a következő sajátérték–sajátvektor egyenlethez juthatunk:

$$[\underline{U}^{-1}(\underline{R} - \underline{U}^2)\underline{U}^{-1} - \mu^2\underline{I}]\underline{v} = \underline{0}, \tag{59}$$

ahol $\underline{v} = \underline{U}\,\underline{w}$, és

$$\mu^2 = \frac{\alpha_0}{1 - \alpha_0}. \tag{60}$$

A fenti egyenlet alapján a belső konzisztencia együtthatót kifejezhetjük az i-edik sajátérték segítségével:

$$\alpha_{0i} = \frac{\mu_i^2}{\mu_i^2 + 1}. \tag{61}$$

A (61) azonosságból láthatjuk, hogy nem találunk a változóknak olyan súlyozását, amellyel a közös komponenseket pontosan elő tudnánk állítani, vagyis hogy a megbízhatóság tökéletes legyen. A fentiek alapján azt mondhatjuk, hogy annál nagyobb lesz a faktorelemzés megbízhatósága, minél nagyobb lesz az első sajátértékeknek a nagysága, amihez szükséges feltétel, hogy a változók száma, valamint a változók számának és az $(\underline{U}^{-1}(\underline{R}-\underline{U}^2)\underline{U}^{-1})$ sajátérték mátrix rangjának (r) aránya nagy legyen.

A faktorsúlyok mátrixát a sajátértékek és sajátvektorok alapján számíthatjuk most is ki.

KAISER és CAFFREY (1965) definiált a Cronbach–féle α-hoz hasonló, de csak a közös komponensekre koncentráló mutatót:

$$\alpha_c = \frac{m}{m-1}\left(1 - \frac{\underline{w}'\underline{H}^2\underline{w}}{\underline{w}'(\underline{R} - \underline{U}^2)\underline{w}}\right). \tag{62}$$

Kaiser és Caffrey kereste a változóknak azt a súlyozását, amelyik az α_c mutatót maximálizálja. A maximumfeladatot megoldva a következő egyenlethez jutottak:

$$[\underline{H}^{-1}(\underline{R} - \underline{U}^2)\underline{H}^{-1} - \delta^2\underline{I}]\underline{q} = \underline{0}, \tag{63}$$

ahol $\underline{q} = \underline{H}\,\underline{w}$.

A δ_i sajátérték és az α_{ci} közötti összefüggés:

$$\alpha_{ci} = \frac{m}{m-1}\left(1 - \frac{1}{\delta_i^2}\right) . \tag{64}$$

A megbízhatóság a fenti egyenlet alapján kisebb vagy egyenlő lesz 0-val, ha a sajátérték kisebb vagy egyenlő 1-nél. Ezért az alfa–faktorelemzésnél csak azokat a faktorokat vesszük figyelembe, amelyeknél a sajátértékek nagyobbak 1-nél.

Kaiser és Caffrey alfa–faktorelemző eljárása — mint említettük — a megfigyelt változók közös komponenseire koncentrál. Könnyen megmutatható, hogy tulajdonképpen a közös komponensek korrelációs mátrixának faktorelemzését végzi el.

Tegyük fel a jelölések egyszerűsége miatt, hogy a standardizált változók $\underline{Z}$ mátrixát a megfigyelési egységek száma négyzetgyökének reciprokával súlyoztuk, így két változó között a korrelációt $\underline{z}_i\underline{z}_j/n$ helyett egyszerűen $\underline{z}_i\underline{z}_j$ formulával számíthatjuk.

Bontsuk fel két megfigyelt változó korrelációs együtthatóját a közös komponens és a hiba tag alapján:

$$r_{ij} = \underline{z}_i\underline{z}_j = \underline{c}_i'\underline{c}_j + \underline{e}_i'\underline{e}_j = \underline{c}_i'\underline{c}_j ,$$

mivel a hiba tagok függetlenek.

$\underline{c}_i'\underline{c}_j$ a két közös komponens közötti kovarianciával egyenlő. A közös komponensek varianciáját a megfelelő megfigyelt változók kommunalitása adja ($\underline{c}_i'\underline{c}_i = h_i^2$). Ha a kovarianciákat osztjuk a megfelelő varianciákkal, akkor a korrelációs együtthatóhoz jutunk. Mátrixjelölésekkel:

$$\underline{H}^{-1}(\underline{R} - \underline{U}^2)\underline{H}^{-1} ,$$

ami a megfigyelt változók közös komponenseinek a korrelációs mátrixa.

Az alfa–faktorelemzés iteratív eljárását a kommunalitások becslésével kezdjük. A kommunalitások kezdeti becslései a változók többszörös korrelációs együtthatói, amelyeket a korrelációs mátrix megfelelő indexű diagonális elemeivel kicserélünk, és az így módosított mátrixból kiindulva keressük azt a közös faktort, amelynek megbízhatósági együtthatója a legnagyobb. Ezután a korrelációs mátrixot adjusztáljuk, és az eljárást addig folytatjuk, amíg a kommunalitások nem konvergálnak. A faktorok egyre csökkenő megbízhatóságuak, az első faktornak lesz a legnagyobb a megbízhatósági együtthatója, a második faktornak a második legnagyobb, és így tovább.

A faktoroknak ez a tulajdonsága a rotálásnál megváltozik. Az elemzéshez annyi faktort tartunk meg, amennyinek a sajátértéke nagyobb 1-nél.

2.6 Maximum likelihood faktorelemzés

Négy különböző faktorelemző eljárás is a faktormátrix (faktorsúlyok) elemeinek lényegében azonos becsléséhez vezet: LAWLEY (1940, 1941) maximum likelihood módszere, RAO (1955) kanonikus korreláció elméletén alapuló módszere, BARGMANN (1957) és HOWE (1955) maximális determináns módszere, amely a faktorsúlyok becsléséhez maximalizálja a megfigyelt változók feltételes korrelációs mátrixának determinánsát (a megfigyelt változóknak a faktorok kiszűrésével számított parciális korrelációs mátrixának determinánsát), továbbá JÖRESKOG (1966) lineáris strukturális relációkra kidolgozott modellje (LISREL), amely maximum likelihood becslést ad a faktorsúly mátrix elemeire.

Lawley likelihood egyenletei és Rao kanonikus faktor egyenletei azonos identifikációs feltételeket tartalmaznak: hogy $\underline{\Lambda}'\underline{U}^{-2}\underline{\Lambda}$ mátrix diagonális legyen, és hogy azonos egyértelmű megoldást adjanak a faktormátrixra.

A faktorelemzésnek az az általános feltétele, hogy az első faktor járuljon maximális mértékben hozzá a megfigyelt változók varianciáihoz, a második a maradék lehetőségek közül járuljon hozzá az összvarianciához maximális mértékben azzal a feltételezéssel, hogy független legyen az előző faktortól, és így tovább, azzal az identifikációs feltétellel függ össze, hogy $\underline{\Lambda}'\underline{U}^{-2}\underline{\Lambda}$ diagonális legyen (LAWLEY és MAXWELL (1971), 4. fejezet).

A maximális determináns egyenleteknek nincs egyértelmű megoldása, de ortogonális transzformációval megkaphatjuk azt a megoldást, amelyik kielégíti a fenti identifikációs feltételeket.

Lawley eljárásáról HOWE (1955) kimutatta, hogy sokkal lassabban konvergál, mint a maximális determináns módszer. LORD (1956) viszont azt találta, hogy bizonyos esetekben, ha a kezdeti becslés nem elég jó, Lawley eljárása megáll, és a faktorsúlyokra rossz becslést ad. Ezért Lawley eljárását nem használják.

Rao kanonikus faktorelemző eljárását BROWNE (1968) vizsgálta, és tapasztalata szerint az eljárás lassabban konvergál, mint a maximális determináns módszer, és előfordulhat olyan eset, hogy az egymást követő iterációs lépésekben az u_i^2 értékek között csak nagyon kicsi az eltérés akkor, amikor még ezek az értékek a valódi értékektől nagyon messze vannak.

Bargman és Howe maximális determináns módszere a megfigyelt változók közötti korrelációs együtthatók mátrixának determinánsát maximálja, ahol a kontrollváltozók a faktorok.

A parciális korrelációs mátrix:

$$\underline{R}_{(x|f)} = \underline{U}^{-1}(\underline{R} - \underline{\Lambda}\,\underline{\Lambda}')\underline{U}^{-1}\,. \tag{65}$$

A determináns:

$$|\underline{R}_{(x|f)}| = |\underline{R}||\underline{I} - \underline{\Lambda}'\underline{R}^{-1}\underline{\Lambda}||\underline{U}|^2\,. \tag{66}$$

A determináns akkor lesz maximális, ha $\underline{R}_{(x|f)}$ az egységmátrix.

Az eljárásban a kommunalitások kezdeti becslésének a többszörös korrelációs együtthatók becslésének a négyzetét (R_j^2) tekintik. Maximum likelihood becslés esetén:

$$R_j^2 = 1 - 1/r^{jj},$$

ahol r^{jj} az $\underline{R}^{-1}$ inverz mátrix j-edik diagonális eleme.

Ennek alapján az egyedi varianciák kezdeti becslései:

$$u_j^2 = 1/r^{jj},$$

általában

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

Bargman eljárásában az iterációt akkor állítják le, ha az egymás utáni lépésekben az u_i^2-ek eltérésének abszolút értéke nem nagyobb mint .0001.

Annak a nullhipotézisnek a tesztelésére, hogy r számú faktor elégséges a korrelációs mátrix reprodukálásához, a likelihood hányados statisztikát alkalmazzuk (lásd HOWE, 1957):

$$\alpha_r = -k \log_e |\underline{R}_{(x|f)}|. \tag{67}$$

A k multiplikátort BARTLETT (1950) adta meg:

$$k = n - (2m + 11)/6 - (2r)/3.$$

Az α_r statisztika aszimptotikusan $1/2((m-r)^2 - m - r)$ szabadságfokú khi–négyzet eloszlást követ.

LAWLEY (1940) a likelihood hányados statisztikának könnyen számolható közelítő formáját határozta meg:

$$\alpha_r = k \sum_{j=1}^{m} \sum_{i=1}^{j} r_{ij}^{*2},$$

ahol r_{ij}^{*2} a parciális korrelációs mátrix i-edik sorának j-edik eleme.

A likelihood hányados statisztikát $0, 1, 2, \ldots$ számú faktorra egymásután alkalmazzuk egészen addig, amíg a nullhipotézist már nem kell elvetnünk (az adott valószínűségi szint mellett). A faktorok számának becslése az első olyan lépésben levő faktorszám lesz, amelyben a nullhipotézist nem utasítjuk el.

Ha feltételezzük, hogy a megfigyelt változók többváltozós normális eloszlású valószínűségi változók, akkor a minta variancia–kovariancia mátrixa Wishart eloszlást követ $(n-1)$ szabadságfokkal (lásd például MARDIA, KENT és BIBBY (1979)). A faktormodellben a populáció variancia–kovariancia mátrixát ($\underline{\Sigma}$) a következő egyenlettel fejezzük ki:

$$\underline{\Sigma} = \underline{\Lambda}\, \underline{\Lambda}' + \underline{U}^2. \tag{68}$$

A likelihood függvény logaritmusa (ANDERSON (1958)) a megfigyelésektől függő konstans elhagyása után

$$\log_e L = 1/2(n-1)[\log_e |\underline{\Sigma}| + \mathrm{trace}(\underline{S}\,\underline{\Sigma}^{-1})]\,, \tag{69}$$

ahol

$\underline{S}$ a megfigyelt változók mintából becsült variancia–kovariancia mátrixa.

A (69) és a (68) egyenletek alapján látható, hogy a likelihood függvény logaritmusa a faktorsúlyok és a specifikus variancia függvénye. Ezen paramétereket becsülhetjük a likelihood függvény logaritmusának maximalizálásával úgy, hogy $\underline{\Lambda}'\underline{U}^{-2}\underline{\Lambda}$ mátrix diagonális legyen. A gyakorlatban célszerűbb a likelihood függvény logaritmusa helyett a következő függvényt minimalizálni:

$$F = \mathrm{trace}(\underline{S}\,\underline{\Sigma}^{-1}) + [\log_e |\underline{\Sigma}| - \log_e |\underline{S}|] - m\,. \tag{70}$$

Az F függvény minimalizálása equivalens $\log_e L$ maximalizálásával, mivel $F = -(c_1 \log_e L + c_2)$, ahol c_1 és c_2 konstansok.

Ha $\underline{S}$ és $\underline{\Sigma}$ elemei hasonlóak, akkor az inverzeik is hasonlóak. Ezért $\underline{S}\,\underline{\Sigma}^{-1}$ közelít az egységmátrixhoz, ha $\underline{S}$ és $\underline{\Sigma}$ közelít egymáshoz. Mivel egy mátrix nyoma a diagonális elemeinek összege, az első tag m-hez közelít, ha $\underline{S}$ és $\underline{\Sigma}$ közelít egymáshoz. A második tag $\underline{\Sigma}$ és $\underline{S}$ determinánsa logaritmusainak különbsége, ezért a második tag 0-hoz közelít, ha $\underline{\Sigma}$ közelít $\underline{S}$-hez. A harmadik tag konstans, a megfigyelt változók számával egyenlő. Eszerint ha $\underline{S}$ és $\underline{\Sigma}$ egyenlő, az illeszkedést mérő függvény egyenlő 0-val.

Az F függvény minimalizálására JÖRESKOG (1967) javasolt egy sikeres eljárást, amely először a faktorsúlyok szerint minimalizálja az F függvényt a specifikus varianciák rögzített értékei mellett, majd a faktorsúlyok rögzítése mellett minimalizálja F értékét az $\underline{U}^2$ értékei szerint.

2.7 Legkisebb négyzetek módszere

A klasszikus legkisebb négyzetek módszere (unweighted least squares) minimalizálja a megfigyelt és a modell által reprodukált kovarianciák különbségeit:

$$F_{ULS} = \mathrm{trace}[(\underline{S} - \underline{\Sigma})^2]\,. \tag{71}$$

A legkisebb négyzetek módszere $\underline{\Lambda}$ és $\underline{U}^2$ becsléséhez a reziduális variancia–kovarianciákat minimalizálja.

A legkisebb négyzetek módszerének legnagyobb előnye, hogy nem tételez fel a változók eloszlásáról semmit. Hátránya viszont, hogy egyrészt nem ad statisztikai próbát a modell illeszkedésére, másrészt, hogy skála–függő. Ha egy módszer skála–függő, akkor megváltoztatva a megfigyelt változók mértékegységét (skáláját), az illeszkedést mérő függvény minimumhelye is megváltozik, és a becslésekben történő változás csupán a skála változását tükrözi. A legkisebb négyzetek módszere különböző eredményt adhat, ha a mértékegységeket változtatjuk. Ilyen esetben célszerű a változókat standardizálni, így a korrelációs mátrixot (a variancia-kovariancia mátrix helyett) elemezni.

2.8 Általánosított legkisebb négyzetek módszere

Az általánosított legkisebb négyzetek módszere abban különbözik a klasszikus legkisebb négyzetek módszerétől, hogy az $\underline{S}$ és $\underline{\Sigma}$ közötti különbséget súlyozza az $\underline{S}^{-1}$ elemeivel, így ez a módszer már invariáns a skála (mértékegység) megválasztására.

A modell illeszkedését mérő függvény:

$$F_{GLS} = \mathrm{trace}[(\underline{S} - \underline{\Sigma})\underline{S}^{-1}]^2 . \tag{72}$$

Amennyiben a megfigyelt változók együttes valószínűségeloszlása normális, az általánosított legkisebb négyzetek módszere aszimptotikusan ekvivalens a maximum likelihood módszerrel (lásd LEE, 1977, BOWNE, 1974).

2.9 Faktorelemző eljárások összehasonlítása

A faktorelemzés különböző módszereit alapvetően aszerint hasonlíthatjuk össze, hogy az eljárás mit is optimalizál, hogyan definiálja az egyediségeket, hogyan becsüli a kommunalitásokat és a faktorértékeket. Az összehasonlításra GORSUCH (1983) táblázatát mutatjuk be.

1. *Táblázat* Faktorelemző eljárások összehasonlítása

Az eljárás neve	Az eljárás lényege	Az egyediségek definiálása	A kommunalitás becslése	Faktor-értékek
Főkomponens-elemzés	Maximalizálja a magyarázott varianciát	Nincs	Nem szükséges	Számított
Főfaktor-elemzés	Maximalizálja a magyarázott varianciát	Specifikus faktorok, véletlen hiba	Számos becslő eljárás	Becsült
Image-elemzés	Minimalizálja a reziduális image–eket	Minden változó azon része, amely nem korrelál a többi változóval	Többszörös korreláció négyzete	Számított
Alfa–elemzés	Maximalizálja a megbízhatósági együtthatót	Pszichometrikus hiba	Iteratív	Számított
Maximum–Likelihood (Lawley, Jöreskog)	A reprodukált korrelációs mátrix legjobb becslése	Specifikus faktorok, véletlen hiba	Iteratív	Becsült
Maximum–Likelihood (Rao–féle kanonikus faktorelemzés)	A megfigyelt változókkal maximálisan korreláló faktorok	Specifikus faktorok, véletlen hiba	Iteratív	Becsült

A faktorelemző eljárások összehasonlításakor azt kell elsősorban megállapítani, hogy amennyiben a kommunalitások értéke 1.0, a főkomponenselemzés, a főfaktorok módszere, az alfa- és a maximum likelihood eljárás azonos eredményre vezet.

Amennyiben a változók száma növekszik, a korrelációs mátrix elemei között a diagonális elemek aránya, így a diagonális elemek fontossága is csökken, ezért a kommunalitás becslése helyett a diagonális elemeken kívüli korrelációs együtthatók becslése válik fontosabbá, vagyis a maximum likelihood, image– és alfa–faktorelemző eljárás megoldásai pontosabb eredményre vezetnek.

Az empirikus összehasonlításokban TUCKER, KOOPMAN és LINN (1969) azt tapasztalta, hogy a főkomponens– és a főfaktor–eljárás azonos faktorokat eredményezett, amikor minden változónál (20 változó) magas volt a kommunalitás. KALLINA és HARTMAN (1976) nem talált interpretálható különbséget a főkomponens és a főfaktor között. DZIUBAN és HARRIS (1973) más vizsgálatokkal ellentétben nagy különbségeket talált több módszer faktormátrixában, sőt azt a következtetést tették, hogy a főkomponens–elemzés bizonyos feltételek esetén nem elfogadható eljárás. Azonban, ahogyan ezt VELICER (1977) bizonyította, ez a következtetés nem volt helyes, mivel abból adódott, hogy a faktormátrixban túlságosan sok faktor szerepelt (az 1-nél nagyobb sajátértékű faktorok maradtak meg), és ha más teszteket felhasználva, kevesebb faktorral végezték volna el az összehasonlítást, akkor nagyon hasonló eredményekre vezettek volna a különböző módszerek. VELICER (1977) saját vizsgálatában, amelyben a főkomponens image– és a maximum likelihood módszert hasonlította össze, azt a következtetést vonta le, hogy a három eljárás lényegében azonos eredményt ad, ezen belül a maximum likelihood és az image–eljárás eredménye hasonlított legjobban egymáshoz, és a főkomponens és a maximum likelihood között tapasztalt nagyobb különbséget. A főkomponens–eljárásnak nagy előnye az egyszerűség, mind számítási, mind elméleti értelemben. Velicer szerint az image–eljárás adhatja a legelfogadhatóbb alternatívát. BROWNE (1968) a különböző eljárások pontosságát vizsgálta ismert faktorsúly–mátrix esetén, és eredményül azt kapta, hogy a maximum likelihood becslés pontosabb megoldást adott, de ez a pontosság csak nagy mintánál ($n = 1500$) volt igazán érzékelhető.

Az irodalmi tapasztalatok tanúsága szerint (lásd GORSUCH, 1983, 123 p.) általánosságban az állítható, hogy viszonylag nagy (nagyobb mint 30) változószám esetén és viszonylag elfogadható kommunalitások mellett (nem kisebbek mint 0.4), bármelyik exploratív faktorelemző eljárás praktikusan hasonlóan értelmezhető eredményre vezet.

De mit lehet mondani általánosságban, ha a változók száma nem elég nagy? Sajnos azt, hogy nincs valódi alternatíva. Vagy alkalmas hipotézisünk van a közös faktorokról, és alkalmazhatjuk valamelyik közös faktoreljárást, vagy a kommunalitások nagyságában bízva a főfaktor–eljárást.

A főfaktor eljárás a legkisebb négyzetek módszere elve szerinti becslést ad, míg a kanonikus faktorelemző eljárás maximum likelihood eljárás, ha a változók többdimenziós normális eloszlásúak. A maximum likelihood becslés általában jobb eredményre vezet, ha a változók eloszlása meghatározott. A legkisebb négyzetes becslésnél nincs statisztikai próbafüggvény az illeszkedés vizsgálatára, míg a maximum–likelihood módszernél könnyen számolható tesztünk van. A főfaktor eljárás a legjobb r faktort határozza meg, amelyek a legnagyobb mértékben járulnak hozzá a megfigyelt változók varianciájához, míg a kanonikus faktorelemzés olyan faktorokat keres, amelyek maximálisan korrelálnak a megfigyelt változókkal, így a varianciák helyett a megfigyelt változók közötti korrelációk reprodukciójára koncentrál.

2.10 Faktorstruktúrák összehasonlítása azonos minták esetén

A faktorelemzések eredményét nemcsak akkor hasonlíthatjuk össze, ha ugyanazon mintán azonos változók struktúráját különböző faktorelemző eljárással vizsgáljuk. Érdekes lehet összehasonlítani két különböző változóhalmazt egy adott minta esetén, ha az a feltételezésünk, hogy a két változóhalmazra azonos faktorok hatnak. Két nem azonos változóhalmaz faktorainak összehasonlítását HARMAN (1960) nyomán a faktorok invarianciájának mérésével végezhetjük el.

Eszerint ha a két vizsgálat faktorértékeinek mátrixai ($\underline{S}_1$ és $\underline{S}_2$) között korrelációt számítunk:

$$\begin{aligned}\underline{R} &= \frac{1}{n}\begin{pmatrix}\underline{S}'_1\\ \\ \underline{S}'_2\end{pmatrix}[\underline{S}_1,\underline{S}_2]\\ &= \frac{1}{n}\begin{pmatrix}\underline{S}'_1\underline{S}_1 & \underline{S}'_1\underline{S}_2\\ \underline{S}'_2\underline{S}_1 & \underline{S}'_2\underline{S}_2\end{pmatrix}\\ &= \begin{pmatrix}\underline{R}_{11} & \underline{R}_{12}\\ \underline{R}_{21} & \underline{R}_{22}\end{pmatrix},\end{aligned} \tag{73}$$

akkor $\underline{R}_{12}(=\underline{R}'_{21})$ a két faktormegoldás kongruencia együtthatóit tartalmazza, ezek közvetlenül mérik a faktorok hasonlóságát. Könnyen belátható, hogy ha a két megoldás faktorait transzformáljuk úgy, hogy azok maximálisan egybeessenek (kongruáljanak), és a faktorok megoldásonként ortogonálisak, akkor a kongruencia együttható egyenlő a kanonikus korrelációval.

2.11 Faktorstuktúrák összehasonlítása különböző minták esetén

Két különböző vizsgálat (minta) egy–egy faktora közötti hasonlóságot mérhetjük a faktorértékek (factor scores) között számított korrelációs együtthatóval, amit ebben az esetben invariancia együtthatónak nevezünk. HENRYSSON (1957) nyomán nevezhetjük ezt konfigurációs invarianciának is, tekintettel arra, hogy a faktorértékek között számított korrelációs együttható nem érzékeny a faktorsúlyok nagyságrendbeli különbségeire.

A faktorstruktúra hasonlóságát mérhetjük a faktorsúlyok között számított korrelációs együtthatóval is. Problematikus azonban ennél a mutatónál, hogy ha pl. az egyik faktor súlyai .0 és .85 között mozognak, a másiké pedig –.85 és .85 között, akkor a korreláció számításánál a két standardizált intervallumban a 0 faktorsúly azonos értéket kap az erős negatív –.85 faktorsúllyal. Ezért javasolta PINNEAU és NEWHOUSE (1964), hogy a korrelációt a faktorsúlyok négyzetei között számoljuk. Ez a javaslat azonban ugyanazt a problémát veti fel, ti. két ellentétes értelmű súly (–.85 és .85) azonos értéket vesz fel a számítás során.

Egy másik kézenfekvő mutató a faktorsúlyok eltéréseiből számított négyzetes középérték. Jelöljük az első megoldás faktorsúlyait ${}_1\lambda_{jk}$-val, a második megoldás faktorsúlyait pedig ${}_2\lambda_{jt}$-vel. Ekkor az eltérés négyzetes középértéke:

$$h_{kt} = \sqrt{\frac{1}{m}\sum_{j}^{m}({}_1\lambda_{jk} - {}_2\lambda_{jt})^2}\,. \tag{74}$$

Ez a mutató közelít 0–hoz, ha a faktorsúlyok hasonlók, de a mutató egyszerűségéből adódóan nehéz megmondani, hogy milyen értéke fejezi ki a „jó megegyezést".

TUCKER (1951) javasolta a faktorok hasonlóságának mérésére a kongruencia együtthatót, melynek értelmezési tartománya –1 és +1 közé esik.

A kongruencia együttható

$$\psi_{kt} = \frac{\sum_j^m {}_1\lambda_{jk}{}_2\lambda_{jt}}{\sqrt{(\sum_j {}_1\lambda_{jk}^2)(\sum_j {}_1\lambda_{jt}^2)}}\,. \tag{75}$$

Látható, hogy a fenti együttható a szorzat momentum együtthatóhoz hasonló, de nem egyenlő a korrelációs együtthatóval.

Általánosságban, ha ${}_1\underline{\Lambda}$ az első megoldás $(m \times p_1)$ típusú, ${}_2\underline{\Lambda}$ a második megoldás $(m \times p_2)$ típusú faktorsúly– mátrixa, és a faktorsúly–mátrixok oszlopvektorait normáljuk a következő módon:

$${}_1\underline{\Lambda}^* = {}_1\underline{\Lambda}\begin{pmatrix} \frac{1}{\sqrt{\Sigma_i^m {}_1\lambda_{i1}^2}} & & \\ & \vdots & \\ & & \frac{1}{\sqrt{\Sigma_i^m {}_1\lambda_{ip_1}^2}} \end{pmatrix}, \tag{76}$$

és hasonlóan számítjuk ${}_2\underline{\Lambda}^*$–t, akkor a kongruencia együtthatók $(p_1 \times p_2)$ típusú mátrixa:

$$\underline{\Psi} = {}_1\underline{\Lambda}^{*\prime}{}_2\underline{\Lambda}^*\,. \tag{77}$$

A kongruencia együtthatók megbízhatóságának vizsgálatára szignifikancia próba nem áll rendelkezésre, és az értelmezhetősége is bizonytalan lehet. Nézzük például négy változónak két különböző faktorsúlyait: .01, .03, .02, .04, és .91, .93, .92, .94. A két faktorstruktúra azonos, bár közöttük .90 konstans különbség van, a kongruencia együttható pedig csak .92.

Nézzünk egy másik példát: a két faktor súlyai, .2, .3, .1, .3 és .3, .2, .4, .1. Ebben az esetben a kongruencia együttható .70, és ez mutatja a gyengébb egyezést (strukturálisan).

Általában a kongruencia együttható képletéből adódóan magasabb kongruencia együtthatót kapunk azoknál a faktoroknál, amelyeknél a megfelelő súlyok azonos előjelűek.

Ha a két vizsgálat (vagy a két különböző faktorelemző eljárás) minden faktorpárjára kiszámítjuk a kongruencia együtthatókat, akkor a kongruencia együtthatók mátrixához jutunk (lásd a (77) egyenletet). Tekintsük a továbbiakban azt az esetet, amikor a két vizsgálat során azonos számú faktort határozunk meg, így a kongruencia mátrix kvadratikus lesz – $(p \times p)$ típusú.

Természetesen nem várhatjuk feltétlenül el, hogy az azonos indexű faktorok hasonlítsanak legjobban egymáshoz. Ezért ha a két faktormegoldás hasonlóságát egy összegző mutatóval akarjuk kifejezni, figyelembe kell venni, hogy az első megoldás egyes faktorai a második megoldás mely faktoraihoz hasonlítanak legjobban. Célszerűnek látszik a második megoldás faktorsúly–mátrixát permutálni úgy, hogy az az első faktormegoldás faktoraihoz igazodjon a maximális hasonlóság értelmében. Jelöljük a permutációs mátrixot $\underline{P}$-vel. Ha a két faktormátrix azonos indexű faktorai hasonlítanak legjobban egymáshoz, a $\underline{P}$ mátrix diagonális lesz ($\underline{P} = \underline{I}$). Amennyiben az első faktormátrix i-edik faktora a második faktormátrix j-edik faktorához hasonlít legjobban, akkor a permutáló mátrix i-edik és j-edik diagonális eleme 0 lesz, és az (i, j), (j, i) cellákba kerül 1. Akkor is permutálni kell a második faktormátrix oszlopvektorait, ha az összehasonlítandó faktorsúlyok előjelben különböznek egymástól. Így ha a kongruencia mátrix i-edik sorában a j-edik elem lesz maximális abszolút értékű, de az együttható negatív előjelű, akkor a második faktormátrix j-edik faktorát -1-gyel kell szorozni és az i-edik faktorral kicserélni. Ekkor a permutációs mátrix (i, j)-edik eleme 1, a (j, i)-edik eleme pedig -1 lesz. Ha az így összeállított permutációs– mátrixszal jobbról megszorozzuk a második faktormátrixot, akkor a két faktormátrix közvetlenül összehasonlíthatóvá válik.

Jelölje a permutált faktormátrixot ${}_2\underline{\Lambda}^* = {}_2\underline{\Lambda}\, \underline{P}$.

Az összehasonlítandó két faktormátrix különbség mátrixa:

$$\underline{E} = {}_1\underline{\Lambda} - {}_2\underline{\Lambda}^* , \tag{78}$$

és a különbözőség összegző mutatója:

$$\gamma = \text{trace}\,(\underline{E}'\underline{E}) . \tag{79}$$

A γ mutató tehát a két faktormegoldás különbözőségének mértékét fejezi ki. Tökéletes egyezés esetén értéke 0 lesz, hasonló megoldásoknál közel 0.

A γ mutatót függetlenné tehetjük a változók és a faktorok számától, ha elosztjuk a változók és a faktorok számának szorzatával, így különböző vizsgálatok eltérő változó– és faktorszámára kapott különbözőségeket közvetlenül össze tudunk hasonlítani:

$$\bar{\gamma} = \gamma/\text{mp} = \frac{1}{\text{mp}}\text{trace}(\underline{E}'\underline{E}) . \tag{80}$$

Mérhetjük a különböző faktormegoldások hasonlóságát a kongruencia mátrix elemeitől függő mutatóval is. A kongruencia mátrix diagonális elemeinek átlaga lehet egy ilyen mérőszám. Természetesen a kongruencia mátrix sor, ill. oszlopmaximumai nem feltétlenül a diagonális elemei, így szükséges lehet a fentiekben már definiált permutációs mátrix felhasználásával átrendezni a kongruencia mátrix oszlopvektorait.

Jelölje $\underline{\Psi}^* = \underline{\Psi}\ \underline{P}$ a permutált kongruencia mátrixot.

Két faktormátrix hasonlóságát méri a $\underline{\Psi}^*$ kongruencia mátrix diagonális elemeinek átlaga:

$$\Psi = \frac{1}{p}\mathrm{trace}(\underline{\Psi}^*)\,. \tag{81}$$

A mutató értékkészlete 0 és 1 között van, maximális egyezés esetén az 1 értéket veszi fel.

Egy másik mutató lehet a kongruencia mátrix deteminánsa:

$$\delta = |\underline{\Psi}^*|\,. \tag{82}$$

Amennyiben a két faktormátrix faktorai tökéletesen illeszkednek egymáshoz, akkor a kongruencia mátrix determinánsa egyenlő lesz 1-gyel.

A faktorstruktúrák hasonlóságának mérésére javasolt mutatók alkalmazására a következőkben mutatunk példát.

2.12 Különböző faktorelemző eljárások empirikus összehasonlítása

A következő faktorelemzési példa az MTA Szociológiai Kutató Intézetének Értékszociológiai Műhelye által 1982-ben végzett életút–értékrendszer vizsgálat anyagából készült.

A minta országos, a felnőtt lakosságot reprezentáló minta, a megfigyelési elemszám: $n = 1464$.

A kérdés, amire a választ kértük, a következő volt:

„Néhány tulajdonságot sorolunk fel, amire nevelni lehet a gyerekeket. Melyeket tartja Ön különösen fontosnak? Kérem, válassza ki az öt legfontosabbat!"

A következő tulajdonságok szerepeltek:

(01) — jó magaviselet
(02) — udvariasság
(03) — önállóság
(04) — a kemény munka szeretete
(05) — őszinteség
(06) — felelősségérzet
(07) — türelem

(08) — képzelőerő, fantázia
(09) — mások tisztelete, tolerancia
(10) — vezetőkészség
(11) — önfegyelem
(12) — takarékosság
(13) — határozottság, állhatatosság
(14) — vallásos hit
(15) — önzetlenség
(16) — engedelmesség
(17) — hűség, lojalitás

A válaszokat dichotom változókká alakítottuk, amelyek két kategóriája: választotta és nem választotta (kódjai 5 és 0). A válaszok mögött meghúzódó főbb látens értékválasztási dimenziók felderítésére a faktorelemzés eljárását alkalmaztuk. A számításokat IBM PC/AT számítógépen, az SPSS PC$^+$ programrendszerrel végeztük. A következőkben bemutatjuk a hét különböző faktorelemző eljárás eredményeit, majd a hét eljárás különbözőségét, ill. hasonlóságát a γ, ψ és a δ mutatók alapján a MINISSA módszer segítségével elemezzük. (A MINISSA módszer és az ezen a néven ismert program az MDS(X) programcsomag egyik eljárása. Ezeket a számításokat az IBM 3031 számítógépen végeztük.)

A hét faktorelemző eljárás:

1. PC Principal component analysis (Főkomponenselemzés)
2. PAF Principal axis factoring (Főfaktorelemzés)
3. ALPHA Alpha factoring (Alfa faktorelemzés)
4. IMAGE Image factoring (Image–elemzés)
5. ULS Unweighted least squares (Legkisebb négyzetek módszere)
6. GLS Generalized least squares (Általánosított legkisebb négyzetek módszere)
7. ML Maximum likelihood (Maximum likelihood eljárás)

A hét faktoreljárás összehasonlításának eredményeit az 1–6 ábrák szemléletesen mutatják. Ezek a MINISSA eljárás háromdimenziós megoldásainak az első két dimenzióit tartalmazzák. Tulajdonképpen a kétdimenziós megoldás illeszkedése is kitűnő volt, de néha túlságosan is közel, fedésbe kerültek az eljárásokat reprezentáló pontok azért, mert a harmadik dimenziónak a szórása is beleolvadt a kétdimenziós megoldásba. (Később még utalni fogunk arra, hogy mely módszert különített el leginkább a harmadik dimenzió.)

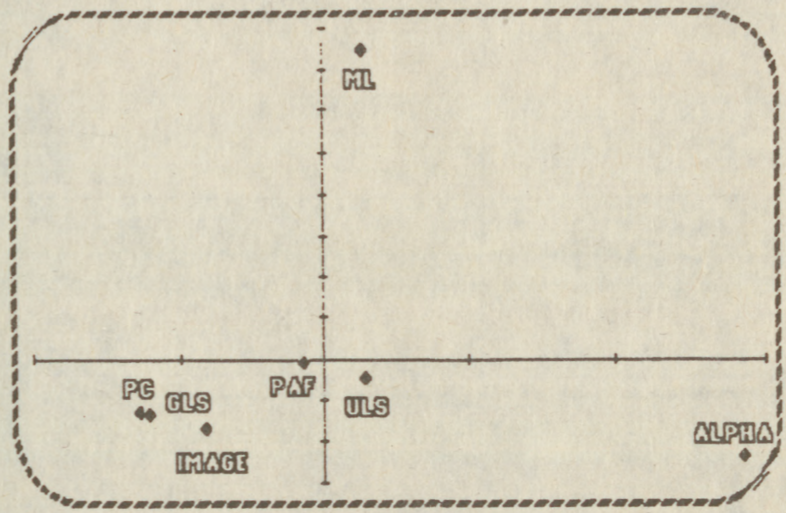
A háromdimenziós megoldás illeszkedését mérő mutató, a nyomaték (stress) értéke a különbözőségi, kongruencia átlag és a kongruencia mátrix determinánsa együtthatóknál rendre a következő volt: a rotálatlan faktoroknál .0000021518, 0.0, .0000019681, rotált faktoroknál .00053942, .0000018542, 0.0. Megállapíthatjuk, hogy a megoldások szinte tökéletes illeszkedést adtak.

Összességében a leghasonlóbb, lényegében azonos megoldást adta a főkomponens (PC) és az általánosított legkisebb négyzetek (GLS) módszere (a negyedik tizedesjegyben voltak eltérések, néha a harmadikban), majdnem ugyanilyen mértékű egyezést találtunk a főfaktor (PAF) és a klasszikus legkisebb négyzetek módszere

(ULS) között, de ez a két páros egymáshoz is közel állt. A legnagyobb különbözőséget tapasztaltuk az alfa (ALPHA) és a maximum likelihood eljárás (ML) között, és ez a két módszer a többitől is hasonló mértékben tért el.

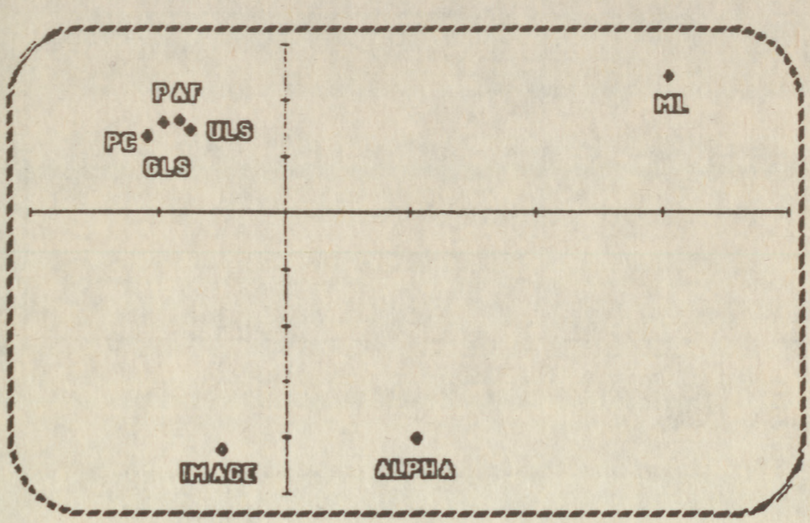
1. ábra

Faktorelemző eljárások összehasonlítása
(különbözőségi együttható, rotálatlan faktorok)



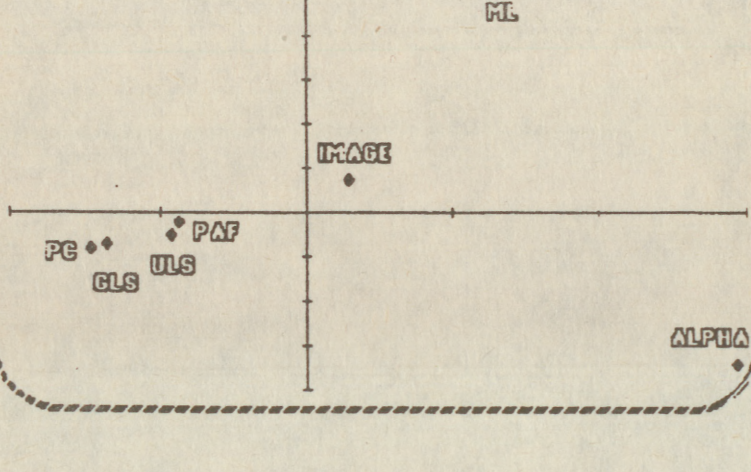
2. ábra

Faktorelemző eljárások összehasonlítása
(kongruencia együttható, rotálatlan faktorok)



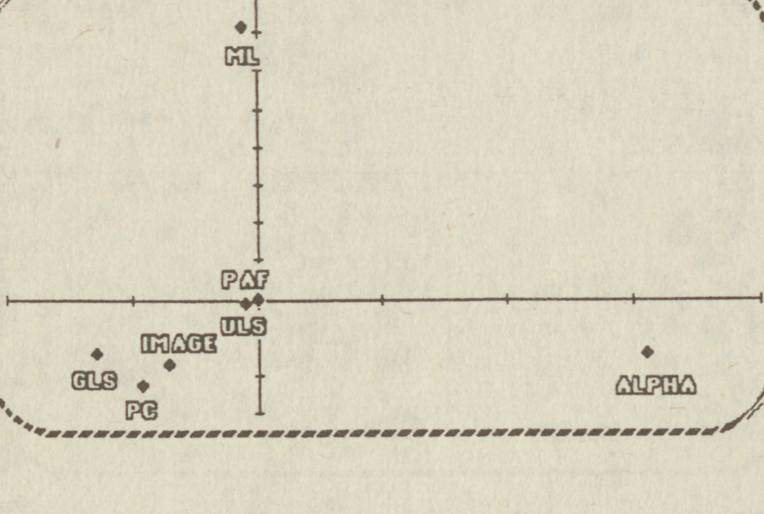
3. ábra

Faktorelemző eljárások összehasonlítás
(kongruencia mátrix determinánsa, rotálatlan faktorok)



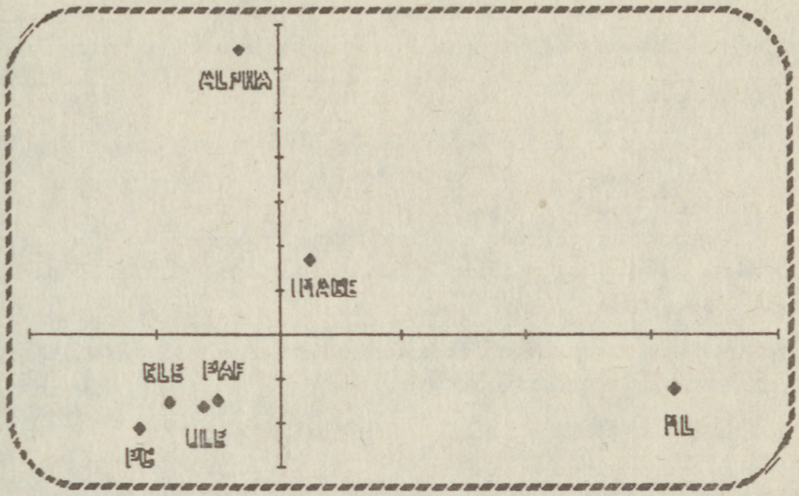
4 ábra

Faktorelemző eljárások összehasonlítása
(különbözőségi együttható, rotált faktorok)



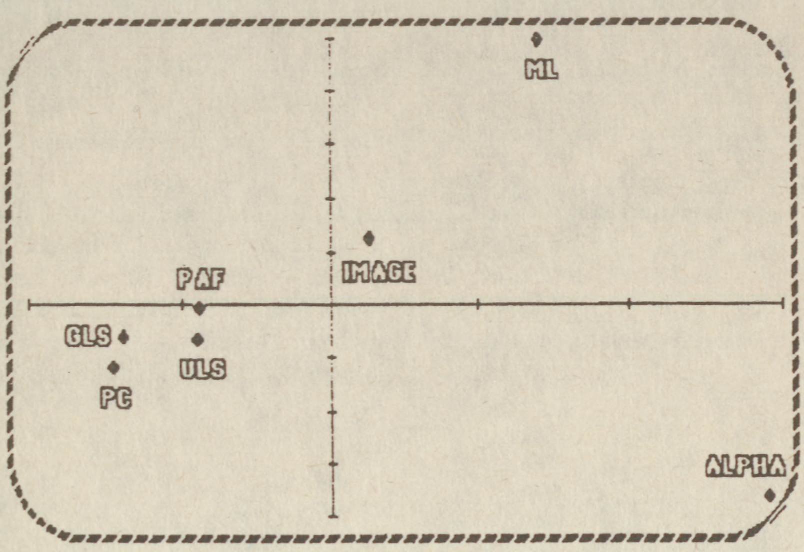
5. ábra

Faktorelemző eljárások összehasonlítása
(kongruencia együttható, rotált faktorok)



6. ábra

Faktorelemző eljárások összehasonlítás
(kongruencia mátrix determinánsa, rotált faktorok)



A fenti eredmény mindhárom vizsgált mutató esetében megegyező volt. Az image (IMAGE) eljárás a PC, GLS, PAF, ULS és az ML, ALPHA módszerek között helyezkedett el, bár az átlagos kongruencia együttható ítélete szerint az első csoporttól eltávolodott, és az ALPHA eljáráshoz került közelebb (ez az eltérés azonban a rotációval gyengébbé vált).

Érdemes megemlíteni, hogy az IMAGE eljárás a harmadik dimenzióban lényegesen elkülönült a többi eljárástól, és ez magyarázza azt, hogy a kétdimenziós vetületben a sík közepe tájékán mozgott — és ezért került relatíve közelebb a PC, GLS, PAF és ULS eljárásokhoz, vagyis az ML, az ALPHA és az IMAGE egymástól és a többi eljárástól is a legnagyobb különbözőséget adta. Meg kell említenünk, hogy a GLS módszernek a várttól eltérő eredménye azzal magyarázható, hogy a vizsgált változók együttes eloszlása nem volt normális, még ha feltételezzük a választás folytonosságát, a durva mérés mögött akkor is csak egy erősen ferde, aszimmetrikus eloszlású változóhalmazt vizsgáltunk. A faktorok tartalmát a főkomponenselemzés (PC) megoldásával mutatjuk be.

Először összefoglalóan megadjuk a legfőbb értékdichotómiákat, majd táblázatokba rendezve a faktorok súlyait.

Az első rotálatlan faktor a legfőbb értékpolaritást fejezi ki:

Modern személyiségértékek ⟷ Hagyományos közösségi értékek.

Ezt az alapvető értékdimenziót figyeltünk meg a Rokeach értékteszt elemzése során 1980 és 1982-ben is (intellektuális és autonómia értékek álltak szemben a hagyományos közösségi és öröm–értékekkel; lásd HANKISS, et. al., 1983). Ebben a dimenzióban a „belülről irányított" ember típusa áll szemben a „tradícióktól irányított" ember tulajdonságaival.

A második rotálatlan faktor:

Munka, biztonság, hit ⟷ Emberi kapcsolatok, autonómia.

A pozitív oldalon kétfajta értékek keverednek: a munkával kapcsolatos és a vallásos értékek, ezen belül az elsőknek van nagyobb, meghatározóbb súlya. Ebben az összefüggésben a vallásos hit a kötelesség, felelősség, becsület fogalmaihoz kapcsolódik. A munka értéke bizonyos fajta biztonságra való törekvéssel fonódik össze: anyagi szempontból a takarékossággal, lelki szempontból a vallásos hittel és a hűséggel. A másik, negatív oldalon az emberi kapcsolatokat szabályozó tulajdonságok szerepelnek. Olyan értékek, amelyek a személyes autonómia fenntartása mellett a másik ember számára is könnyebbé, zökkenőmentesebbé teszik az együttélést.

A harmadik rotálatlan faktor dichotómiája:

Hagyományos keresztény értékek ⟷ rotestáns értékek.

A pozitív póluson a hagyományos keresztény értékek szerepelnek, középpontban mások tisztelete, az alkalmazkodás értékei, mindössze a határozottság keveredik közéjük. Ennek a dimenziónak az ellentétes oldala az evilági élet, ahol az anyagi, protestáns értékek a fontosak.

A faktorok rotálása az első faktort lényegében nem változtatta meg, talán egy kicsit még tisztábbá tette a belülről kontra kívülről irányított ember értékdichotómiát. A második faktornál a rotálás során elmaradt a pozitív oldalról a felelősségérzet és a fantázia, de ezeknek egyébként is gyenge volt a szerepük. Erre az oldalra jellemző WEBER (1982) protestáns etikája: „... ennek az etikának summum bonumja a pénzszerzés, egyre több pénz szerzése, kérlelhetetlenül megtagadva minden elfogulatlan élvezetet...". Máshol a protestáns aszkézisről azt írja: „Erkölcsileg ugyanis akkor vagyunk igazán megvetendők, ha megnyugszunk birtokunkban, élvezzük a gazdagságot, haszontalanul és érzékien élünk, s főleg akkor, ha már nem törekszünk a 'szent' életre... Nem a kényelem és az élvezet, hanem csakis a cselekvés szolgálja Isten egyértelműen kinyilvánított akarata szerint dicsőségének gyarapítását."

A harmadik rotált faktorban tisztábban jelenik meg az emberekhez, a társadalomhoz való igazodás kétféle változata, a hagyományos keresztény és a pragmatikusabb, materiálisabb, de az elfogadás fontosságát is valló szemlélet.

A továbbiakban a részletes faktormátrixokat közöljük.

Gyermeknevelési elvek első, rotálatlan faktora (PC elemzés)

(sajátérték = 1.96, 11.7 %)

+		−	
felelősségérzet	.50	jó magaviselet	.60
önállóság	.43	engedelmesség	.54
határozottság	.40	vallásos hit	.39
fantázia	.39	udvariasság	.33
önfegyelem	.30	takarékosság	.27
önzetlenség	.27	munkaszeretet	.15
őszinteség	.17		
vezetőkészség	.12		

Modern személyiség-értékek ⟷ **Hagyományos közösségi értékek**

Gyermeknevelési elvek második, rotálatlan faktora (PC elemzés)

(sajátérték = 1.40, 8.2 %)

+		−	
munkaszeretet	.52	őszinteség	.61
vezetőkészség	.37	udvariasság	.52
takarékosság	.28	önállóság	.30
türelem	.21	jó magaviselet	.11
hűség	.20		
vallásos hit	.18		
felelősségérzet	.16		
fantázia	.10		

Munka, biztonság, hit	**Emberi kapcsolatok autonómia**

Gyermeknevelési elvek harmadik, rotálatlan faktora (PC elemzés)

(sajátérték = 1.27, 7.5 %)

+		−	
tolerancia	.62	takarékosság	.53
önzetlenség	.35	önállóság	.42
határozottság	.31	önfegyelem	.19
vallásos hit	.26	vezetőkészség	.17
hűség	.18	udvariasság	.14
engedelmesség	.15		

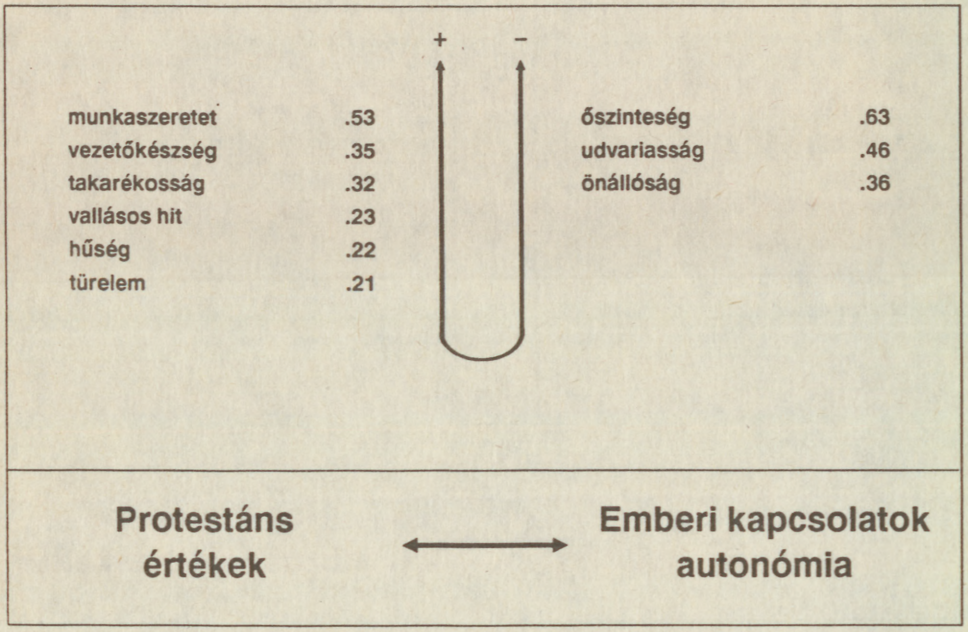
Hagyományos keresztény értékek ⟷ **Protestáns értékek**

Gyermeknevelési elvek első, rotált faktora (PC elemzés)

+		−	
felelősségérzet	.50	engedelmesség	.57
önállóság	.50	jó magaviselet	.56
fantázia	.38	vallásos hit	.42
önfegyelem	.35	udvariasság	.34
határozottság	.28	tolerancia	.21
vezetőkészség	.22		
önérzetesség	.14		

Modern személyiség-értékek ⟷ **Hagyományos közösségi értékek**

Gyermeknevelési elvek második, rotált faktora (PC elemzés)



Gyermeknevelési elvek harmadik, rotált faktora (PC elemzés)

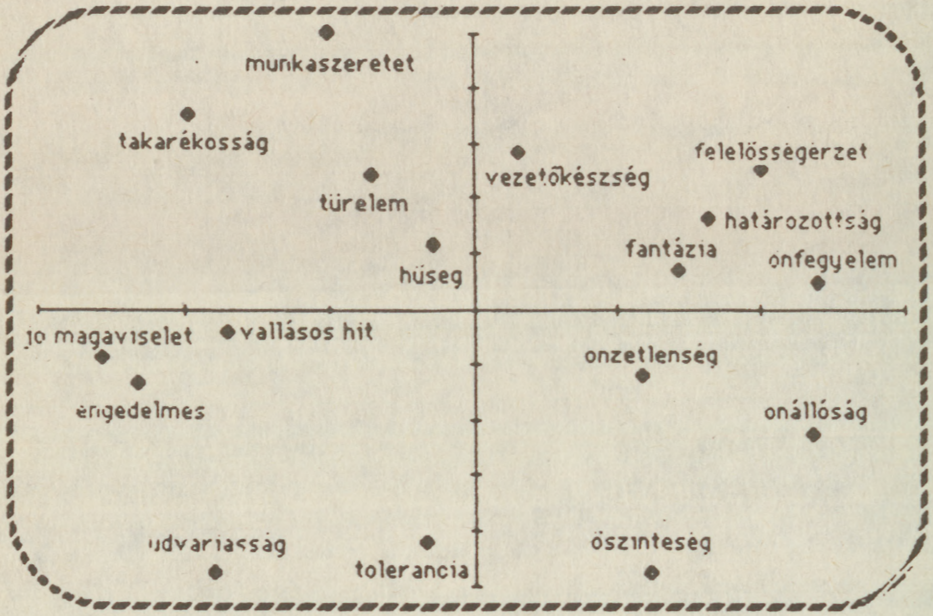
+		−	
tolerancia	.58	takarékosság	.58
határozottság	.42	jó magaviselet	.28
önzetlenség	.41	önállóság	.27
hűség	.18	udvariasság	.25
felelősségérzet	.14	vezetőkészség	.10
vallásos hit	.13	munkaszeretet	.10
fantázia	.12		

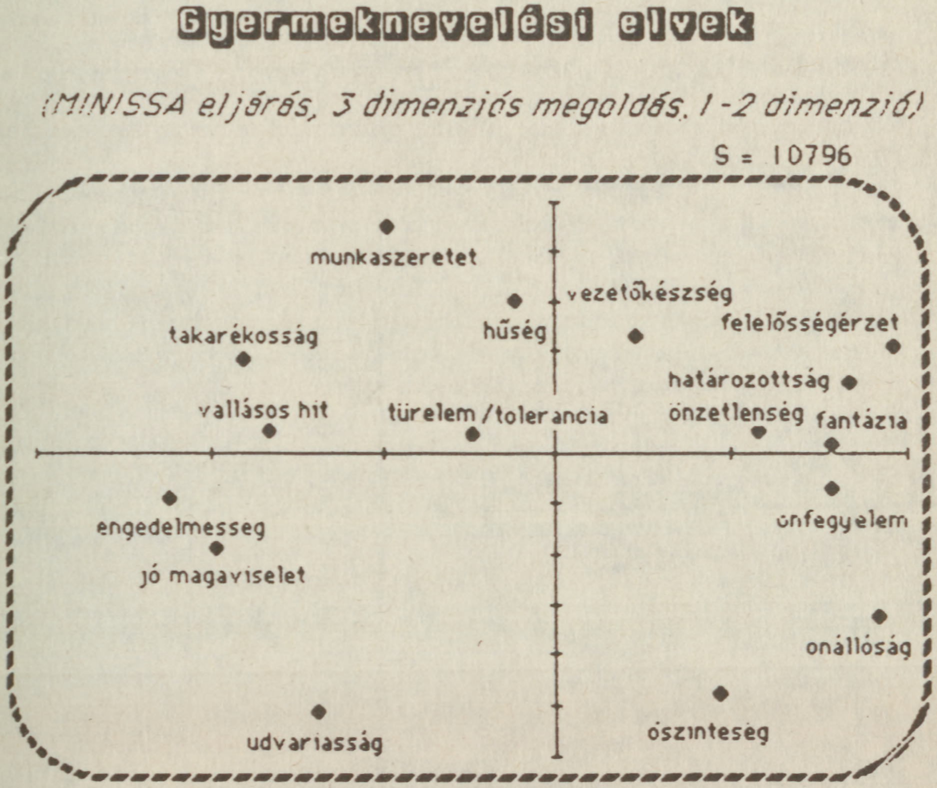
Hagyományos keresztény értékek ⟷ **Protestáns értékek**

Gyermeknevelési elvek

(MINISSA eljárás, 2 dimenziós megoldás)

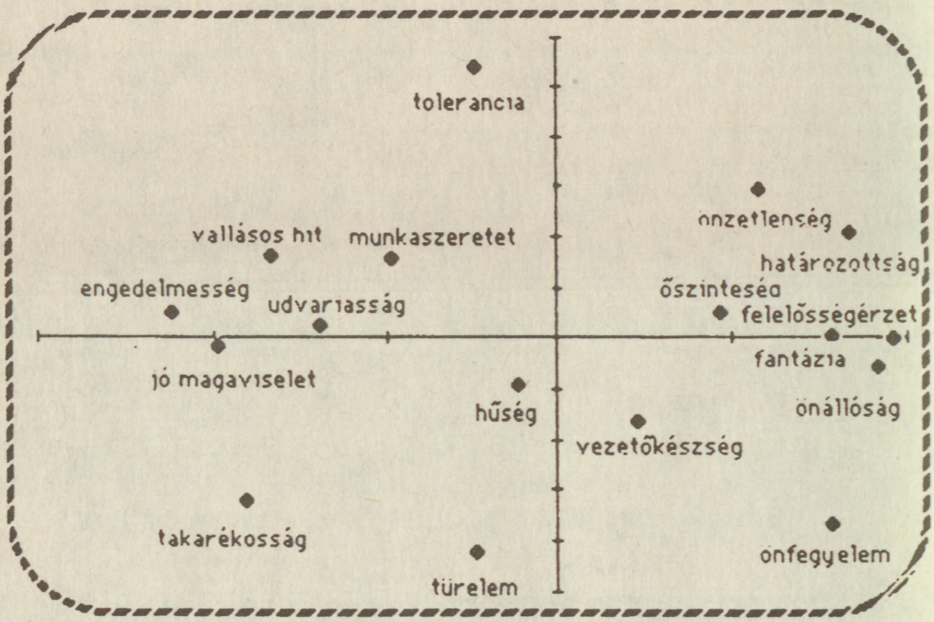
S = 1941




Gyermeknevelési elvek
(MINISSA eljárás, 3 dimenziós megoldás, 1-2 dimenzió)
S = 10796
munkaszeretet
vezetőkészség
hűség
felelősségérzet
takarékosság
határozottság
vallásos hit
türelem / tolerancia
önzetlenség
fantázia
engedelmesség
önfegyelem
jó magaviselet
önállóság
udvariasság
őszinteség

Gyermeknevelési elvek

(MINISSA eljárás, 3 dimenziós megoldás, 1-3 dimenzió)



Monte Carlo vizsgálat

Az előző vizsgálatban egy empirikus elemzés mintabeli korrelációs mátrixa alapján hasonlítottuk össze a kérdéses hét faktorelemző eljárást. A következőkben Monte Carlo elemzéssel vizsgáljuk a különböző faktorelemző eljárásokat.

Kiindultunk két különböző faktorstruktúrából. A faktormátrixok alapján generáltunk az SPSS PC$^+$ programrendszer segítségével 50, 100, 200 és 400 elemű véletlen mintákat, amelyekre teljesült a változók normalitása. A négy különböző mintanagyságra azután elvégeztük a hét különböző faktorelemző eljárást, majd a kapott faktormátrixok hasonlóságát a γ (különbözőségi együttható), ψ (kongruencia együttható) és a δ (a kongruencia mátrix determinánsa) mutatók alapján hasonlítottuk össze. A végeredményt a következő táblázatban közöljük.

Faktorelemző eljárások összehasonlítása
két különböző faktorstruktúra és négy véletlen minta alapján

(a táblázatban azt jelöljük, hogy melyik faktorelemző eljárás adott a "várttól" eltérő eredményt)

50-es minta

Különböző faktor–struktúrák	Különbözőségi mutató 1	Kongruencia mutató 2	Kongruencia mátrix determinánsa 3
1	OK	OK	ML,IMAGE
2	ML	ML	ML

100-as minta

	1	2	3
1	ML	ML	ML
2	ML	OK	ML

200-as minta

	1	2	3
1	OK	OK	ML,IMAGE
2	OK	OK	ML

400-as minta

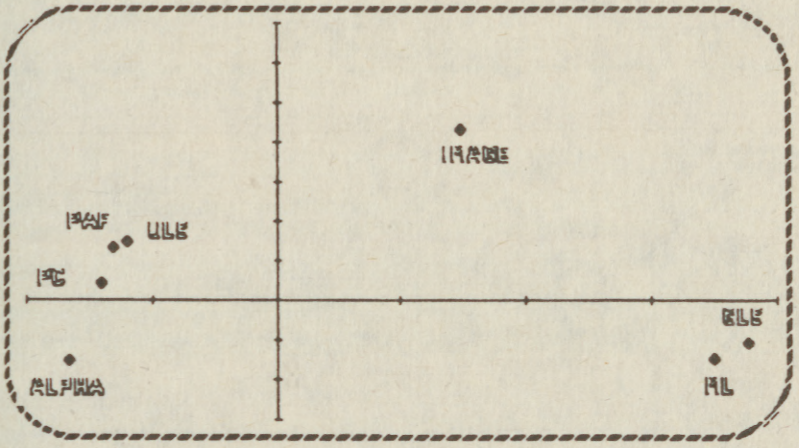
	1	2	3
1	OK	OK	ML,IMAGE
2	OK	OK	ML

A hét faktorelemző eljárás hasonlóságait a 2–dimenziós MINISSA térben ábrázolva jól láthatók a következő összefüggések (lásd a következő négy ábrát, amelyeket a kongruencia együttható alapján készítettünk):

- lényegében azonos eredményeket adott a PC, ULS és PAF,
- az ALPHA is lényegében közeli struktúrát eredményezett, bár kissé elmozdult az ML irányába a második dimenzióban,
- a legnagyobb különbség a PC, ULS, PAF, ALPHA és az IMAGE, ML, GLS módszerek között figyelhető meg, vagyis alapvető eltérést eszerinti csoportosításban találtunk,
- az IMAGE, ML, GLS csoporton belül az IMAGE elemzés tért el legjobban a másik kettőtől is.

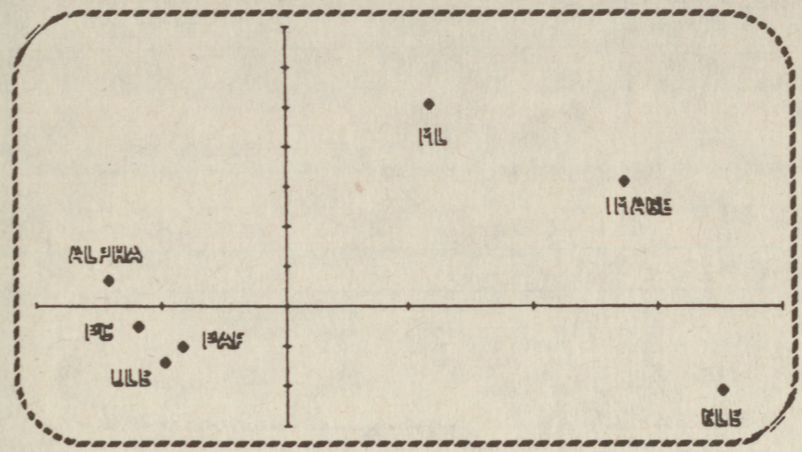
Faktorelemző eljárások összehasonlítása

(Monte Carlo vizsgálat, n = 50, Kongruencia együttható)



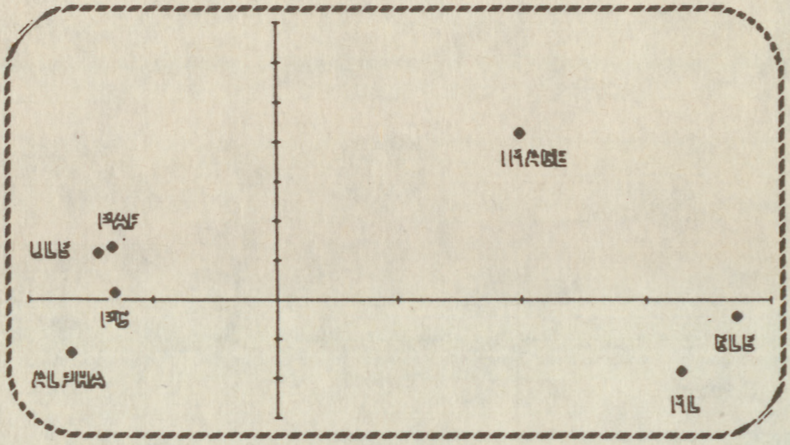
Faktorelemző eljárások összehasonlítása

(Monte Carlo vizsgálat, n = 100, Kongruencia együttható)



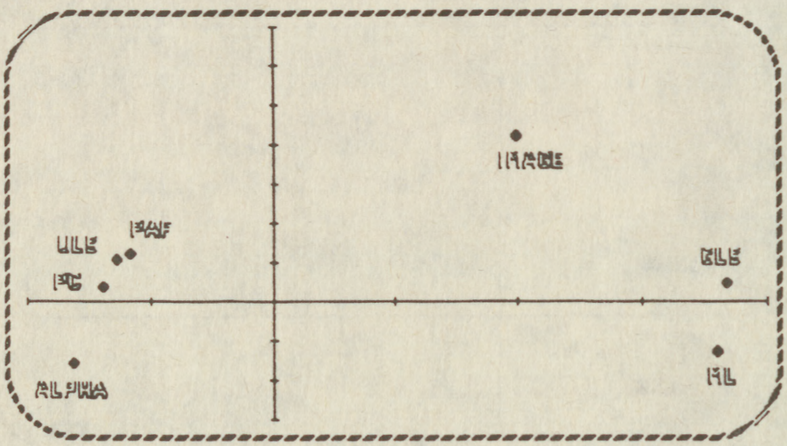
Faktorelemző eljárások összehasonlítása

(Monte Carlo vizsgálat, n = 200, Kongruencia együttható)



Faktorelemző eljárások összehasonlítása

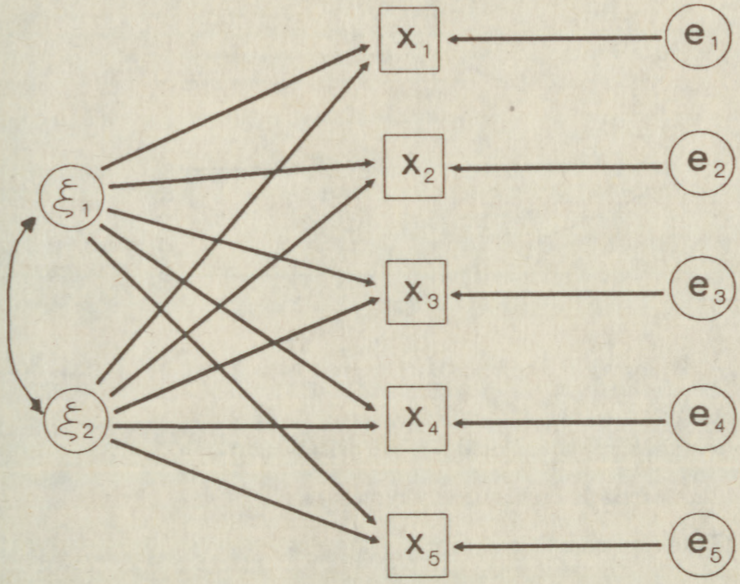
(Monte Carlo vizsgálat, n = 400, Kongruencia együttható)



3. A konfirmatív faktorelemzés

A faktorelemzést eddig mint exploratív elemző eljárást tárgyaltuk. Abból az alapfeltételezésből indultunk ki, hogy a megfigyelt változók kifejezhetők látens változók, faktorok lineáris függvényeként. A manifeszt változók megfigyelt összefüggéseit a faktorok magyarázzák, így ha a faktorok hatásait a változókból kiszűrjük, akkor a változók korrelálatlanok lesznek.

Az exploratív faktorelemzés modelljét mutatja a következő ábra:



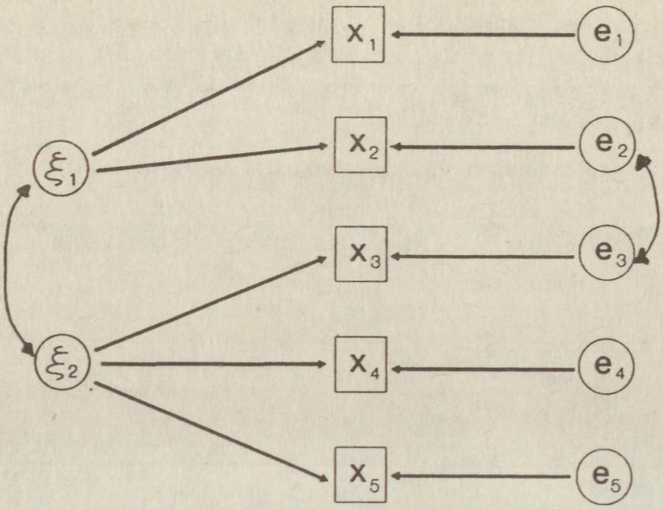
1. ábra. *Az exploratív faktorelemzés modellje*

Ahogyan a fenti ábrából is láthatjuk, az exploratív faktorelemzésnél feltételezzük, hogy

- a közös faktorok (ξ) korrelálnak egymással (vagy korrelálatlanok, más esetben),
- mindegyik közös faktor mindegyik változóra hatással van,
- az egyedi faktorok vagy hiba komponensek (e_i) korrelálatlanok egymással,
- mindegyik megfigyelt változóhoz tartozik egy egyedi faktor (hiba komponens),
- a közös faktorok korrelálatlanok az egyedi faktorokkal.

Láthatjuk, hogy az exploratív faktorelemzés modellje nem tartalmaz feltételezéseket a modell struktúrájára, nincs *a priori* hipotézisünk, hogy a közös faktorok mely manifeszt változókra hatnak és melyekre nem, a közös faktorok között melyek vannak kapcsolatban egymással és melyek nem. Ha van *a priori* ismeretünk a manifeszt és a látens változók közötti összefüggésekről, akkor ezeket beépíthetjük a modellbe, és az így felállított konfirmatív faktorelemzés modelljét illesztjük az adatokhoz.

A következő ábra egy ilyen modellt mutat:



2. ábra. *A konfirmatív faktorelemzés modellje*

A konfirmatív faktorelemzés modelljében feltételezhetjük, hogy

- mely közös faktorpárok korrelálnak egymással,
- melyik közös faktor melyik megfigyelt változóra hat,
- melyik megfigyelt változóra hat egyedi faktor,
- az egyedi faktorok közül melyek korrelálnak egymással.

A fenti típusú feltételekkel felállított konfirmatív modell illeszkedését a megfigyelt adatokhoz statisztikai próba segítségével vizsgálhatjuk.

A konfirmatív faktorelemzést JÖRESKOG (1967, 1969), JÖRESKOG és LAWLEY (1968) fejlesztette ki, és Jöreskog nevéhez fűződik a LISREL (Linear Structural Relationships by the Method of Maximum Likelihood) számítógépes program is. A későbbiekben látni fogjuk, hogy a LISREL alkalmas a közös faktorok közötti strukturális kapcsolatok elemzésére is.

Matematikai jelölésekkel a modell a következőképpen írható fel:

$$\underline{x} = \underline{\Lambda}\,\underline{f} + \underline{e}, \tag{1}$$

ahol

$\underline{x}$ a megfigyelt változók vektora ($m \times 1$ típusú),

$\underline{f}$ a közös faktorok vektora ($r \times 1$ típusú),

$\underline{e}$ az egyedi faktorokat vagy hiba komponenseket tartalmazza ($m \times 1$ típusú),

$\underline{\Lambda}$ a faktorsúlyok mátrixa ($m \times r$ típusú).

Feltételezzük, hogy $E(\underline{x}) = E(\underline{f}) = E(\underline{e}) = \underline{0}$, vagyis mind a megfigyelt, mind a látens változót a várható értéküktől való eltérésükkel jellemezzük. Feltételezzük továbbá, hogy $E(\underline{f}\,\underline{e}') = \underline{0}$, vagyis a közös faktorok és az egyedi faktorok (hibatagok) korrelálatlanok.

A modell feltétele alapján a megfigyelt változók variancia–kovariancia mátrixa ($\underline{\Sigma}$) a következőképpen fejezhető ki:

$$\underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}, \tag{2}$$

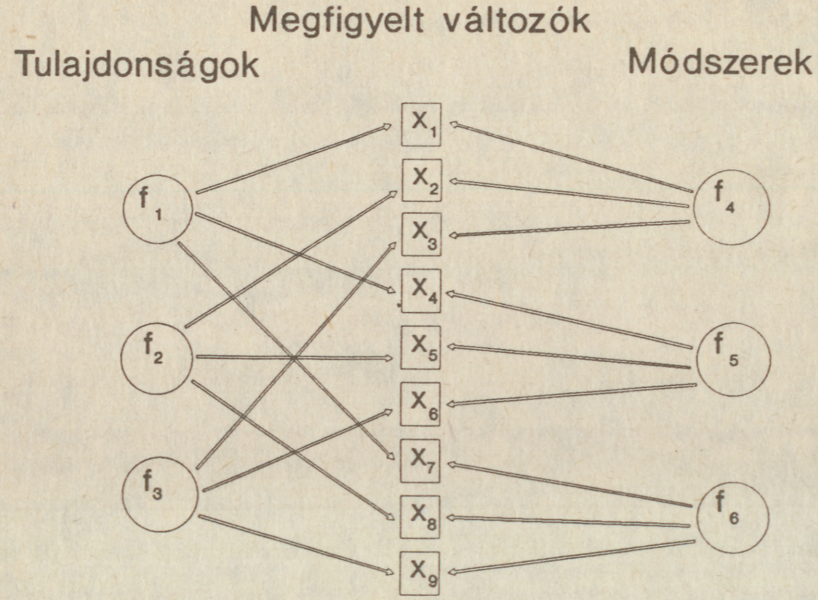
ahol

$\underline{\Phi}$ a közös faktorok variancia–kovariancia mátrixa ($r \times r$ típusú), amelyre $\underline{\Phi} = E(\underline{f}\,\underline{f}')$,

$\underline{\Theta}$ az egyedi faktorok vagy hibatagok variancia–kovariancia mátrixa, amelyre $\underline{\Theta} = E(\underline{e}\,\underline{e}')$.

A konfirmatív faktorelemző modellben a $\underline{\Lambda}$, $\underline{\Phi}$ és $\underline{\Theta}$ paramétermátrixok elemeire teszünk feltételezéseket.

Például tekintsük a több módszer–több tulajdonság (Multi–Method Multi–Trait, MMMT modellt), amelyben feltételezzük, hogy a látens jellemzőket, tulajdonságokat (trait) több módszerrel mérjük. Ha feltételezünk három látens jellemzőt (ezeknek megfelel az f_1, f_2, f_3 faktor), és mindegyiket három módszerrel mérjük (a módszereknek megfelelő faktorok f_4, f_5, f_6), valamint kilenc megfigyelt változónk közül x_1, x_2, és x_3 a megfelelő látens tulajdonság (trait) manifeszt indikátora az első módszerrel mérve, és az x_4, x_5, x_6 és x_7, x_8, x_9 változók a három látens trait indikátorai a másik két módszerrel mérve, akkor az MMMT modell ábrája és együtthatómátrixai a következőképpen írhatók fel:



3. ábra. *Az MMMT (Multi–Method Multi–Trait) modell*

A faktorsúlyok mátrixa:

$$
\begin{array}{c|c}
\begin{array}{ccc} & \text{trait súlyok} & \\ f_1 & f_2 & f_3 \end{array} & \begin{array}{ccc} & \text{módszer súlyok} & \\ f_4 & f_5 & f_6 \end{array}
\end{array}
$$

$$
\underline{\Lambda} = \begin{bmatrix}
\lambda_{11} & 0 & 0 & \lambda_{14} & 0 & 0 \\
0 & \lambda_{22} & 0 & \lambda_{24} & 0 & 0 \\
0 & 0 & \lambda_{33} & \lambda_{34} & 0 & 0 \\
\lambda_{41} & 0 & 0 & 0 & \lambda_{45} & 0 \\
0 & \lambda_{52} & 0 & 0 & \lambda_{55} & 0 \\
0 & 0 & \lambda_{63} & 0 & \lambda_{65} & 0 \\
\lambda_{71} & 0 & 0 & 0 & 0 & \lambda_{76} \\
0 & \lambda_{82} & 0 & 0 & 0 & \lambda_{86} \\
0 & 0 & \lambda_{93} & 0 & 0 & \lambda_{96}
\end{bmatrix}
\begin{array}{|c}
x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9
\end{array}
$$

A faktorok variancia–kovariancia mátrixa

$$\underset{(6\times 6)}{\Phi=}\left(\begin{array}{cc} \text{trait/trait} & \text{trait/módszer} \\ (3\times 3) & (3\times 3) \\ & \\ \text{módszer/trait} & \text{módszer/módszer} \\ (3\times 3) & (3\times 3) \end{array}\right).$$

A faktorok közötti kovarianciák közül az első részmátrixban találhatók a trait–trait–kovarianciák, amelyek mentesek a módszerhatástól. A változók közötti kovarianciákat könnyen felbonthatjuk a trait és a módszer hatásaival összefüggő tagokra.

Pl.

$$E(x_1,x_2)=\sigma_{12}=\lambda_{11}\lambda_{22}\phi_{12}+\lambda_{14}\lambda_{22}\phi_{24}+ \\ +\lambda_{11}\lambda_{24}\phi_{14}+\lambda_{14}\lambda_{24}\phi_{44}$$

Az MMMT modellben a mérési módszert elkülönítjük a látens jellemző, szubsztantív vonástól (trait–től), és mindkét szisztematikus látens komponens hatását (faktorsúlyait) a konfirmatív faktorelemzéssel vizsgálhatjuk.

3.1 *A faktormodell identifikálhatósága*

A modell paramétereinek a becslését mindig meg kell hogy előzze a modell identifikálhatóságának vizsgálata.

A modell identifikálhatósága a modell paramétereinek egyértelmű meghatározhatóságát jelenti. Miután a $\underline{\Sigma}$ variancia–kovariancia mátrix a $\underline{\Lambda}, \underline{\Phi}, \underline{\Theta}$ paraméterek függvénye, a paraméterek akkor határozhatók meg egyértelműen, ha ez a függvény egyértékű függvény, vagyis $\underline{\Sigma}$-t egy és csakis egy paraméterhalmaz (paraméterstruktúra) generálja. Ha egy modell nem identifikálható, akkor a modell $\underline{\Lambda}, \underline{\Phi}$ és $\underline{\Theta}$ paraméterei a $\underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\Lambda' + \underline{\Theta}$ kovariancia egyenletből ismert $\underline{\Sigma}$ esetén sem határozhatók meg egyértelműen, vagyis több paraméterstruktúrához is tartozhat ugyanaz a $\underline{\Sigma}$ mátrix.

Az identifikálhatóság alapvetően a megfigyelt adatokból származó információ (a faktormodellben a variancia–kovariancia mátrix független elemei) és a modell független paraméterei számának különbségével függ össze. A modell alulidentifikált (nem identifikálható), ha a paraméterei nem határozhatók meg egyértelműen. Az éppen identifikálható modellben a paraméterek száma pontosan megegyezik a megfigyelt varianciák és kovarianciák számával. A modell paraméterei ilyenkor becsülhetők (a paraméterek egyértelműen meghatározhatók), azonban egy ilyen modellnek nincs diszkriminatív ereje, tudományosan nem érdekes, mivel soha nem lehet elvetni, minden mintához illeszthető. Az a modell érdekes, amelyik túlidentifikált, paramétereinek (a független paramétereknek) száma kevesebb, mint a megfigyelt varianciák és kovarianciák száma, melynek paraméterei becsülhetők, és amely statisztikai próba segítségével elvethető vagy elfogadható.

Könnyen belátható, hogy az exploratív faktorelemzés modellje nem identifikálható. Kiindulva az $\underline{x} = \underline{\Lambda}\,\underline{f} + \underline{e}$ és $\underline{\Sigma}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}$ egyenletekből és a $\underline{\Lambda}, \underline{\Phi}$ és $\underline{\Theta}$ paramétermátrixokból, bármely ortogonális mátrixszal ($\underline{M}$, $(r \times r)$ típusú) a faktorokat és faktorsúlymátrixot transzformálhatjuk anélkül, hogy a modell két egyenletét megsértenénk.

Jelölje a transzformált faktorokat $\underline{f}^* = \underline{M}\,\underline{f}$, a transzformált faktorsúlyokat $\underline{\Lambda}^* = \underline{\Lambda}\,\underline{M}'$, a transzformált faktorok variancia–kovariancia mátrixát $\underline{\Phi}^* = \underline{M}\,\underline{\Phi}\,\underline{M}'$, ekkor könnyen belátható, hogy mind a $\underline{\Lambda}, \underline{\Phi}, \underline{\Theta}$, mind a $\underline{\Lambda}^*, \underline{\Phi}^*, \underline{\Theta}$ mátrixok kielégítik a faktormodell alapegyenleteit:

$$
\begin{aligned}
\underline{x} &= \underline{\Lambda}^*\,\underline{f}^* + \underline{e} = (\underline{\Lambda}\,\underline{M}')\,\underline{M}\,\underline{f} + \underline{e} \\
&= \underline{\Lambda}\,(\underline{M}'\,\underline{M})\underline{f} + \underline{e} \\
&= \underline{\Lambda}\,\underline{f} + \underline{e},
\end{aligned}
\tag{3}
$$

és

$$
\begin{aligned}
\underline{\Sigma} &= \underline{\Lambda}^*\,\underline{\Phi}^*\,\underline{\Lambda}^{*\prime} + \underline{\Theta} \\
&= (\underline{\Lambda}\,\underline{M}')(\underline{M}\,\underline{\Phi}\,\underline{M}')(\underline{M}\,\underline{\Lambda}') + \underline{\Theta} \\
&= \underline{\Lambda}(\underline{M}'\,\underline{M})\,\underline{\Phi}(\underline{M}'\,\underline{M})\underline{\Lambda}' + \underline{\Theta} \\
&= \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}\,.
\end{aligned}
\tag{4}
$$

A faktorok ilyen transzformációs lehetőségét a faktorok rotálásánál tárgyaljuk még. A faktorok rotálása az exploratív faktorelemzésnél a faktorok egyértelműbb értelmezését segíti. A fentiekben láthattuk, hogy pótlólagos feltételezések nélkül a faktorelemzés modellje nem oldható meg egyértelműen. Feltételezve a faktorok korrelálatlanságát és a véletlen komponensek korrelálatlanságát, megmutatható (lásd LAWLEY és MAXWELL, 1971, 4. fejezet), hogy az exploratív faktorelemzésben az egyértelmű megoldás feltétele, hogy $\underline{\Lambda}'\,\underline{\Theta}^{-1}\,\underline{\Lambda}$ diagonális legyen. $\underline{\Lambda}'\,\underline{\Theta}^{-1}\,\underline{\Lambda}$ r diagonális elemét a következőképpen értelmezhetjük. Ha a hibakomponensek varianciája egységnyi, akkor λ_{ik}^2/Θ_i az i-edik változó varianciájának a k-adik faktorral összefüggő részét fejezi ki. Ha figyelembe vesszük az összes változót, akkor a megfigyelt változók összes varianciájához a k-adik faktor $\Sigma_i(\lambda_{ik}^2/\Theta_i)$ mértékben járul hozzá. Ez a $\underline{\Lambda}'\,\underline{\Theta}^{-1}\,\underline{\Lambda}$ mátrix k-adik diagonális eleme. A faktorokat úgy választjuk, hogy az első faktor maximális mértékben járuljon hozzá a megfigyelt változók együttes varianciájához, a második faktor a maradék lehetőség közül magyarázza maximálisan a megfigyelt változók varianciáit, miközben korrelálatlan az első faktorral (ha $\underline{\Phi}$ egységmátrix), és így tovább.

Az identifikálhatóság szükséges feltételét a független kovariancia egyenletek és a független paraméterek számának az összehasonlításával határozhatjuk meg.

A kovariancia egyenletek száma ((4) egyenlet) $\frac{1}{2}m(m+1)$.

A független paraméterek száma az r–faktoros modellben $mr + m - \frac{1}{2}r(r-1)$, ha $\underline{\Phi}$ egységmátrix.

A különbség a kovariancia mátrix elemei és a független paraméterek száma között:

$$\begin{aligned} s &= \frac{1}{2}m(m+1) - \{mr + m - \frac{1}{2}r(r-1)\} \\ &= \frac{1}{2}[(m-r)^2 - (m+r)]\,. \end{aligned} \tag{5}$$

Az exploratív faktorelemzés identifikálhatóságának szükséges feltétele, hogy $s \geq 0$.

Az (5) egyenlet az identifikálható modellben a faktorok számára ad felső korlátot. Eszerint a faktorok száma nem lehet nagyobb, mint

$$r \leq \{2m + 1 - (8m+1)^{\frac{1}{2}}\}\,. \tag{6}$$

Az (5) feltétel az identifikálhatóságnak csak szükséges feltétele, és nem biztosítja, hogy a véletlen komponensek varianciái ($\underline{\Theta}$ elemei) becslései ne legyenek negatívak. KANO (1986) közöl szükséges és elégséges feltételt a modell identifikálhatóságára a maximum likelihood és az általánosított legkisebb négyzetek becslési módszerek esetén.

A konfirmatív faktorelemzésnél az identifikálhatóság szükséges feltételénél nem tételezhetjük fel általánosan, hogy a $\underline{\Phi}$ mátrix egységmátrix, így a $\underline{\Lambda}, \underline{\Phi}$ és $\underline{\Theta}$ paramétermátrixok összesen $mr + \frac{1}{2}r(r+1) + \frac{1}{2}m(m+1)$ elemet tartalmaznak.

A konfirmatív faktorelemzés identifikálhatóságának szükséges feltétele:

$$\frac{1}{2}m(m+1) - \{mr + \frac{1}{2}r(r+1) + \frac{1}{2}m(m+1)\} \geq 0\,.$$

A paraméterek lehetséges számából természetesen le kell vonni az adott modell kötött paramétereinek számát, és a szabad független paraméterek számát kell kivonni a variancia–kovariancia mátrix független elemeinek számából.

JÖRESKOG és SÖRBOM (1981) az identifikáció vizsgálatát beépítették számítógépes programjukba (LISREL). Az ún. információs mátrix alapján — amely a maximum likelihood becslő függvény paraméterek szerinti második parciális deriváltjait tartalmazza, és függ a paraméterek becsléseinek variancia–kovarianciáitól — azt lehet mondani, hogy ha az információs mátrix pozitív definit, akkor a modell (majdnem biztosan) identifikálható, és ha az információs mátrix szinguláris, akkor a modell nem identifikálható (JÖRESKOG és SÖRBOM, 1981, 11. old.)

3.2 Skála-invariancia

A gyakorlatban a faktormodellt általában a megfigyelt változók mintabeli korrelációs mátrixához illesztjük a variancia-kovariancia mátrix helyett. Ennek az oka elsősorban az, hogy a manifeszt változók mértékegysége gyakran önkényes a társadalomtudományokban, így indokoltabb a változókat standardizálni a mintabeli szórásokkal.

Ha a változók mérésénél a mértékegységet változtatjuk, akkor ez hatással van a változók szórásaira. Ha a szórásokkal standardizált változók között számítjuk a kovarianciákat, akkor azok a korrelációs együtthatókkal lesznek egyenlőek. Ez biztosítja, hogy az x változók mértékegységeinek (skálájának) változtatása nincs hatással az eredményekre. Azonban a korrelációs együtthatók együttes eloszlása nem azonos a kovarianciák együttes eloszlásával, így nem nyilvánvaló, hogy a becslések a valódi maximum likelihood becsléseket adják-e, és hogy az illeszkedés jóságát mérő statisztika érvényes-e.

Induljunk ki a faktorelemzés alapegyenleteiből, és nézzük meg, hogy a mérési skála mértékegységének változtatása hogyan változtatja meg a modellt. Legyen $\underline{C}$ diagonális mátrix, diagonális elemei legyenek pozitívok, és jelentsék az egyes változók skála-transzformációit, vagyis

$$\underline{x}^* = \underline{C}\,\underline{x}.$$

Az új, transzformált modell:

$$\underline{x}^* = \underline{C}\,\underline{\Lambda}\,\underline{f} + \underline{C}\,\underline{e} \tag{7}$$

$$\mathrm{cov}(\underline{x}^*) = \underline{\Sigma}^* = \underline{C}\,\underline{\Lambda}\,\underline{\Lambda}'\,\underline{C} + \underline{C}\,\underline{\Theta}\,\underline{C}.$$

Ha $\underline{C} = (\mathrm{diag}\underline{\Sigma})^{-\frac{1}{2}}$, vagyis az x változókat a szórásukkal osztjuk, akkor a (7) egyenletet a következőképpen írhatjuk:

$$\underline{x}^* = \underline{\Lambda}^*\,\underline{f} + \underline{e}^*, \tag{8}$$

és

$$\underline{\Sigma}^* = \underline{\Lambda}^*\,\underline{\Lambda}^{*\prime} + \underline{\Theta}^*,$$

ahol

$$\underline{\Lambda}^* = \underline{C}\,\underline{\Lambda} \text{ és } \quad \underline{\Theta}^* = \underline{C}\,\underline{\Theta}\,\underline{C}.$$

A fentiekből látható, hogy $\underline{x}$ mérési skálaegységeinek változtatása nincs hatással $\underline{x}^*$ értékeire, és így a $\underline{\Lambda}^*$ és $\underline{\Theta}^*$ paraméterek becslése is független $\underline{x}$ mérési skálájától, így ezek a paraméterek skála-invariánsak.

A $\underline{\Lambda}^*$ és $\underline{\Theta}^*$ paramétereket becsülhetjük úgy, hogy először becsüljük a $\underline{\Lambda}$ és $\underline{\Theta}$ paramétereket, és ezek alapján határozzuk meg $\underline{\Lambda}^*$ és $\underline{\Theta}^*$ becsléseit:

$$\widehat{\underline{\Lambda}}^* = (\mathrm{diag}\widehat{\underline{\Sigma}})^{-\frac{1}{2}}\,\widehat{\underline{\Lambda}} \quad \text{és} \quad \widehat{\underline{\Theta}}^* = (\mathrm{diag}\widehat{\underline{\Sigma}})^{-\frac{1}{2}}\widehat{\underline{\Theta}}(\mathrm{diag}\widehat{\underline{\Sigma}})^{-\frac{1}{2}}.$$

Egy becslési eljárás skálafüggetlen, ha a modell paramétereitől függő $\underline{\Sigma}$ mátrix és a mintabeli $\underline{S}$ mátrix illeszkedését mérő függvény minimuma nem függ a változók skálájától. Eszerint skálafüggetlen becslési eljárásnál az illesztett függvény minimuma megegyezik akár a kovariancia mátrixot, akár a korrelációs mátrixot elemezzük.

A klasszikus legkisebb négyzetek módszere függ a manifeszt változó skálájától, így ez a becslési eljárás skálafüggő.

Könnyen belátható (lásd KRANE és MCDONALD, 1978), hogy a maximum likelihood becslési eljárás invariáns a skála–transzformációval szemben.

3.3 A konfirmatív faktormodell paramétereinek becslése

Az identifikáció vizsgálata után — ha a modell identifikálhatónak bizonyult — kezdhetünk a modell paramétereinek becsléséhez. Keressük a paramétereknek azt a becslését, amelyekkel a lehető legjobban reprodukálni tudjuk a manifeszt változók mintabeli variancia–kovariancia mátrixát. A populáció variancia–kovariancia mátrixa ($\underline{\Sigma}$) a faktormodell szerint a $\underline{\Lambda}$, $\underline{\Phi}$ és $\underline{\Theta}$ paramétermátrixok függvénye: $\underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}$.

A populáció variancia–kovariancia mátrixának becsléséhez a populáció paramétereinek becslésével juthatunk: $\widehat{\underline{\Sigma}} = \widehat{\underline{\Lambda}}\,\widehat{\underline{\Phi}}\,\widehat{\underline{\Lambda}}' + \widehat{\underline{\Theta}}$. A becsült paramétermátrixoknak ki kell elégíteni az *a priori* feltételeket, és a becslési eljárásnak olyan $\underline{\Sigma}$ becslést kell generálni, melynek elemei a lehető legjobban illeszkednek a mintabeli variancia–kovariancia mátrix megfelelő elemeihez. Az illeszkedés jóságát különböző függvényekkel mérhetjük. Általánosan az illeszkedést mérő függvényt jelölje $F(\underline{S}, \underline{\Sigma})$ vagy $F(\underline{S}, \underline{\Sigma}(\underline{\Lambda}, \underline{\Phi}, \underline{\Theta}))$.

Azokat a $\underline{\Lambda}, \underline{\Phi}, \underline{\Theta}$ paramétermátrixokat tekintjük a sokaság paraméterei becslésének (ezeket $\widehat{\underline{\Lambda}}, \widehat{\underline{\Phi}}, \widehat{\underline{\Theta}}$ jelöli), amelyek esetén az F függvény felveszi minimumát.

A konfirmatív faktorelemzésnél háromféle illeszkedő függvényt alkalmaznak. Ezek a függvények háromféle statisztikai becslési eljáráshoz kapcsolódnak:
1) a súlyozatlan vagy klasszikus legkisebb négyzetek módszere (ULS),
2) az általánosított legkisebb négyzetek módszere (GLS), és 3) a maximum likelihood (ML) módszere.

3.3.1 A legkisebb négyzetek módszere

A legkisebb négyzetek módszere (ULS) a mintabeli variancia–kovariancia mátrixnak ($\underline{S}$) és a konfirmatív faktormodell variancia–kovariancia mátrixának ($\underline{\Sigma}$) megfelelő elemei közötti eltérés négyzetösszegét minimalizálja:

$$F_{ULS}(\underline{S}, \underline{\Sigma}) = tr[(\underline{S} - \underline{\Sigma})^2]\,. \qquad (9)$$

Az ULS becslés nem tételez fel a megfigyelt változók eloszlásáról semmit. Ez a legnagyobb előnye. Hátránya viszont, hogy így az illeszkedés jóságát nem tudjuk statisztikai próbával vizsgálni. Másik hátránya, hogy a becslési eljárás skálafüggő, vagyis függ a megfigyelt változók mértékegységétől.

3.3.2 Az általánosított legkisebb négyzetek módszere

Az általánosított legkisebb négyzetek módszere (GLS) a mintabeli variancia-kovariancia mátrixnak ($\underline{S}$) és a konfirmatív faktormodell $\underline{\Sigma}$ mátrixának megfelelő elemei közötti különbséget súlyozza az $\underline{S}^{-1}$ mátrix elemeivel, és a súlyozott eltérés-négyzetösszeget számítja:

$$F_{GLS}(\underline{S}, \underline{\Sigma}) = tr[(\underline{S} - \underline{\Sigma})\underline{S}^{-1}]^2 . \qquad (10)$$

Bár az általánosított legkisebb négyzetek módszere nem tételezi fel, hogy a megfigyelt változók valamilyen meghatározott valószínűségeloszlást követnek, viszont nagy előnye, hogy ha $\underline{x}$ normális eloszlású, a GLS becslés aszimptotikus tulajdonságú, és (a mintabeli becslés várható értéke közelíti a populáció paraméterének értékét, valamint a standard hiba a lehető legkisebb), és aszimptotikusan közelít a maximum likelihood becsléshez.

3.3.3 A maximum likelihood becslés

A maximum likelihood (ML) becslésnél feltételezzük, hogy a megfigyelt változók együttes eloszlása normális:

$$\begin{aligned} f(\underline{x}) &= (2\pi)^{-m/2}|\underline{\Sigma}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\underline{x} - \underline{\mu})'\underline{\Sigma}^{-1}(\underline{x} - \underline{\mu})\} = \\ &= (2\pi)^{-m/2}|\underline{\Sigma}|^{-\frac{1}{2}} \exp\{-\frac{1}{2}tr[(\underline{x} - \underline{\mu})(\underline{x} - \underline{\mu})'\underline{\Sigma}^{-1}]\} , \end{aligned} \qquad (11)$$

ahol

$$\underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}.$$

Általában feltételezzük, hogy a megfigyelt változókat várható értéküktől való eltérésükkel mérjük, így $\underline{\mu} = 0$, és $f(\underline{x})$ egyszerűbb alakra hozható:

$$f(\underline{x}) = (2\pi)^{-m/2}|\underline{\Sigma}|^{-\frac{1}{2}} \exp[-\frac{1}{2}tr\left(\underline{x}\,\underline{x}'\underline{\Sigma}^{-1}\right)] . \qquad (12)$$

Ha az x változók együttes eloszlása normális, a minta

$$\underline{S} = \frac{1}{n-1}\sum_{i=1}^{n} \underline{x}_i \underline{x}_i' ,$$

variancia–kovariancia mátrixa $(n-1)$ szabadságfokú Wishart eloszlást követ, ahol

n a minta elemszáma,

$\underline{x}_i$ az i-edik megfigyelés m változóra vonatkozó mérési értékeit tartalmazza (az x változókról feltételezzük, hogy standardizáltak).

A likelihood függvény logaritmusa:

$$\log_e L = -\frac{1}{2}(n-1)\{m \log_e 2\pi + \log_e |\underline{\Sigma}| + tr\left(\underline{S}\,\underline{\Sigma}^{-1}\right)\}. \tag{13}$$

A log–likelihood függvény maximalizálása helyett praktikus megfontolások miatt a gyakorlatban a következő F_{ML} függvényt minimalizáljuk:

$$F_{ML} = tr(\underline{S}\,\underline{\Sigma}^{-1}) + [\log_e |\underline{\Sigma}| - \log_e |\underline{S}|] - m, \tag{14}$$

($F_{ML} = -(c_1 \log_e L + c_2)$, ahol c_1 és c_2 konstansok.) F_{ML} minimalizálására JÖRESKOG (1967) javasolt jól működő eljárást. Az egyszerűség kedvéért tételezzük fel, hogy a közös faktorok korrelálatlanok ($\underline{\Phi} = \underline{I}$), és így a $\underline{\Sigma} = \underline{\Lambda}\,\underline{\Lambda}' + \underline{\Theta}$ modell paramétereit becsüljük. Jöreskog eljárása kétlépcsős, először az F_{ML} függvényt $\underline{\Lambda}$ szerint minimalizáljuk rögzített $\underline{\Theta}$ érték mellett, második lépésként a minimalizálást $\underline{\Theta}$ szerint végezzük rögzített $\underline{\Lambda}$ mellett.

Az első fázisban feltételezzük, hogy létezik F_{ML} minimuma rögzített $\underline{\Theta}$ mellett, ezt $f(\underline{\Theta})$-val jelöljük, és Λ_Θ-val azt a Λ-t, amelynél a minimum található.

$$f(\underline{\Theta}) = \min_{\underline{\Lambda}} F_{ML}(\underline{S}, \underline{\Sigma}(\underline{\Lambda}, \underline{\Theta})). \tag{15}$$

Második fázisban $\underline{\Theta}$ szerint minimalizálunk:

$$\min_{\underline{\Theta}} f(\underline{\Theta}) = \min_{\underline{\Lambda},\underline{\Psi}} F_{ML}. \tag{16}$$

Az adott $\underline{\Theta}$ mellett F_{ML} minimalizálása visszavezethető (a parciális deriváltakat egyenlővé téve $\underline{0}$-val) egy sajátérték–sajátvektor feladatra. Eszerint $\underline{\Lambda}_\Theta$ kifejezhető $\underline{\Theta}^{-\frac{1}{2}}\underline{S}\,\underline{\Theta}^{-\frac{1}{2}}$ sajátértékeivel és sajátvektoraival:

$$\underline{\Lambda}_\Theta = \underline{\Theta}^{\frac{1}{2}}\underline{\Gamma}(\underline{\Delta} - \underline{I})^{\frac{1}{2}},$$

ahol

$\underline{\Delta}$ diagonális mátrix, amelynek diagonálisa a $\underline{\Theta}^{-\frac{1}{2}}\underline{S}\,\underline{\Theta}^{-\frac{1}{2}}$ mátrix első r sajátértékét tartalmazza csökkenő sorrendben,

$\underline{\Gamma}$ $\underline{\Theta}^{-\frac{1}{2}}\underline{S}\,\underline{\Theta}^{-\frac{1}{2}}$ sajátvektorait tartalmazza a sajátértékeknek megfelelő sorrendben.

Könnyen belátható, hogy $\underline{\Lambda}_{\Theta}$ kielégíti a következő két feltételt:

$$\begin{aligned}\underline{\Lambda}_{\Theta}\underline{\Lambda}'_{\Theta} &= \underline{\Theta}^{\frac{1}{2}}\underline{\Gamma}(\underline{\Delta} - \underline{I})^{\frac{1}{2}}(\underline{\Delta} - \underline{I})^{\frac{1}{2}}\underline{\Gamma}'\underline{\Theta}^{\frac{1}{2}} \\ &= \underline{\Theta}^{\frac{1}{2}}[\underline{\Theta}^{-\frac{1}{2}}\underline{S}\,\underline{\Theta}^{-\frac{1}{2}} - \underline{I}]\underline{\Theta}^{\frac{1}{2}} \\ &= \underline{S} - \underline{\Theta},\end{aligned}$$

ami a modell alapegyenletének egy becslése, és

$$\begin{aligned}\underline{\Lambda}'_{\Theta}\underline{\Theta}^{-1}\underline{\Lambda}_{\Theta} &= (\underline{\Delta} - \underline{I})^{\frac{1}{2}}\underline{\Gamma}'\underline{\Gamma}(\underline{\Delta} - \underline{I})^{\frac{1}{2}} \\ &= (\underline{\Delta} - \underline{I}),\end{aligned}$$

ami diagonális mátrix, és ezzel $\underline{\Lambda}_{\Theta}$ a modell identifikálhatóságához bevezetett pótlólagos feltételt is kielégíti.

$\underline{\Lambda}_{\Theta}$ meghatározása után az $f(\underline{\Theta})$ függvényt minimalizáljuk a FLETCHER és POWELL (1963) eljárás szerint. Az iteratív eljárás során az egymást követő $\underline{\Theta}^{(1)}, \underline{\Theta}^{(2)}, \ldots$ mátrixokra teljesül, hogy

$$f(\underline{\Theta}^{(h+1)}) < f(\underline{\Theta}^{(h)}),$$

és minden lépésben újraszámítjuk a $\underline{\Theta}^{(h)}$-hoz tartozó $\underline{\Lambda}_{\Theta}^{(h)}$ mátrixot. Az eljárás gyorsan konvergál a végső becslés $\widehat{\underline{\Lambda}}$ és $\widehat{\underline{\Theta}}$ mátrixaihoz. Az eljárás első lépésénél szükségünk van a $\underline{\Theta}$ mátrix kezdeti becslésére, amit Jöreskog a következőképpen határozott meg:

$$\psi_{ii}^{(0)} = (1 - \frac{1}{2}r/m)(1/s^{ii}), \tag{17}$$

ahol

s^{ii} az $\underline{S}^{-1}$ mátrix i-edik diagonális eleme.

Az eljárásnak két lehetséges problémája van. Az egyik, hogy lokális minimumot találjunk a globális minimum helyett. Ez a gyakorlatban tapasztalatunk szerint ritkán fordul elő (lásd pl. JÖRESKOG és SÖRBOM, 1981). A másik lehetséges probléma, hogy az f függvény minimumhelyén a reziduális varianciák között előfordul negatív érték. Ezt az esetet az irodalomban *Heywood- esetnek* nevezik (HEYWOOD, 1931).

A Heywood-eset elsősorban a minta hibája, ezért elsősorban a minta elemszámával függ össze. Kis minta esetén ($n \leq 100$) nagy, nagy minta esetén ($n \geq 500$) kicsi az előfordulásának esélye.

Ha a modell egyébként helyesen definiált, a Heywood-eset előfordulásának valószínűsége közelít 0-hoz $n \longrightarrow \infty$ esetén.

Adott mintanagyság esetén a negatív variancia előfordulásának esélye csökken, ha a változók száma nő. Nő viszont a Heywood–eset bekövetkezésének valószínűsége, ha a változópárok korrelációi között előfordulnak nagyon magas értékek. Az ilyen változók közül az egyiket minden különösebb veszteség nélkül elhagyhatjuk. Ez azonban nem mindig szünteti meg a problémát.

A Heywood–eset előfordulása összefügghet természetesen a modellel magával is. Legkézenfekvőbb, amikor több faktort szerepeltetünk a modellben, mint amennyi elméletileg lehetséges.

Összefügghet a negatív variancia előfordulása a változók normális eloszlására vonatkozó feltétel megsértésével is.

A Heywood–eset előfordulása lehet a helytelen hiányzó adat–kezelés következménye is. A gyakorlati alkalmazásokban a kovariancia- vagy korrelációs mátrixot legtöbbször a páronkénti hiányzó adat–kezeléssel számítjuk. Ez azt jelenti, hogy minden megfigyelést figyelembe veszünk, ha az éppen vizsgált két változóra érvényes mérési eredménnyel rendelkeznek. Ebből következően minden kovarianciát vagy korrelációt más–más mintából becsülünk. Különösen akkor alkalmazzuk a páronkénti hiányzó adat–kezelést, ha a minta nagysága kicsi, ami még jobban növeli a Heywood–eset előfordulásának valószínűségét.

A páronkénti hiányzó adat–kezelést csak akkor alkalmazhatjuk, ha a hiányzó adatok száma mind a változóknál, mind a megfigyeléseknél kicsi.

Természetesen ha pontosan tudjuk a Heywood–eset előfordulásának okát, azt könnyen megszüntethetjük. Egyébként a problémát megkerülve vagy azt tehetjük, hogy az iterációs eljárást leállítjuk még azelőtt, hogy a hiba variancia negatívvá válna (valamilyen rögzített kicsi Θ_i érték, pl. .05 van .01 mellett), vagy ahhoz a változóhoz, amelynél a hiba komponens negatív varianciája előfordult, hozzáadunk egy ismert (σ^2) varianciájú véletlen változót, és az új változót alkalmazzuk a modellben:

$$var(x_i^*) = \sum_j^r \lambda_{ij} + \Theta_i + \sigma^2 ,$$

ahol

σ^2-et úgy kell meghatározni, hogy az x_i^* változók mellett a Heywood–eset ne forduljon elő.

Az x változó helyett az x^* változót alkalmazva az új korrelációs mátrixot úgy kaphatjuk meg, hogy a diagonálison kívüli elemeket megszorozzuk $(1+\sigma^2)^{-\frac{1}{2}}$-nel. Ebben az esetben az eredeti és a módosított modell paraméterei között a kapcsolat:

$$\lambda_{ij} = \lambda_{ij}^*(1+\sigma^2)^{\frac{1}{2}} \quad \text{és} \quad \Theta_i = 1 - (1+\sigma^2)\sum_j^r \lambda_{ij}^{*2} ,$$

ahol

λ_{ij}^*-ok jelölik a módosított modell faktorsúlyait.

3.4 A modell illeszkedésének vizsgálata

Az általánosított legkisebb négyzetek és a maximum likelihood módszer alkalmazása esetén — ha a becslési módszerek feltételei teljesülnek — khi–négyzet próbával tesztelhetjük a modell illeszkedését a mintabeli variancia–kovariancia mátrixhoz.

Legyen H_0 az adott modell nullhipotézise (a modell tökéletesen illeszkedik a populáció kovarianciamátrixához). H_0 alternatív hipotézise $L(H_1)$ azt fejezi ki, hogy $\underline{\Sigma}$ bármely pozitív definit, szimmetrikus mátrix lehet, amelynek a rangja m.

A khi–négyzet próba méri a modell becslésével kapott $\widehat{\underline{\Sigma}} = \widehat{\underline{\Lambda}}\,\widehat{\underline{\Phi}}\,\widehat{\underline{\Lambda}}' + \widehat{\underline{\Theta}}$ és a mintabeli variancia–kovariancia mátrix $(\underline{S})$ közötti eltérést. A khi–négyzet statisztikához tartozó szabadságfok a konfirmatív faktormodellben:

szfok= (a független paraméterek száma a H_1 hipotézis mellett)
— (a független paraméterek száma a H_0 hipotézis mellett)

$$\text{szfok} = \tfrac{1}{2}m(m+1) - t\,, \tag{18}$$

ahol

m a megfigyelt változók száma,

t a modell ismeretlen paramétereinek száma.

A modell becslésénél láttuk, hogy a paraméterek identifikálhatósága miatt fel kellett tételeznünk, hogy a $\underline{\Lambda}'\,\underline{\Theta}^{-1}\,\underline{\Lambda}$ mátrix diagonális mátrix, így a modell $\frac{1}{2}r(r-1)$ pótlólagos feltételt tartalmaz a paraméterekre. Ha feltételezzük, hogy a közös faktorok korrelálatlanok ($\underline{\Phi} = \underline{I}$), akkor a $\underline{\Lambda}, \underline{\Theta}$ paramétermátrixok

$$m + mr - \frac{1}{2}r(r-1)$$

szabad, ismeretlen paramétert tartalmaznak, így a szabadságfok:

$$\begin{aligned}\text{szfok} &= \frac{1}{2}m(m+1) - [m + mr - \frac{1}{2}r(r-1)] \\ &= \frac{1}{2}[(m-r)^2 - (m+r)]\,. \end{aligned} \tag{19}$$

Természetesen a konfirmatív faktormodellben általában nem tételezzük fel a közös faktorok függetlenségét, és a $\underline{\Lambda}$, $\underline{\Phi}$ és $\underline{\Theta}$ paramétermátrixok elemeire lehetnek pótlólagos feltételezéseink, pl. ha $\underline{\Phi}$ szimmetrikus mátrix, akkor a ϕ_{21} paramétert nem becsülhetjük, mivel $\phi_{12} = \phi_{21}$ a szimmetrikusság miatt. Vagy feltételezhetjük, hogy bizonyos faktorsúlyok egyenlőek egymással, pl. $\lambda_{11} = \lambda_{12} = \lambda_{13}$. Így egy adott modell független paramétereinek számát a konkrét feltételezések figyelembevételével határozhatjuk meg.

A becsült ($\hat{\underline{\Sigma}}$) és a mintabeli variancia–kovariancia mátrix ($\underline{S}$) eltérésének statisztikai tesztelésére a hagyományos khi–négyzet statisztika (χ^2) mellett a modell becslésénél definiált likelihood függvény is alkalmas.

Jelölje Ω az összes lehetséges $\underline{\Sigma}$ mátrixok halmazát (a pozitív definit, szimmetrikus, m rangú mátrixok halmazát). Legyen ω ennek azon részhalmaza, amelyhez tartozó kovarianciamátrixok megfelelnek a H_r hipotézisnek (a (2) egyenlettel definiált modellnek és a modell feltételeinek). Jelölje L_Ω és L_ω a megfelelő maximum likelihood függvények maximumát. Mivel az L függvény maximumát a Ω halmazon akkor veszi fel, ha $\underline{\Sigma} = \underline{S}$, így a (13) egyenlet a következő lesz:

$$\log_e L_\Omega = -\frac{1}{2}(n-1)(\log_e |\underline{S}| + m). \tag{20}$$

L_ω-t megkaphatjuk, ha $\underline{\Sigma}$-t helyettesítjük $\hat{\underline{\Sigma}} = \hat{\underline{\Lambda}}\hat{\underline{\Lambda}}' + \underline{\Theta}$ becslésével (feltételezve az egyszerűség kedvéért megint, hogy $\underline{\Phi} = \underline{I}$):

$$\log_e L_\omega = -\frac{1}{2}(n-1)[\log_e |\hat{\underline{\Sigma}}| + tr(\underline{S}\,\hat{\underline{\Sigma}}^{-1})]. \tag{21}$$

A likelihood hányados:

$$\lambda = L_\omega / L_\Omega. \tag{22}$$

Bár λ pontos valószínűségeloszlását nem ismerjük, de nagy minták esetén ismert, hogy $-2\log_e\lambda$ eloszlása közelítőleg khi–négyzet eloszlású, ha H_r igaz.

A (20) és (21) egyenletek alapján:

$$\begin{aligned} -2\log_e\lambda &= -2\log_e L_\omega + 2\log_e L_\Omega \\ &= (n-1)[\log_e |\hat{\underline{\Sigma}}| + tr(\underline{S}\,\hat{\underline{\Sigma}}^{-1}) - \log_e |\underline{S}| - m]. \end{aligned} \tag{23}$$

A (23) egyenletet összehasonlítva a (14) egyenlettel világosan látszik, hogy a likelihood hányados logaritmusának minusz kétszerese ($-2\log\lambda$) egyenlő F_{ML} minimumának $(n-1)$-szeresével. Vagyis a maximum likelihood becslés F_{ML} függvénye felhasználható a modell illeszkedésének statisztikai tesztelésére:

$$v = (n-1)\min F_{ML}. \tag{24}$$

aszimptotikusan khi–négyzet eloszlású, a szabadságfok egyenlő a mintabeli variancia–kovariancia mátrix független elemei számának és a becsült független paraméterek számának különbségével, általánosan

$$\text{szfok} = m(m+1)/2 - t,$$

ahol

t a modell becsült független paramétereinek száma.

A v statisztika a minta elemszámának, valamint $\underline{\hat{\Sigma}}$ és $\underline{S}$ illeszkedésének a függvénye.

Nagyon sok esetben fontos tudni, hogy a modell változtatása javított–e a modell illeszkedésén. Ha a vizsgált modellek hierarchikusak, vagyis ha a H_m modellt a H_b modellből kaphatjuk meg úgy, hogy a H_b modell egy vagy több paraméterére feltételeket teszünk, tehát ha a H_m modell lehetséges $\underline{\Sigma}_m$ variancia–kovariancia mátrixai részhalmazát képezik a H_b modell lehetséges $\underline{\Sigma}_b$ mátrixai halmazának (lásd pl. a ω és a Ω halmazokat), akkor a modelleknek megfelelő likelihood hányados statisztikák különbsége is khi–négyzet eloszlást követ:

$$(n-1)(F_m - F_b) \tag{25}$$

khi–négyzet eloszlású szfok$_m$ - szfok$_b$ szabadságfokkal.

Mivel a likelihood hányados statisztika (és a khi–négyzet statisztika is) függ a minta nagyságától, és eszerint kis minta esetén még a nyilvánvalóan nem megfelelő modell is elfogadható illeszkedést mutat, míg nagy minta esetén figyelembe kell venni majdnem minden lehetséges közös faktort, hogy elfogadhatóan illeszkedő modellt kapjunk, ami nem egyezik az elméleti törekvéssel, a khi–négyzet statisztikát próbálták transzformálni olymódon, hogy kevésbé legyen érzékeny a minta elemszámára.

Bentler és Bonett (1980) javasolta, hogy a khi– statisztikák felhasználásával 0 és 1 értékek közé eső mutatót számoljunk a modell kiértékelésére.

Az illeszkedés Bentler–Bonett–féle normált mutatója:

$$\triangle_1 = \frac{\chi_b^2 - \chi_m^2}{\chi_b^2}, \tag{26}$$

vagy

$$\triangle_1 = \frac{F_b - F_m}{F_b}, \tag{27}$$

ahol

χ_b^2 egy alapmodell (nullhipotézis, amely szerint a változók között nincs összefüggés) illeszkedését mérő χ^2 statisztika,

χ_m^2 a hipotetikus, javított modell χ^2 értéke,

F_b és F_m az illesztett függvény értékei a két modellnél.

$\triangle_1$ méri a hipotetikus, vizsgált modellilleszkedésének javulását az alapmodellhez viszonyítva.

$\triangle_1$ értéke 0 és 1 között mozog, közeledve 1–hez jobb modell illeszkedést mutat. Maximális értékét akkor veszi fel, ha χ_m^2 (vagy F_m) egyenlő nullával. χ_m^2 nagy minta esetén aszimptotikusan khi–négyzet eloszlást követ, és így χ^2_m várható értéke közelítőleg a szabadságfokával (szf_m) lesz egyenlő.

A χ^2_b becslő függvény csak bizonyos feltételek mellett (lásd STEIGER, 1985) követ nem–centrális khi– négyzet eloszlást.

A ($\bar{\chi}^2_b - \mathrm{szf}_m$) kifejezés (ahol $\bar{\chi}^2_b$ az alapmodell khi–négyzet értékének az átlaga) átlagértéket ad a vizsgált modell esetére, így a ($\chi^2_b - \chi^2_m$) különbséget viszonyíthatjuk ehhez a várható különbséghez. A $\triangle_2$ mutató (a $\bar{\chi}^2_b$ átlag helyett annak becslését, χ^2_b-t helyettesítettük):

$$\triangle_2 = \frac{\chi^2_b - \chi^2_m}{\chi^2_b - \mathrm{szf}_m}. \tag{28}$$

A (28) mutató magasabb értéket ad, ha kevesebb paraméterrel kapjuk ugyanazt a becsült χ^2_m értéket.

Az illesztett függvényértékekre:

$$\triangle_2 = \frac{F_b - F_m}{F_b - (\mathrm{szf}_m/(n-1))}, \tag{29}$$

ahol

n a minta elemszáma.

Ha $\triangle_2$ számlálója és nevezője is pozitív, akkor $\triangle_2$ nagyobb mint $\triangle_1$. $\triangle_2$ nem normalizált mutató, mivel lehetséges, hogy a 0 és 1 határokon kívüli értéket vesz fel. $\triangle_2$ maximuma $\chi^2_b/(\chi^2_b - \mathrm{szf}_b)$, feltéve hogy $\chi^2_b > \mathrm{szf}_b$. $\triangle_2$ minimuma 0 hierarchikus modellek esetén.

A $\triangle_2$ mutató általánosított formáját kapjuk, mikor a számlálóba az alapmodell helyett egy kevésbé korlátozott modell (de nem annyira, mint a vizsgált) khi–négyzet értékét tesszük.

Ekkor

$$\triangle_2 = \frac{\chi^2_r - \chi^2_m}{\chi^2_b - \mathrm{szf}_m}.$$

BOOLEN (1986) közölt a $\triangle_1$-hez hasonló, a szabadságfokokhoz viszonyított eltérést mérő mutatót:

$$\rho_1 = \frac{\left(\frac{\chi^2_b}{\mathrm{szf}_b}\right) - \left(\frac{\chi^2_m}{\mathrm{szf}_m}\right)}{\left(\frac{\chi^2_b}{\mathrm{szf}_b}\right)}, \tag{30}$$

vagy

$$\rho_1 = \frac{\left(\frac{F_b}{\mathrm{szf}_b}\right) - \left(\frac{F_m}{\mathrm{szf}_m}\right)}{\left(\frac{F_b}{\mathrm{szf}_b}\right)}. \tag{31}$$

ρ_1 normált mutató, értéke 0 és 1 között mozog. BENTLER és BONETT (1980) javasolta nem–normalizált változatát:

$$\rho_2 = \frac{\left(\frac{\chi_b^2}{\text{szf}_b}\right) - \left(\frac{\chi_m^2}{\text{szf}_m}\right)}{\left(\frac{\chi_b^2}{\text{szf}_b}\right) - 1} \tag{32}$$

$$\rho_2 = \frac{\left(\frac{F_b}{\text{szf}_b}\right) - \left(\frac{F_m}{\text{szf}_m}\right)}{\left(\frac{F_b}{\text{szf}_b}\right) - (1/(n-1))} . \tag{33}$$

A minta nagysága kétféleképpen hat az illeszkedés jóságát mérő mutatókra.

Az egyik hatás, amikor a minta elemszáma (n) közvetlenül részt vesz a számításnál.

Belátható (lásd pl. BOOLEN, 1986), hogy ρ_2 és $\triangle_2$ közvetlenül függ a mintanagyságtól, míg ρ_1 és $\triangle_1$ nem.

A másik hatás, amikor az illeszkedést mérő mutató mintabeli eloszlásának várható értéke függ a minta elemszámától. Ha egy illeszkedő modell mellett különböző elemszámú véletlen mintákat veszünk, és kiszámoljuk az illeszkedést mérő mutatókat és azok eloszlásait, akkor azt tapasztaljuk, hogy $\triangle_1$ és ρ_1 mintabeli eloszlásának várható értéke pozitív kapcsolatban van a minta elemszámával (n-nel), míg $\triangle_2$ és ρ_2 függősége közel nulla.

BOOLEN (1989) közölt Monte Carlo kísérletet a $\triangle_1, \triangle_2$, ρ_1 és ρ_2 mutatók eloszlásának vizsgálatára.

Vizsgálta az

$$\underline{x} = \underline{\Lambda}\,\underline{f} + \underline{e}$$

modellt, ahol a megfigyelt változók (x) száma $m = 5$, a faktorok (f) száma $r = 2$ volt. Boolen a konfirmatív faktormodellt alkalmazta, ahol a modell paramétereire a következő feltételezéseket tette:

$$\underline{\Lambda}_x = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} , \qquad \underline{\Phi} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} ,$$

$$\text{diag}\underline{\Theta} = [1, 1, 1, 1, 1] ,$$

ahol

$\underline{\Phi}$ a közös faktorok variancia–kovariancia mátrixa,

$\underline{\Theta}$ a hiba komponensek variancia–kovariancia mátrixa.

A megfigyelt változók eloszlása rendre normális eloszlású volt. Boolen három mintanagyságot vizsgált (75, 150, 300), és minden mintanagyságra tizennégyszer végezte el a számításokat a maximum likelihood becslést alkalmazva (F_{ML}). A következő táblázat mutatja a modell illeszkedését mérő mutatók átlagait és szórásait:

Az illeszkedés jóságát mérő mutatók	Várható érték (szórás) Minta		
	75	150	300
$\triangle_2$	1.003	1.002	1.000
	(.019)	(.012)	(.008)
$\triangle_1$	.967	.985	.992
	(.018)	(.012)	(.008)
ρ_2	1.008	1.005	1.001
	(.052)	(.030)	(.021)
ρ_1	.918	.963	.980
	(.046)	(.030)	(.020)

A táblázatból látható, hogy $\triangle_1$ és ρ_1 várható értéke függ a minta elemszámától, míg $\triangle_2$ és ρ_2 várható értéke közel esik 1-hez mindegyik mintanagyság esetén.

A másik jellemző tanulság a fenti vizsgálatból, hogy a $\triangle$ mutatók szórása lényegesen kisebb (kevesebb mint fele), mint a ρ mutatóké.

Ezenkívül látható még, hogy a $(\triangle_2 - \triangle_1)$ és a $(\rho_2 - \rho_1)$ különbségek csökkennek a minta elemszámának növekedésével.

3.5 A faktorok értelmezése és transzformálása

Miután a faktormodellt illesztettük az adatokhoz, a paraméterek becslése ismeretében a paraméterek és a faktorok értelmezése következhet. A továbbiakban is feltételezzük, hogy a közös faktorok korrelálatlanok.

Az értelmezést a következő egyenlettel kezdjük:

$$var(x_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \ldots + \lambda_{ir}^2 + \Theta_i \tag{34}$$

$$var(x_i) = \sum_{k=1}^{r} \lambda_{ik}^2 + \Theta_i \qquad i = 1, \ldots, m.$$

A (34) egyenlet az i-edik megfigyelt változó varianciáját két tagra bontja. Az első tag az i-edik változó varianciájának az r-számú közös faktorral magyarázható része, a második tag a hiba komponens (vagy egyedi faktor) által magyarázott rész.

Az első tag az adott változó kommunalitása. Általában a változókat standardizáljuk, így

$$1 = var(x_i) = \sum_{k=1}^{r} \lambda_{ik}^2 + \Theta_i .$$

Ha az együttes varianciát vizsgáljuk, összegezve a (34) egyenleteket $i = 1, \ldots, m$-re, a következőket kapjuk:

$$\sum_i^m var(x_i) = \sum_i^m \lambda_{i1}^2 + \sum_i^2 \lambda_{i2}^2 + \ldots + \sum_i^m \lambda_{ir}^2 + \sum_i^m \Theta_i , \qquad (35)$$

ahol a $\Sigma_i^m \lambda_{ik}^2$ a k-adik faktor hozzájárulását fejezi ki a teljes varianciához. Ezt kifejezhetjük relatív formában is:

$$\frac{\Sigma_i^m \lambda_{ik}^2}{\Sigma_i^m var(x_i)} .$$

Szokás ezt a kifejezést százalékos formában is használni. Ilyenkor azt mondjuk, hogy az adott faktor az összvariancia

$$\frac{\Sigma \lambda_{ik}^2}{\Sigma var(x_i)} \cdot 100 \text{ \%-át magyarázza} .$$

A faktorok értelmezését a λ_{ik} faktorsúlyok segítségével adhatjuk meg. A λ_{ik} faktorsúlyt mint standardizált regressziós együtthatót értelmezhetjük. A λ_{ik} kifejezhető az i-edik változó és a k-adik faktor közötti korrelációs együtthatóval, ha a közös faktorok korrelálatlanok. A k-adik faktor értelmezéséhez a λ_{ik} faktorsúlyok relatív nagyságát és előjelét kell figyelembe vennünk. Minél nagyobb a λ_{ik} érték, annál szorosabban kapcsolódik az i-edik változó a k-adik faktorhoz. Az azonos előjelű, nagy relatív súlyu változókat összegyűjtve az ezen változókban lévő közös tulajdonság, közös jellemző fejeződik ki a faktorban. Ha minden változó azonos mértékben és előjellel kapcsolódik a faktorhoz, akkor az adott faktort általános faktornak nevezhetjük. Gyakran előfordul, hogy a változók egy része pozitívan, egy másik része negatívan kapcsolódik az adott faktorhoz. Ilyenkor az adott faktor egyik irányban azt fejezi ki, amiben az egyik változóhalmaz közös ugyanakkor, amikor a másik változóhalmaz közös jellemzője hiányzik, míg másik irányban fordítva mér. Ezt a faktort bipoláris faktornak nevezzük.

Azokat a faktorokat lehet könnyen értelmezni, amelyekben λ egy része közel esik 0-hoz, a másik része hasonló nagyságú és előjelű. Ha a faktorsúlyokat nem lehet eszerint szeparálni, akkor a faktorteret transzformálhatjuk úgy, hogy a faktorok értelmezése lehetséges (vagy könnyebb) legyen.

A paraméterek becslése — mint láttuk — kielégíti azt a feltételt, hogy a $\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda}$ mátrix diagonális legyen. Ez azt jelenti, hogy a faktorokat a főtengelyek mentén határozzuk meg. Az első faktor a pontok legnagyobb szóródása irányába mutat, a második a lehetséges második legnagyobb szóródás irányába, de merőlegesen az elsőre, és így tovább. A gyakorlatban nem biztos, hogy ez a legjobb választás, és így a faktorok értelmezése meglehetősen nehéz lehet. A faktorteret transzformálhatjuk úgy, hogy a modell illeszkedése ne változzon meg, de a faktorsúlyok mátrixa lehetőleg egyszerű struktúrát mutasson. Egy faktormátrixot egyszerű struktúrának nevezünk, ha a változók a lehető legkevesebb faktorral kapcsolódnak 0-tól különböző súllyal.

Tekintsük az irodalomban sokszor szereplő példát a faktortér transzformációjára (lásd pl. EVERITT, 1984).

A következő táblázat a hat tantárgy iskolai eredményei közötti korrelációt mutatja egy 220 iskolás fiú adatait tartalmazó mintában.

Táblázat. Iskolai tantárgyak korrelációs mátrixa

	Iskolai tantárgyak					
	1	2	3	4	5	6
1 Francia nyelv	1.00					
2 Angol nyelv	.44	1.00				
3 Történelem	.41	.35	1.00			
4 Számtan	.29	.35	.16	1.00		
5 Algebra	.33	.32	.19	.59	1.00	
6 Geometria	.25	.33	.18	.47	.46	1.00

A faktormodell illesztésekor két közös faktort találtak. Az eredményeket a következő táblázat tartalmazza.

Táblázat. Iskolai tantárgyak 2–faktoros modelljének eredményei

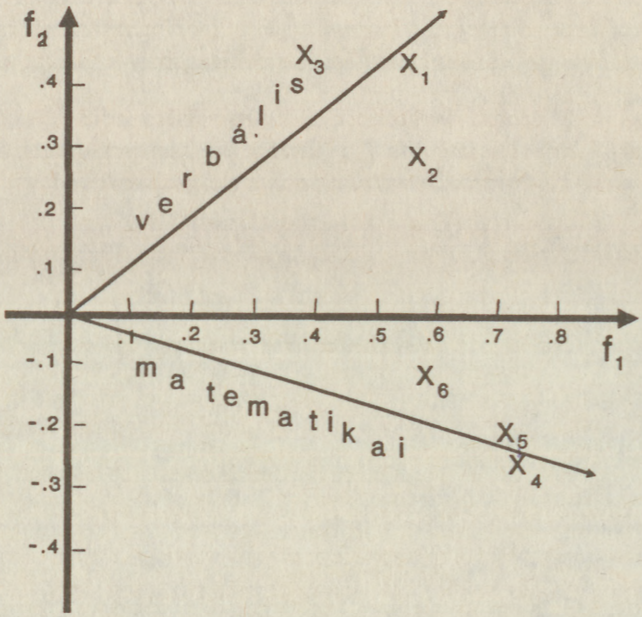
Tantárgyak	Faktorsúlyok		Kommunalitások
1 Francia nyelv	.55	.43	.49
2 Angol nyelv	.57	.29	.41
3 Történelem	.39	.45	.36
4 Számtan	.74	–.27	.62
5 Algebra	.72	–.21	.57
6 Geometria	.59	–.13	.37

A fenti faktormátrixból látható, hogy az első faktor, amelyik általános faktornak tekinthető, mindegyik változóval pozitívan, viszonylag erősen korrelál, s általános intelligencia faktornak nevezhetjük el.

A második faktor a változókat két részre bontja, egyik részével pozitívan, másikkal negatívan korrelál, ezért ez a faktor bipoláris, és a verbális kontra matematikai képességet fejezi ki.

A második faktor azt mutatja, hogy egy adott általános intelligenciaszinten azok a tanulók, akik jó eredményeket érnek el a verbális tárgyakból, kevésbé jók a matematikai tárgyakból, és fordítva.

A faktorsúlyok mátrixát ábrázolhatjuk úgy, hogy a tengelyek a faktorokat reprezentálják, a faktorsúlyok a változóknak a faktorokra vonatkozó koordinátái:



A faktorok grafikus rotálásánál a tengelyeket úgy forgatjuk, hogy lehetőleg elkerüljük a negatív faktorsúlyokat, és a lehető legkevesebb nullától különböző súly maradjon.

A faktortengelyeket rotálva a két változócsoporthoz, eredményül két korrelált faktort kapunk, az egyiket a verbális képességek, a másodikat a matematikai képességek faktorának nevezhetjük.

Láttuk korábban, hogy a faktorok ortogonális transzformációja a modell illeszkedését nem változtatja meg. Általánosságban a faktorok grafikus rotálását az

$$\underline{M} = \begin{pmatrix} \cos\Theta & \sin\Theta \\ -\sin\Theta & \cos\Theta \end{pmatrix}$$

ortogonális mátrix segítségével végezhetjük Θ forgásszöggel az óra járásával azonos irányban. Az új faktorsúlyok mátrixa:

$$\underline{\Lambda}^* = \underline{\Lambda}\,\underline{M}.$$

A forgásszöget megkaphatjuk úgy, hogy az ábráról lemérjük. Az óra járásával ellentétes irányú forgatásmódot adja az alábbi ortogonális transzformációs mátrix:

$$\underline{M} = \begin{pmatrix} \cos\Theta & -\sin\Theta \\ \sin\Theta & \cos\Theta \end{pmatrix}.$$

A rotálás numerikus módszerei közül a leggyakrabban használt a *varimax* módszer. A varimax módszer célja olyan súlyokat találni, hogy lehetőleg csak abszolút értékben viszonylag nagy vagy viszonylag kicsi faktorsúlyok forduljanak elő. A varimax eljárás minden változóra maximalizálja a faktorsúlyok négyzeteinek

$$S_v = \sum_k^r \left[\frac{m\Sigma_i^m (\lambda_{ik}^2)^2 - (\Sigma_i^m \lambda_{ik}^2)^2}{m^2} \right] \tag{36}$$

varianciáját. A varimax módszer részletes leírását lásd pl. LAWLEY és MAXWELL (1971) 6. fejezetében.

A varimax rotálás ortogonális faktorokat eredményez. Hasonlóan ortogonális transzformációs eljárás a *quartimax*. A quartimax rotálásnál a faktorsúlyok negyedik hatványának

$$S_q = \sum_i^m \sum_k^r \lambda_{ik}^4 \tag{37}$$

összegét maximalizáljuk. A nem ortogonális transzformációk közül, amelyek korrelált, ferdeszögű faktorokat eredményeznek, a leggyakrabban az oblimin rotálást alkalmazzuk. Az oblimin rotálásnál a különböző faktorok faktorsúlyai négyzeteinek

$$S_{d0} = \sum_\ell^r \sum_{k \neq \ell}^r \left[\sum_i^m \lambda_{i\ell}^2 \lambda_{ik}^2 - \frac{1}{m} \sum_i^m \lambda_{i\ell}^2 \sum_i^m \lambda_{ik}^2 \right] \tag{38}$$

kovarianciáit minimalizáljuk. Ha a faktorok korrelálnak egymással, akkor a faktorsúlyokat már nem értelmezhetjük mint a változók és a faktorok közötti korrelációs együtthatókat.

3.6 Ipszatív változók faktorelemzése

Az ipszatív transzformáció fogalmát CATTELL (1944) vezette be (a latin ipse = ő maga szóból) egy változóhalmaz értékeinek a megfigyelési egységek átlagaihoz történő centirozására. Az ipszatív transzformációt HORST (1965) jobboldali centirozásnak nevezte, megkülönböztetve a változók normatív transzformálásától (amikor a változókat saját átlagukhoz centirozzuk), amit baloldali centirozásnak nevezett. Amikor a változók értékeit transzformáljuk mind a két módon, akkor dupla centirozásról beszélünk.

Az ipszatív transzformálás eredményeként a változók értékeinek összege minden megfigyelési egységnél egyenlő egy konstanssal. A konstans a gyakorlatban egyenlő 0-val, mivel a megfigyelési egységek ipszatív értékei a vizsgált változóhalmaz értékeinek eltérései az adott megfigyelési egység átlagától. Általában az ipszatív fogalmat használjuk tekintet nélkül a konstans értékére, ha a változók értékeinek az összege minden megfigyelésnél egyenlő a konstans értékkel.

A gyakorlatban egy változóhalmaznak lehet ipszatív tulajdonsága akkor is, ha az nem ipszatív transzformáció eredménye, például rangsorolt változók esetében.

Legyen $\underline{y}$ m megfigyelt valószínűségi változó vektora. Tételezzük fel, hogy mindegyik változót azonos mértékegységgel, azonos skálán mérjük, és hogy $\underline{y}$ variancia–kovariancia mátrixa pozitív definit.

Az $\underline{y}$ változó ipszatív transzformációját általánosan a következőképpen írhatjuk:

$$\underline{x} = \underline{y} - \underline{1}w, \tag{39}$$

ahol

$\underline{1}$ összegzővektor,

w skalár,

$\underline{x}$ $(m \times 1)$ típusú vektor, ipszatív tulajdonsággal.

Általában $w = m^{-1}(\underline{1}'\underline{y} - c)$, ahol c konstans, a változók értékeinek összege minden megfigyelésnél. c értéke általában indifferens, így célszerű 0-nak választani. Ekkor $w = m^{-1}\underline{1}'\underline{y}$. A (39) transzformációban a változók átlagát választjuk skalárnak, és így

$$\begin{aligned} \underline{x} &= \underline{y} - \underline{1}(m^{-1}\underline{1}'\underline{y}) \\ &= \underline{y} - m^{-1}\underline{1}\,\underline{1}'\underline{y} \\ &= (\underline{I} - m^{-1}\underline{U})\underline{y}, \end{aligned} \tag{40}$$

ahol

$\underline{U} = \underline{1}\,\underline{1}'$ $(m \times m)$ típusú összegzőmátrix.

Az ipszatív transzformáció $(\underline{I} - \frac{1}{m}\underline{U})$ mátrixa szimmetrikus, idempotens mátrix, a rangja $(m-1)$. (Lásd pl. HORST, 1965, 288–289).

Jelöljük az ipszatív transzformáció mátrixát $\underline{A}$-val:

$$\underline{A} = (\underline{I} - \frac{1}{m}\underline{U}), \qquad \underline{x} = \underline{A}\,\underline{y}.$$

Az $\underline{x}$ variancia–kovariancia mátrixa:

$$\Sigma_x = \underline{A}\,\Sigma_y\,\underline{A}, \tag{41}$$

ahol

$\underline{\Sigma}_y$ a megfigyelt változók variancia-kovariancia mátrixa.

Mivel $\underline{A}$ rangja $(m-1)$, (41) $\underline{\Sigma}_y$ szinguláris transzformációja, ezért $\underline{\Sigma}_x$ is szinguláris lesz, így nem elemezhető ugyanúgy, mint $\underline{\Sigma}_y$.

Az $\underline{\Sigma}_x$ ipszatív variancia–kovariancia mátrix tulajdonságai (lásd CLEANANS, 1956):

- $\underline{\Sigma}_x$ sorainak (vagy oszlopainak) az összege nullával egyenlő,
- egy ipszatív változóhalmaz és egy kritérium, függő változó kovarianciáinak az összege egyenlő 0-val,
- ha m, a változók száma nő, $\underline{\Sigma}_x$ közelít $\underline{\Sigma}_y$-hoz.

Mivel $\underline{\Sigma}_x$ nem pozitív definit, nem alkalmazhatjuk a faktorelemzés modelljét közvetlenül a $\underline{\Sigma}_x$ mátrixra.

Tételezzük fel viszont, hogy a megfigyelt változókra a faktorelemzés modellje illeszthető:

$$\underline{y} = \underline{\Lambda}_y \underline{f} + \underline{e}. \tag{42}$$

A modell feltételei:

$$\begin{aligned} E(\underline{y}) &= E(\underline{f}) = E(\underline{e}) = \underline{0} \\ E(\underline{f}\,\underline{e}') &= \underline{0} \\ E(\underline{f}\,\underline{f}') &= \underline{\Phi}_y \\ E(\underline{e}\,\underline{e}') &= \underline{\Theta}_y . \end{aligned}$$

A faktormodell kovariancia egyenlete:

$$E(\underline{y}\,\underline{y}') = \underline{\Sigma}_y = \underline{\Lambda}_y \underline{\Phi}_y \underline{\Lambda}_y + \underline{\Theta}_y . \tag{43}$$

A (40) egyenlet szerinti ipszatív transzformációt elvégezve, az ipszatív változók faktormodelljét kapjuk:

$$\begin{aligned} \underline{x} &= (\underline{\Lambda}_y \underline{f} + \underline{e}) - \underline{1} m^{-1} \underline{1}'(\underline{\Lambda}_y \underline{f} + \underline{e}) \\ &= \underline{\Lambda}_y \underline{f} + \underline{e} - \underline{1}(\overline{\underline{\lambda}'}\underline{f} + \bar{e}) , \end{aligned} \tag{44}$$

ahol

$\overline{\underline{\lambda}'}$ az $\underline{\Lambda}_y$ faktorsúlymátrix oszlopainak átlagvektora,

$\bar{e}$ a hiba komponensek átlaga.

Tovább egyszerűsítve a (44) egyenletet, az ipszatív változók faktormodelljét a következőképpen írhatjuk:

$$\underline{x} = (\underline{\Lambda}_y - \underline{1}\,\overline{\underline{\lambda}'})\underline{f} + \underline{A}\,\underline{e}, \tag{45}$$

ahol

$$\underline{A}\,\underline{e} = \underline{e} - \underline{1}\bar{e},$$

$$\underline{A} = (\underline{I} - m^{-1}\underline{U}),$$

vagy

$$\underline{x} = \underline{\Lambda}_x \underline{f} + \underline{A}\,\underline{e}, \tag{46}$$

ahol

$$\underline{\Lambda}_x = (\underline{\Lambda}_y - \underline{1}\,\underline{\lambda}'),$$

$$\underline{A} = (\underline{I} - m^{-1}\underline{U}) = (\underline{I} - m^{-1}\underline{1}\,\underline{1}').$$

Az ipszatív változók faktorsúlymátrixa az $\underline{\Lambda}_y$ faktormátrixnak és az átlagos faktorsúlyok mátrixának a különbsége. A hiba komponensek az ipszatív modellben korrelálnak egymással (mivel $\underline{A}$ nem diagonális mátrix).

Az ipszatív $\underline{x}$ változók kovarianciamátrixa:

$$E(\underline{x}\,\underline{x}') = \underline{\Sigma}_x = \underline{\Lambda}_x \underline{\Phi}_y \underline{\Lambda}_x + \underline{A}\,\underline{\Theta}_y \underline{A}. \tag{47}$$

Látható, hogy az ipszatív változók faktormodelljében a faktorok megegyeznek a preipszatív modell faktoraival, ugyanígy a faktorok kovarianciamátrixa is változatlan maradt. Az ipszatív faktorsúlyok mátrixa az eredeti faktormátrix elemeinek eltérését tartalmazza a faktorsúlymátrix megfelelő oszlopainak átlagától. A hiba komponensek az ipszatív transzformáció hatása miatt korreláltakká váltak.

A gyakorlatban ipszatív tulajdonságú változókhoz általában nem ipszatív transzformációval, hanem elsősorban rangsorolással jutunk akkor, amikor a vizsgált személyektől azt kérjük, hogy valamilyen szempont szerint rangsorolják a változóhalmaz elemeit. Így a mérés eredménye rangszám lesz, és a változóhalmaz ipszatív tulajdonságú.

Mivel $\underline{\Sigma}_x$ ebben az esetben szinguláris, a $\underline{\Sigma}_x$ mátrixból el kell hagynunk egy sort és egy oszlopot, önkényesen elhagyva az $\underline{x}$ változók közül egyet. Jelölje $\underline{x}^*$ az 1-gyel csökkentett számú ipszatív változókat ($\underline{x}^*$ $(m-1)$ változót tartalmaz). A $\underline{x}^*$ változók $\underline{\Sigma}_{x^*}$ kovarianciamátrixához illeszthetjük a faktormodellt. Az elhagyott változó faktorsúlyát a faktormátrix minden oszlopában megkaphatjuk úgy, hogy az $(m-1)$ változó faktorsúlya összegének az ellentettjét vesszük (−1-gyel szorozzuk) minden oszlopban, és ezt feleltetjük meg a faktormátrix hiányzó változójának súlyával.

ALWIN és JACKSON (1981) három feltételt nevezett meg, amely teljesülése esetén az ipszatív tulajdonságú változókra alkalmazhatjuk a faktorelemzés modelljét: (1) az adott ipszatív tulajdonságú változóhalmaz mellett létezik a populációban egy ezzel összefüggő, hipotetikus nem–ipszatív változóhalmaz, (2) létezik az ipszatív transzformáció a két változóhalmaz között, (3) teljesülnek a faktormodell feltételei a hipotetikus nem–ipszatív változókra.

3.7 Dichotom változók faktorelemzése

Jelölje $\underline{y}^*$ a megfigyelt manifeszt változókat, számuk m, $\underline{\eta}$ pedig a faktorokat (számuk r).

A faktorelemzés modellje:

$$\underline{y}^* = \underline{\Lambda}\,\underline{\eta} + \underline{\varepsilon},$$

ahol

$\underline{\Lambda}$ a faktorsúlyok mátrixa $(m \times r)$ típusú,

$\underline{\varepsilon}$ az m számú reziduálisokat jelöli.

A $\underline{y}^*$ megfigyelt változó variancia–kovariancia mátrixa a szokásos feltételezésekkel:

$$V(\underline{y}^*) = \underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta},$$

ahol

$\underline{\Phi}$ a faktor kovariancia mátrix,

$\underline{\Theta} = \underline{I} - \operatorname{diag}(\underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}')\underline{\Theta}$ a reziduális kovariancia mátrix (diagonális).

Mivel a faktorokat és a reziduálisokat folytonos változóknak tekintjük, így y_i^*-k is folytonosak, azonos skálájú változók. Tegyük fel, hogy minden y_i^* változóhoz tartozik egy dichotom ((0,1) értékű) y változó. A klasszikus faktorelemzés alkalmazása esetén feltételezhető, hogy $y = y^*$ minden változóra. Ez azonban eléggé önkényes feltételezés, mivel a modell jobboldalán folytonos, intervallum mérési skálájú változók, a baloldalon diszkrét, dichotom változók állnak. Ez a probléma jól ismert a regresszióelemzés alkalmazásánál és mint az OLS regresszió versus logisztikus regresszió polémia jelenik meg az irodalomban. A probléma lényege az, hogy $\underline{y}^*$ és $\underline{\eta}$ között a modell lineáris, míg $\underline{y}$ és $\underline{\eta}$ között nem lineáris a kapcsolat.

Tegyük fel, hogy y_i^* változó egy pozitív irányú látens hatást tükröz. Ha ez a hatás eléri az adott változóra jellemző küszöbértéket, akkor a változó kifejez egy szimptómát, egyébként nem. Ha τ_i jelöli ezt a küszöbértéket, akkor ezt a következőképpen írhatjuk:

$$y_i = \begin{cases} 0, & \text{ha } y_i^* < \tau_i \\ 1, & \text{egyébként} \end{cases}.$$

Az y_i változó binomiális eloszlású, feltételezve, hogy a reziduálisok függetlenek $\underline{\eta}$–tól és normális, vagy logisztikus eloszlásúak, az y_i és $\underline{\eta}$ közötti nemlineáris kapcsolat:

$$P(y_i = 0 \mid \underline{\eta}) = P(y_i^* < \tau_i) = F(\tau_i - \underline{\lambda}'\underline{\eta}),$$

ahol

F standard normális, vagy logisztikus eloszlásfüggvény, és feltételezzük, hogy a reziduális variancia standardizált.

Ha az $y = y^*$ megfeleltetést vesszük, és a faktorelemzés modelljét alkalmazzuk, akkor a Pearson–féle korrelációt alkalmazzuk a dichotom változókra, ami azt jelenti, hogy a lineáris modell alapján becsüljük a paramétereket.

Tudjuk, hogy dichotom változók esetén a Pearson–féle ϕ együtthatók (product moment coefficients) nem mérik jól az asszociációt, a maximumuk függ a peremeloszlásoktól, (a ϕ együttható nemcsak a változók kapcsolatának erősségétől függ, hanem a változók várható értékétől is, az 1 értéket csak akkor veszi fel, ha a két változónak azonos a várható értéke.) Azonban, ha a változók rendre pozitívan kapcsolódnak egymáshoz, és az eloszlásuk közel azonos irányba és mértékbe ferdül, a faktorelelmzés a ϕ együtthatók alkalmazása esetén lényegében nem ad torz eredményeket, a faktorstruktúra közel azonos lesz, bár a faktorsúlyok kisebbek lesznek, így a megbízhatóságról hamis képet kapunk. Ha a nemlineáris modellt alkalmazzuk, akkor a tetrakorikus korrelációt kell számítanunk a megfigyelt dichotom változók között, a faktormodellt ehhez kell illesztenünk.

A modell paramétereit $(\underline{\tau}\, \underline{\lambda}\, \underline{\phi})$ becsülhetjük a következő függvény minimalizálásával:

$$F = (\underline{s} - \underline{\sigma})' \underline{W}^{-1} (\underline{s} - \underline{\sigma})\,,$$

ahol

$\underline{s}$ első m eleme a küszöbértékeket, a további $m(m-1)/2$ elem a tetrakorikus korrelációkat tartalmazza,

$\underline{\sigma}$ elemei az $\underline{s}$ megfelelői a teljes populációban,

$\underline{W}$ az $\underline{s}$ mintabeli kovariancia mátrixának a becslése.

Ha $\underline{W} = \underline{I}$, akkor a klasszikus legkisebb négyzetek módszerével egyezik F minimalizálása, egyébként a reziduálisokat súlyozzuk a mintabeli ingadozással, így a paraméterek standard hibája kisebb lesz.

Az általánosított legkisebb négyzetek módszerének előnye továbbá, hogy az F minimuma χ négyzet eloszlású (MATHÉN, 1978). F minimalizálásánál nem okoz ugyan gondot, ha a tetrakorikus korrelációs mátrix nem pozitív definit, de ez jelzi, hogy a változók mögötti eloszlás nem normális eloszlású. Ilyenkor fel kell tételeznünk azt is, hogy a mögöttes látens faktorok eloszlása nem normális. MUTHÉN (1989) javasolta a többváltozós probit regresszió alkalmazását erre az esetre. MUTHÉN (1989) példát is ad mindkét modell alkalmazására.

3.8 Szimultán faktorelemzés a faktoriális invariancia vizsgálatára

Ha azonos változóhalmaz faktormodelljét becsüljük több mintában, felmerül a faktorok összehasonlításának problémája, a faktorok hasonlóságának mérése.

Az exploratív faktorelemzésnél vizsgáltuk a faktorok hasonlóságának mérésére kidolgozott mutatókat. Ezeket elsősorban az exploratív elemzés eredményeinek összehasonlításánál alkalmazhatjuk. A faktoriális invariancia problémája a konfirmatív faktorelemzés témaköréhez kötődik. JÖRESKOG (1971), LAWLEY és MAXWELL (1971), SÖRBOM és JÖRESKOG (1976) javasolta a szimultán faktorelemzés módszerét a faktoriális invariancia vizsgálatára.

Ha például két országban vizsgáljuk a mérési modellt, a következő hipotéziseket tesztelhetjük:

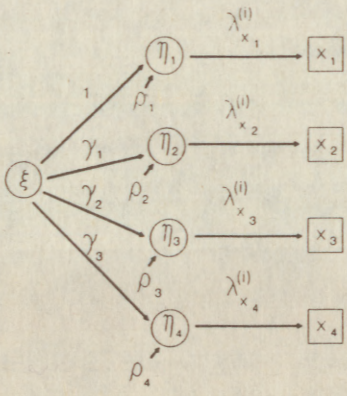
a) $\underline{\Lambda}^{(1)} = \underline{\Lambda}^{(2)}$ a két országban azonos a mérési modell,

b) $\underline{\Theta}^{(1)} = \underline{\Theta}^{(2)}$ a hiba komponensek varianciái egyenlőek,

c) $\underline{\Phi}^{(1)} = \underline{\Phi}^{(2)}$ a faktorok variancia–kovariancia mátrixai egyenlőek.

A két ország különbségét a congeneric mérési modellt feltételezve úgy írjuk le, hogy különbség van a két ország között variabilitásban, de egyébként a faktormodell invariáns a két országban. Ez azt jelenti, hogy a szórások különbözhetnek az országok között, de a korrelációs mátrix megegyezik. Másképpen fogalmazva, a kovarianciamátrix csak a megfigyelt változók skálafaktoraiban különbözik.

Ezt mutatja a következő ábra:



ahol

γ-k a faktorsúlyok,

λ-k a skála faktorok.

3.8.1 Szimultán faktorelemzés

Feltételezzük, hogy a különböző populációkban (vagy részpopulációkban) azonos m manifeszt változó faktormodelljét a következőképpen írhatjuk:

$$\underline{x}_g = \underline{\nu}_g + \underline{\Lambda}_g \underline{\xi}_g + \underline{\varepsilon}_g \,, \tag{48}$$

ahol $g = 1, 2, \ldots, s$,

$\underline{\nu}_g$ az m manifeszt változó mérési skáláinak középpontjait, nullpontjait tartalmazó oszlopvektor,

$\underline{\xi}_g$ a látens (hipotetikus) közös faktorok $(r \times 1)$ típusú vektora,

$\underline{\varepsilon}_g$ az egyedi faktorok (hiba komponensek) $(m \times 1)$ típusú vektora,

$\underline{\Lambda}_g$ a faktorsúlyok $(m \times r)$ típusú mátrixa.

$\underline{\nu}_g$ általánosan a változók mérési skálájának középpontja, kiindulópontja, nullpontja, a gyakorlatban legtöbbször az m változó várható értékeinek vektora (SÖRBOM és JÖRESKOG (1976) lokális paraméternek nevezi).

Általában azt szokták feltételezni, hogy $E(\underline{\xi}_g) = \underline{0}$, a következőkben azonban a faktorokat értelmezhető skálán mérjük.

A faktormodell feltételei a g-edik populációban:

1. $E(\underline{\xi}_g) = \underline{\theta}_g$ a faktorok várható érték–vektora
2. $E(\underline{\varepsilon}_g) = \underline{0}$
3. $E(\underline{\varepsilon}_g \underline{\varepsilon}'_g) = \underline{\Theta}_g$ diagonális
4. $E(\underline{\xi}_g \underline{\varepsilon}'_g) = \underline{0}$.

A 2–4. feltételek figyelembevételével $\underline{x}_g$ variancia–kovariancia mátrixát következőképpen írhatjuk:

$$\underline{\Sigma}_g = \underline{\Lambda}_g \underline{\Phi}_g \underline{\Lambda}'_g + \underline{\Theta}_g \,, \tag{49}$$

ahol

$E(\underline{\xi}_g \underline{\xi}'_g) = \underline{\Phi}_g$ a látens közös faktorok variancia–kovariancia mátrixa,

$E(\underline{\varepsilon}_g \underline{\varepsilon}'_g) = \underline{\Theta}_g$ az egyedi faktorok (hiba komponensek) variancia–kovariancia mátrixa (feltételezzük, hogy diagonális).

A modell paramétereinek becslése előtt a modell identifikálhatóságát kell vizsgálni ahhoz, hogy a paraméterek egyértelmű becslését megkaphassuk. Az identifikálhatóságnál a $\underline{\Lambda}_g, \underline{\Phi}_g, \underline{\Theta}_g$ paramétermátrixok és a $\underline{\Sigma}_g$ variancia–kovariancia mátrix egyértelmű megfeleltetését kell biztosítani. Elégséges feltétele az idenfitikálhatóságnak, hogy a modell egyenleteinek száma nagyobb vagy egyenlő legyen független paramétereinek számánál.

A becsléshez feltételezzük, hogy $\underline{x}_g$ többdimenziós normális elsozlású. A g-edik mintában a likelihood függvény logaritmusa

$$\log_e L_g = -\frac{1}{2}(n_g - 1)[\log_e |\underline{\Sigma}_g| + tr(\underline{S}_g \underline{\Sigma}_g^{-1})],$$

ahol

n_g a g-edik minta elemszáma,

$\underline{S}_g$ a mintabeli variancia–kovariancia mátrix.

Mivel a minták függetlenek, az összes mintára a likelihood függvény logaritmusa:

$$\log_e L = \sum_{g=1}^{s} \log_e L_g .$$

A gyakorlatban $\log_e L$ maximalizálása helyett az

$$F = \sum_{g=1}^{s} (n_g - 1) \left[tr(\underline{S}\underline{\Sigma}_g^{-1}) + \left(\log |\underline{\Sigma}_g| - \log |\underline{S}_g|\right) - m_g \right]$$

függvényt minimalizáljuk. A FLETCHER és POWELL (1963)–féle eljárással minimalizáljuk az F függvényt $\underline{\Lambda}_g, \underline{\Phi}_g$ és $\underline{\Theta}_g$ ismeretlen paraméterei szerint.

A modell illeszkedését khi négyzet statisztikával tesztelhetjük.

$$v = \min F(\underline{S}_g, \underline{\Sigma}_g(\underline{\Lambda}_g, \underline{\Phi}_g, \underline{\Theta}_g))$$

nagy minták esetén khi–négyzet eloszláshoz tart

$$d = \sum_{g=1}^{s} \frac{1}{2} m_g (m_g + 1) - t$$

szabadságfok mellett,

ahol

t a független paraméterek száma a modellben.

Az $\underline{x}_g$ megfigyelt változók kovarianciastruktúrája független a változók mérésétől, $\underline{\nu}_g$-től és a faktorok várható értékétől, $\underline{\theta}_g$-től.

A megfigyelt változók várható értéke a faktorok várható értékének lineáris függvénye:

$$E(\underline{x}_g) = \underline{\mu}_g = \underline{\nu}_g + \underline{\Lambda}_g \underline{\theta}_g . \tag{50}$$

Ha $\underline{\theta}_g = \underline{0}$ (a faktorok várható értéke nulla), akkor $\underline{\mu}_g = \underline{\nu}_g$, (a megfigyelt változók várható értéke $\underline{\nu}_g$-vel egyenlő). Ebben a modellben a faktorok mérése érdektelen. Ha $\underline{\nu}_g = 0$, akkor $\underline{\mu}_g = \underline{\Lambda}_g \underline{\theta}_g$, vagyis a megfigyelt változók várható értéke a faktorok várható értékének lineáris függvénye, de a lineáris függvényben nincs additív konstans. Ezt a modellt akkor használjuk, amikor a faktorok várható értékei fontosak, de a megfigyelt változók mérésének lokalizálása érdektelen. A továbbiakban feltételezzük, hogy $\underline{\nu}_g = \underline{0}$.

A közös faktormodell paramétereinek becsléséhez feltételezzük, hogy a modell identifikálható. Az identifikálhatóság szükséges feltétele, hogy a $\underline{\Lambda}_g, \underline{\Phi}_g$ és $\underline{\Theta}_g$ paraméterekre a modell legalább r^2 független feltételt tartalmazzon (lásd JÖRESKOG, 1971, SÖRBOM, 1974, SÖRBOM és JÖRESKOG, 1976).

Faktoriális invariancia

A faktoriális invariancia vizsgálatánál négy kérdést vizsgálunk. Az első kettő a mérési struktúra ($\underline{\Lambda}_g$ és $\underline{\Theta}_g$) invarianciáját vizsgálja, a másik kettő a faktortérben vizsgálja a modell ($\underline{\theta}_g$ és $\underline{\Phi}_g$) invarianciáját.

A faktoriális invariancia Thurstone–féle vizsgálata (THURSTONE, 1947) leszűkül a faktorsúlymátrix invarianciájának vizsgálatára (HARMAT, 1967), RUMMEL, 1970), a $\underline{\Lambda}_g$ mátrix azonosságának vizsgálatára két vagy több populációban. Az exploratív faktorelemzésnél a faktorok hasonlóságának mérőszámait a standardizált faktorokra alkalmaztuk. Ha $\underline{S} = (\text{diag}\underline{\Phi})^{-\frac{1}{2}}$ és $\underline{D} = (\text{diag}\underline{\Sigma})^{-\frac{1}{2}}$, a $\underline{\Lambda}$ faktorsúlymátrixot transzformálhatjuk, hogy elemei ± 1 között mozogjanak, ahogyan az exploratív faktorelemzésnél ezt feltételezzük:

$$\underline{\Lambda}^* = \underline{D}\,\underline{\Lambda}\,\underline{S}^{-1}. \tag{51}$$

A $\underline{\Lambda}^*$ „standarizált" faktorsúlymátrix függ mind a $\underline{\Lambda}$, mind a $\underline{\Phi}$ mátrixtól (és természetesen $\underline{D}$-től is, a változók varianciáitól is), így a $\underline{\Lambda}^*$ mátrixok hasonlóságának vizsgálata nem különíti el $\underline{\Lambda}$ és $\underline{\Phi}$ hatásait, összemossa $\underline{\Lambda}$ és $\underline{\Phi}$ invarianciájának vizsgálatát.

Először JÖRESKOG (1971) és SÖRBOM (1974) a szimultán faktorelemzés modelljében különítette el $\underline{\Lambda}_g, \underline{\Phi}_g$ és $\underline{\Theta}_g$ invarianciájának vizsgálatát, ezt egészítette ki ALWIN és JACKSON (1981) $\underline{\theta}_g$ invarianciájának vizsgálatával.

A következőkben a faktoriális invariancia hipotéziseit vizsgáljuk.

1.1 Hipotézis

Az első hipotézis, hogy a modell három paramétermátrixa, $\underline{\Lambda}_g, \underline{\Phi}_g$ és $\underline{\Theta}_g$ invariáns a különböző populációkban, amiből következik az a hipotézis, hogy a populációk megfigyelt kovarianciastruktúrái egyenlőek:

$$1.1 \qquad H_0 : \underline{\Sigma}_1 = \underline{\Sigma}_2 = \ldots = \underline{\Sigma}_s\,.$$

A hipotézis teszteléséhez a statisztikát JÖRESKOG (1971) dolgozta ki.

A statisztika közelítőleg χ^2 eloszlást követ $\frac{1}{2}(m-1)p(p+1)$ szabadságfok mellett $(\chi^2_{\underline{\Sigma}})$.

Ha a hipotézist nem vetjük el, akkor bizonyos hiba megengedésével azt tételezhetjük fel, hogy

$$\underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s, \qquad \underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_s \qquad \text{és}$$
$$\underline{\Theta}_1 = \underline{\Theta}_2 = \ldots = \underline{\Theta}_s$$

A továbbiakban ezeknek a mátrixoknak az invarianciáját teszteljük.

1.2 Hipotézis

Ha elfogadjuk azt a hipotézist, hogy a különböző populációk kovarianciastruktúrája egyenlő, vagyis egyenlőek a faktorsúlymátrixok, megvizsgálhatjuk a faktorok várható értékeinek egyenlőségét:

$$1.2 \qquad H_0 : \underline{\theta}_1 = \underline{\theta}_2 = \ldots = \underline{\theta}_s .$$

A hipotézist úgy teszteljük, hogy megvizsgáljuk a

$$H_0 : \underline{\Sigma}_1 = \underline{\Sigma}_2 = \ldots = \underline{\Sigma}_s \quad \text{és} \quad \underline{\mu}_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_s .$$

hipotézisnek megfelelő statisztikát, amely $\frac{1}{2}(s-1)m(m+1)+m(s-1)$ szabadságfokú χ^2 eloszlást követ, jelöljük $\chi^2_{\underline{\Sigma}\,\underline{\theta}}$-vel. Ezt a $\chi^2_{\underline{\Sigma}\,\underline{\theta}}$ értéket összehasonlítjuk a $\chi^2_{\underline{\Sigma}}$ értékkel (1.1 hipotézis), és így az 1.2 H_0 hipotézist a

$$\chi^2_{\underline{\theta}|\underline{\Sigma}} = \chi^2_{\underline{\Sigma}\,\underline{\theta}} - \chi^2_{\underline{\Sigma}} \qquad m(s-1) \text{ szabadságfokú}$$

χ^2 statisztikával teszteljük.

Ha $\chi^2_{\underline{\theta}|\underline{\Sigma}}$ szignifikáns, akkor a faktorok várható értékei különbözőek, ha a statisztika nem szignifikáns, akkor mivel $\underline{\mu}_g = \underline{\Lambda}_g \underline{\theta}_g$, azt mondhatjuk, hogy $\underline{\theta}_1 = \underline{\theta}_2 = \ldots = \underline{\theta}_s$.

2.1 Hipotézis

Ha az 1.1 hipotézist elvetjük, akkor külön–külön kell megvizsgálni a paramétermátrixok invarianciáját. Először a faktorsúlymátrix invarianciáját vizsgáljuk. A nullhipotézis

$$2.1 \qquad H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s .$$

Ez a hipotézis feltételezi, hogy a populációkban a közös faktorok száma megegyezik és a faktorok struktúrája is azonos.

Ha χ^2_r jelöli azt a statisztikát, amely a faktorok azonos számát teszteli, $\chi^2_{\underline{\Lambda}}$ pedig annak a modellnek az illeszkedését méri, amelynek paramétermátrixai $\underline{\Lambda}, \underline{\theta}_1, \underline{\theta}_2, \ldots, \underline{\theta}_s, \underline{\Phi}_1, \underline{\Phi}_2, \ldots, \underline{\Phi}_s, \underline{\Theta}_1, \underline{\Theta}_2, \ldots, \underline{\Theta}_s$, akkor a 2.1 H_0 hipotézist a

$$\chi^2_{\underline{\Lambda}|r} = \chi^2_{\underline{\Lambda}} - \chi^2_r$$

$\text{szf}_{\underline{\Lambda}|r} = \text{szf}_{\underline{\Lambda}} - \text{szf}_r$ szabadságfokú χ^2 eloszlású statisztika teszteli,

ahol

$$\text{szf}_r = s[\tfrac{1}{2}m(m+1) + m - mr + q - \tfrac{1}{2}r(r-1) - m - r],$$

q a $\underline{\Lambda}_g$-ben rögzített elemek száma,

$$\text{szf}_{\underline{\Lambda}} = \tfrac{1}{2}sm(m+1) - mr + q - \tfrac{1}{2}sr(r+1) - sm.$$

Ha a χ^2 érték szignifikánsan nagyobb a feltételeket tartalmazó modellben, mint a feltételeket nem tartalmazó modellben ($\chi^2_{\underline{\Lambda}|r} > 0$), akkor a faktorsúlymátrixok invarianciájának hipotézisét elvetjük.

2.2 Hipotézis

Ha a 2.1 hipotézist nem vetettük el, akkor megvizsgáljuk a faktorok várható értékei egyenlőségét. A hipotézis:

$$2.2 \qquad H_0 : \underline{\theta}_1 = \underline{\theta}_2 = \ldots = \underline{\theta}_s.$$

Először illesztjük azt a modellt, amelynek paramétermátrixai $\underline{\Lambda}, \underline{\theta}, \underline{\Phi}_1, \underline{\Phi}_2, \ldots,$ $\ldots, \underline{\Phi}_s, \underline{\Theta}_1, \underline{\Theta}_2, \ldots, \underline{\Theta}_s$. A modell illeszkedését a $\chi^2_{\underline{\Lambda}\,\underline{\theta}}$ statisztika méri. A $\chi^2_{\underline{\Lambda}\,\underline{\theta}}$ statisztikának és a 2.1 hipotézist tesztelő $\chi^2_{\underline{\Lambda}}$ statisztikának a

$$\chi^2_{\underline{\theta}|\underline{\Lambda}} = \chi^2_{\underline{\Lambda}\,\underline{\theta}} - \chi^2_{\underline{\Lambda}}$$

$$\text{szf}_{\underline{\theta}|\underline{\Lambda}} = \text{szf}_{\underline{\Lambda}\,\underline{\theta}} - \text{szf}_{\underline{\Lambda}}$$

különbsége teszteli a 2.2 hipotézist, ahol

$$\text{szf}_{\underline{\theta}|\underline{\Lambda}} = \tfrac{1}{2}sm(m+1) - mr - r + q - \tfrac{1}{2}sr(r+1),$$

q a $\underline{\Lambda}$ mátrix kötött elemeinek a száma.

2.3 Hipotézis

Ha a faktorstruktúra invariáns, akkor a faktorok várható értékei azonossága után (mellett) a hiba komponensek kovarianciastruktúrájának invarianciáját vizsgálhatjuk:

$$2.3 \qquad H_0 : \underline{\Theta}_1 = \underline{\Theta}_2 = \ldots \underline{\Theta}_s.$$

Ezt a hipotézist is feltételes hipotézisként értelmezzük. Illesztjük a $\underline{\Lambda}, \underline{\Theta}$, $\underline{\theta}_1, \underline{\theta}_2, \ldots, \ldots, \underline{\theta}_s$, $\underline{\Phi}_1, \underline{\Phi}_2, \ldots, \underline{\Phi}_s$ paraméterű modellt. Az illeszkedés jóságát a $\chi^2_{\underline{\Lambda}\,\underline{\Theta}}$ statisztika méri, $\mathrm{szf}_{\underline{\Lambda}\,\underline{\Theta}} = 1/2sm(m+1) + sm - mr + q - 1/2sr(r+1) - m - sr$.

A 2.3 H_0 hipotézist a

$$\chi^2_{\underline{\theta}|\underline{\Lambda}} = \chi^2_{\underline{\Lambda}\,\underline{\Theta}} - \chi^2_{\underline{\Lambda}}$$

$$\mathrm{szf}_{\underline{\Theta}|\underline{\Lambda}} = \mathrm{szf}_{\underline{\Lambda}\,\underline{\Theta}} - \mathrm{szf}_{\underline{\Lambda}} \quad \text{szabadságfokú}$$

χ^2 statisztikával teszteljük.

2.4 Hipotézis

Ha a faktorstruktúra invarianciájának hipotézisét (2.1) elfogadjuk, a harmadik paramétermátrix, amit tesztelnünk kell, a $\underline{\Phi}$. A nullhipotézis:

$$2.4 \qquad H_0 : \underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_s .$$

Az előzőekhez hasonlóan először a $\underline{\Lambda}, \underline{\Phi}$, $\underline{\theta}_1, \underline{\theta}_2, \ldots \underline{\theta}_s$, $\underline{\Theta}_1, \underline{\Theta}_2, \ldots, \underline{\Theta}_s$ modellt illesztjük, és a $\underline{\Phi}_g$-re vonatkozó invariancia hipotézist a következő statisztikával teszteljük:

$$\chi^2_{\underline{\Phi}|\underline{\Lambda}} = \chi^2_{\underline{\Lambda}\,\underline{\Phi}} - \chi^2_{\underline{\Lambda}}$$

$$\mathrm{szf}_{\underline{\Phi}|\underline{\Lambda}} = \mathrm{szf}_{\underline{\Lambda}\,\underline{\Phi}} - \mathrm{szf}_{\underline{\Lambda}} ,$$

ahol

$$\mathrm{szf}_{\underline{\Lambda}\,\underline{\Phi}} = \tfrac{1}{2}sm(m+1) - mr + m - \tfrac{1}{2}r(r+1) - sr.$$

Ha $\chi^2_{\underline{\Phi}|\underline{\Lambda}}$ szignifikáns, a faktorok kovarianciastruktúrájának invarianciájára vonatkozó hipotézist el kell vetnünk.

3.1 Hipotézis

JÖRESKOG (1971) javasolta a fenti hipotézisek mellett $\underline{\Phi}_g$ invarianciájának vizsgálatát a $\underline{\Lambda}_g$ és $\underline{\Theta}_g$ invariáns paraméterek feltételezésével.

Először a $\underline{\Lambda}, \underline{\Theta}$, $\underline{\Phi}$, $\underline{\theta}_1, \underline{\theta}_2, \ldots, \underline{\theta}_s$ modellt illesztjük, majd a $\underline{\Lambda}, \underline{\Theta}$, $\underline{\theta}_1, \underline{\theta}_2, \ldots, \underline{\theta}_s$, $\underline{\Phi}_1, \underline{\Phi}_2, \ldots, \underline{\Phi}_s$ modellt, és a két modell illeszkedésének különbségével teszteljük $\underline{\Phi}_g$ invarianciáját. A statisztika:

$$\chi^2_{\underline{\Phi}|\underline{\Lambda}\,\underline{\Theta}} = \chi^2_{\underline{\Lambda}\,\underline{\Theta}\,\underline{\Phi}} - \chi^2_{\underline{\Lambda}\,\underline{\Theta}}$$

$$\mathrm{szf}_{\underline{\Phi}|\underline{\Lambda}\,\underline{\Theta}} = \mathrm{szf}_{\underline{\Lambda}\,\underline{\Theta}\,\underline{\Phi}} - \mathrm{szf}_{\underline{\Lambda}\,\underline{\Theta}} ,$$

ahol

$$\mathrm{szf}_{\underline{\Lambda}\,\underline{\Theta}\,\underline{\Phi}} = \tfrac{1}{2}sm(m+1) + sm - mr + q - \tfrac{1}{2}r(r+1) - m - sr.$$

A fenti hipotéziseket áttekintve azt láthatjuk, hogy az előzetesen említett faktoriális invariancia vizsgálat két csoportba sorolt négy kérdése előtt két előzetes hipotézist is megvizsgáltunk, nevezetesen $\underline{\Sigma}_g$ és $\underline{\mu}_g$ invarianciájának hipotéziseit. A faktoriális invariancia tesztelése csak két előzetes hipotézis elfogadása után értelmes.

A mérési modell invarianciájának tesztelése a $\underline{\Lambda}_g$ és $\underline{\Theta}_g$ paramétermátrixok invarianciájának tesztelését jelenti. Először $\underline{\Lambda}_g$ invarianciáját teszteltük (2.1 hipotézis), majd $\underline{\Theta}_g$ invarianciáját azzal a feltételezéssel, hogy a 2.1 hipotézist nem vetettük el. A faktorok várható értékének ($\underline{\theta}_g$) invarianciáját fogalmazta meg a 2.2 hipotézis, míg a faktorkovariancia–struktúra invarianciáját ($\underline{\Phi}$) a 2.4 hipotézis. Ez utóbbi kettő már nem a mérési modellel, hanem a látens faktorok tulajdonságaival függ össze. JÖRESKOG (1971) javasolta, hogy $\underline{\Phi}_g$ invarianciáját mind $\underline{\Lambda}_g$, mind $\underline{\Theta}_g$ invarianciájának feltevésével teszteljük. Ezt fejezte ki a 3.1 hipotézis.

Összefoglalóan a faktoriális invariancia vizsgálatára a következő hipotéziseket fogalmaztuk meg.

Előzetes hipotézisek:

1.1 $H_0 : \underline{\Sigma}_1 = \underline{\Sigma}_2 = \ldots = \underline{\Sigma}_s$

1.2 $H_0 : \underline{\Sigma}_1 = \underline{\Sigma}_2 = \ldots = \underline{\Sigma}_s$ és $\underline{\mu}_1 = \underline{\mu}_2 = \ldots = \underline{\mu}_s$.

Specifikus invariancia hipotézisek:

2.1 $H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s$

2.2 $H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s$ és $\underline{\theta}_1 = \underline{\theta}_2 = \ldots = \underline{\theta}_s$

2.3 $H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s$ és $\underline{\Theta}_1 = \underline{\Theta}_2 = \ldots = \underline{\Theta}_s$

2.4 $H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s$ és $\underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_s$

3.1 $H_0 : \underline{\Lambda}_1 = \underline{\Lambda}_2 = \ldots = \underline{\Lambda}_s$, $\underline{\Theta}_1 = \underline{\Theta}_2 = \ldots = \underline{\Theta}_s$ és $\underline{\Phi}_1 = \underline{\Phi}_2 = \ldots = \underline{\Phi}_s$.

A 2.2, 2.3 és 2.4 hipotézisek feltételezik, hogy a 2.1 hipotézist nem utasítottuk el, a 3.1 hipotézisnél pedig azt tételeztük fel, hogy a 2.3 hipotézist nem utasítottuk el.

3.9 A faktorértékek becslése

A faktormodell $\underline{\Lambda}$, $\underline{\Phi}$, $\underline{\Theta}$ paramétereinek becslése és értelmezése után érdekelhet bennünket, hogy a közös faktorok látens értékei hogyan becsülhetők, a közös faktorokon az egyes megfigyelések hol helyezkednek el. Elméletileg ez a probléma a megfigyelési egységek $\underline{x}$ vektorainak leképezését jelenti a $\underline{\xi}$ faktortérbe. Ilyen értelemben nem is helyes a faktorértékek becsléséről beszélni, sokkal inkább a $\underline{\xi}$ valószínűségi változó $\underline{x}$ valószínűségi változóra vonatkozó feltételes *a posteriori* eloszlásán alapuló leképezésről. Először természetesen fel kell tételeznünk a faktorok *a priori* eloszlását. A közös faktorok eloszlásáról feltételezzük, hogy közelítőleg normális eloszlású, $\underline{0}$ a várható értékkel, $\underline{I}$ variancia–kovariancia mátrixszal (vagyis a faktorok korrelálatlanok és egységnyi varianciájúak):

$$\underline{\xi} \sim N_r(\underline{0}, \underline{I}) . \tag{52}$$

A faktormodell ($\underline{x} = \underline{\mu} + \underline{\Lambda}\,\underline{\xi} + \underline{\varepsilon}$) feltételezésével az $\underline{x}$ valószínűségi változó feltételes eloszlása normális eloszlást követ:

$$\underline{x}|\underline{\xi} \sim N_m(\underline{\mu} + \underline{\Lambda}\,\underline{\xi}, \underline{\Theta}) , \tag{53}$$

ahol $\underline{\Lambda}$ a faktorsúlyok $(m \times r)$ típusú mátrixa,

$\underline{\Theta}$ a hiba komponensek varianciáinak diagonális mátrixa,

$\underline{\mu}$ a megfigyelt változók várható értékeinek vektora.

A megfigyelt változók valószínűségeloszlása többdimenziós normális eloszlású:

$$\underline{x} \sim N_m(\underline{\mu}, \underline{\Lambda}\,\underline{\Lambda}' + \underline{\Theta}) , \tag{54}$$

feltételezve a közös faktorok korrelálatlanságát.

A fenti eloszlások alapján a közös faktorok $\underline{\xi}$ *a posteriori* valószínűségeloszlása szintén normális:

$$\underline{\xi}|\underline{x} \sim N_r\{\underline{\Lambda}'\Sigma^{-1}(\underline{x} - \underline{\mu}), (\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda} + \underline{I})^{-1}\} . \tag{55}$$

A közös faktorok feltételes várható értéke:

$$\begin{aligned} E(\underline{\xi}|\underline{x}) &= \underline{\Lambda}'\underline{\Sigma}^{-1}(\underline{x} - \underline{\mu}) \\ &= (\underline{I} + \underline{\Gamma})^{-1}\underline{\Lambda}'\underline{\Theta}^{-1}(\underline{x} - \underline{\mu}) , \end{aligned} \tag{56}$$

ami a következő azonosság felhasználásával könnyen igazolható:

$$\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Sigma} = \underline{\Lambda}'\underline{\Theta}^{-1}(\underline{\Lambda}\,\underline{\Lambda}' + \underline{\Theta}) = (\underline{I} - \underline{\Gamma})\underline{\Lambda}' .$$

ahol

$$\underline{\Gamma} = \underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda}.$$

A faktorértékeket tehát megkaphatjuk $\underline{\xi}$ *a posteriori* várható értéke alapján. Természetesen a paramétermátrixoknak csak a becsléseit ismerjük, ezért a faktorértékeknek is egy becslését kaphatjuk így meg.

A faktorértékek azonos becsléséhez jutunk a THOMSON (1951) által javasolt módszer alapján, amelyik a faktorértékek regressziós becslését keresi. Induljunk ki az $\underline{x} = \underline{\Lambda}\,\underline{\xi} + \underline{\varepsilon}$ faktormodellből. Most az egyszerűség kedvéért feltételezzük, hogy $E(\underline{x}) = \underline{\mu} = \underline{0}$ és hogy a közös faktorok korrelálatlanok. Ekkor

$$E(\underline{x}\,\underline{\xi}') = E((\underline{\Lambda}\,\underline{\xi} + \underline{\varepsilon})\underline{\xi}') = \underline{\Lambda}\,E(\underline{\xi}\,\underline{\xi}') = \underline{\Lambda},$$

vagyis korrelálatlan faktorok esetén a faktorsúlyok mátrixa ($\underline{\Lambda}$) a változók és a faktorok közötti korrelációs együtthatókat tartalmazza.

Továbbá a megfigyelt vektorok variancia–kovariancia mátrixa:

$$E(\underline{x}\,\underline{x}') = \underline{\Sigma} = \underline{\Lambda}\,\underline{\Lambda}' + \underline{\Theta}.$$

A faktorelemzés $\underline{x} = \underline{\Lambda}\,\underline{\xi} + \underline{\varepsilon}$ egyenletéből a faktorokat kifejezhetjük mint a megfigyelt változók lineáris függvényét:

$$(\underline{B}\,\underline{\Lambda})^{-1}\underline{B}\,\underline{x} = \underline{\xi} + (\underline{B}\,\underline{\Lambda})^{-1}\underline{B}\,\underline{\varepsilon}, \tag{57}$$

vagy

$$\underline{\widehat{\xi}} = \underline{A}\,\underline{x} = \underline{\xi} + \underline{u},$$

ahol

$$\underline{A} = (\underline{B}\,\underline{\Lambda})^{-1}\underline{B}.$$

Keressük tehát azt a $\underline{\widehat{\xi}}$-t, amelyik jó becslését adja $\underline{\xi}$-nek.

Tekintsük a k-adik faktor, ξ_k becslését:

$$\widehat{\xi}_k = \underline{a}'_k\underline{x} = \underline{x}'\underline{a}_k,$$

ahol

$\underline{a}'_k$ az $\underline{A}$ együtthatómátrix k-adik sorvektora, m–elemű.

Keressük azt az $\underline{a}_k$ vektort, amely mellett $(\widehat{\xi}_k - \xi_k)$ varianciája minimális.

Ezt a varianciát a következőképpen írhatjuk:

$$F_k = E(\widehat{\xi}_k - \xi_k)^2 = E(\underline{x}'\underline{a}_k - \xi_k)^2.$$

Az F_k függvényt deriváljuk $\underline{a}_k$ szerint:

$$\frac{\partial F_k}{\partial \underline{a}_k} = E[2\underline{x}(\underline{x}'\underline{a}_k - \xi_k)] = 2(\underline{\Sigma}\,\underline{a}_k - \underline{\lambda}_k),$$

ahol

$\underline{\lambda}_k$ a $\underline{\Lambda}$ mátrix k-adik oszlopvektora.

A parciális deriváltakat egyenlővé tesszük 0–val:

$$\underline{\Sigma}\,\underline{a}_k = \underline{\lambda}_k$$
$$\underline{a}_k = \underline{\Sigma}^{-1}\underline{\lambda}_k\,,$$

és így a k-adik faktor becslése

$$\widehat{\xi}_k = \underline{\lambda}_k'\underline{\Sigma}^{-1}\underline{x}\,.$$

Ha $\widehat{\underline{\xi}}$ jelöli a $\widehat{\xi}_1, \widehat{\xi}_2, \ldots, \widehat{\xi}_r$ faktorbecslések vektorát, általánosan a faktorértékek regressziós becslése:

$$\widehat{\underline{\xi}} = \underline{\Lambda}'\underline{\Sigma}^{-1}\underline{x}\,. \tag{58}$$

A $\underline{\Lambda}\,\underline{\Sigma}^{-1}$ együtthatómátrix — mint azt az (56) egyenletnél láttuk — egyenlő a $(\underline{I} - \underline{\Gamma})^{-1}\underline{\Lambda}'\,\underline{\Theta}^{-1}$-val, így az (58) egyenlet alternatív formája

$$\widehat{\underline{\xi}} = (\underline{I} + \underline{\Gamma})^{-1}\underline{\Lambda}'\,\underline{\Theta}^{-1}\underline{x}\,, \tag{59}$$

ahol

$$\underline{\Gamma} = \underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda}\,.$$

A becslések $E(\widehat{\underline{\xi}}\,\widetilde{\underline{\xi}}')$ kovarianciamátrixa és $E(\widehat{\underline{\xi}}\,\underline{\xi}')$ megegyezik:

$$\begin{aligned} E(\widehat{\underline{\xi}}\,\widetilde{\underline{\xi}}') = E(\widehat{\underline{\xi}}\,\underline{\xi}') &= \underline{\Lambda}'\underline{\Sigma}^{-1}\underline{\Lambda} \\ &= \underline{I} - (\underline{I} - \underline{\Gamma})^{-1}\,. \end{aligned} \tag{60}$$

A becslés hibája $(\widehat{\underline{\xi}} - \underline{\xi})$, és a hiba variancia mátrixa:

$$\begin{aligned} E[(\widehat{\underline{\xi}} - \underline{\xi})(\widehat{\underline{\xi}} - \underline{\xi})'] &= \underline{I} - \underline{\Lambda}'\underline{\Sigma}^{-1}\underline{\Lambda} \\ &= (\underline{I} + \underline{\Gamma})^{-1}\,. \end{aligned} \tag{61}$$

A fenti egyenletből látható, hogy a becslés akkor lesz jó, ha $(\underline{I} + \underline{\Gamma})^{-1}$ diagonális elemei kicsik. A $\underline{\Gamma}$ mátrix ($\underline{\Gamma} = \underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda}$) a faktormodell identifikációs feltételei szerint diagonális, és a diagonális elemei lehetőleg nagyok. Ez igaz az első elemekre, és így az első faktorok értékeinek becslései jók lesznek, de ahogyan csökken a faktorok varianciája, úgy romlik a faktorértékek becslésének jósága.

A $\widehat{\underline{\xi}}$ regressziós becslésnek az elméleti $\underline{\xi}$ értékekre vonatkozó feltételes várható értéke eltér az elméleti értékektől ($\underline{\xi}$-től), így ez a becslés torzított:

$$\begin{aligned} E(\widehat{\underline{\xi}}|\underline{\xi}) &= E(\underline{\Lambda}'\underline{\Sigma}^{-1}\underline{x}|\underline{\xi}) \\ &= (\underline{\Lambda}'\underline{\Sigma}^{-1}\underline{\Lambda})\xi \\ &= \underline{\xi} - (\underline{I} - \underline{\Gamma})^{-1}\underline{\xi}\,, \end{aligned} \tag{62}$$

ahol felhasználtuk, hogy

$$E(\underline{x}|\underline{\xi}) = \underline{\Lambda}\,\underline{\xi},$$

mivel a hibatagok korrelálatlanok a közös faktorokkal.

Ha feltételezzük, hogy a faktorok korrelálnak egymással, és kovarianciamátrixuk $\underline{\Phi}$, akkor a fenti azonosságok a következőképpen változnak:

$$\begin{aligned} E(\underline{\xi}\,\underline{\xi}') &= \underline{\Phi} \\ E(\underline{x}\,\underline{\xi}') &= E[(\underline{\Lambda}\,\underline{\xi} + \underline{e})\underline{\xi}'] = \underline{\Lambda}\,\underline{\Phi} \\ E(\underline{x}\,\underline{x}') &= \underline{\Sigma} = \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}. \end{aligned}$$

A faktorok becslései:

$$\hat{\underline{\xi}} = \underline{\Phi}\,\underline{\Lambda}\,\underline{\Sigma}^{-1}\underline{x}, \tag{63}$$

vagy

$$\hat{\underline{\xi}} = \underline{\Phi}(\underline{I} + \underline{\Gamma}\,\underline{\Phi})^{-1}\underline{\Lambda}'\underline{\Theta}^{-1}\underline{x}.$$

Továbbá

$$\begin{aligned} E(\hat{\underline{\xi}}\,\hat{\underline{\xi}}') = E(\hat{\underline{\xi}}\,\underline{\xi}') &= \underline{\Phi}\,\underline{\Lambda}'\underline{\Sigma}^{-1}\underline{\Lambda}\,\underline{\Phi} = \underline{\Phi} - \underline{\Phi}(\underline{I} + \underline{\Gamma}\,\underline{\Phi})^{-1} \\ E[(\hat{\underline{\xi}} - \underline{\xi})(\hat{\underline{\xi}} - \underline{\xi}'] &= \underline{\Phi}(\underline{I} + \underline{\Gamma}\,\underline{\Phi})^{-1}, \end{aligned} \tag{64}$$

és

$$\begin{aligned} E(\hat{\underline{\xi}}|\underline{\xi}) &= \underline{\Phi}(\underline{\Lambda}'\underline{\Sigma}^{-1}\underline{\Lambda})\underline{\xi} \\ &= \underline{\Phi}(\underline{I} + \underline{\Gamma}\,\underline{\Phi})^{-1}(\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\Lambda})\underline{\xi} \\ &= (\underline{I} + \underline{\Phi}\,\underline{\Gamma})^{-1}\underline{\Phi}\,\underline{\Gamma}\,\underline{\xi} \\ &= \underline{\xi} - (\underline{I} + \underline{\Phi}\,\underline{\Gamma})^{-1}\underline{\xi}. \end{aligned} \tag{66}$$

BARTLETT (1938) a faktorértékek becslésére egy másik eljárást javasolt, mely nem használja az eloszlásokra vonatkozó feltevést. Az eljárás a legkisebb négyzetek elvét alkalmazza, és a standardizált reziduálisok négyzetösszegét minimalizálja:

$$F = \sum_i^m (\varepsilon_i^2/\theta_i) = \underline{\varepsilon}'\underline{\Theta}^{-1}\underline{\varepsilon} = (\underline{x} - \underline{\Lambda}\,\underline{\xi})'\underline{\Theta}^{-1}(\underline{x} - \underline{\Lambda}\,\underline{\xi}). \tag{67}$$

Az F függvényt minimalizáljuk $\underline{\xi}$ szerint, az első derivált:

$$-2\underline{\Lambda}'\underline{\Theta}^{-1}(\underline{x} - \underline{\Lambda}\,\underline{\xi}) = 2(\underline{\Gamma}\,\underline{\xi} - \underline{\Lambda}'\underline{\Theta}^{-1}\underline{x}),$$

amelyet egyenlővé teszünk $\underline{0}$-val, és $\underline{\xi}$ egyenletet kielégítő becslést $\widehat{\underline{\xi}}^*$-vel jelölve

$$\underline{\Gamma}\widehat{\underline{\xi}}^* = \underline{\Lambda}'\underline{\Theta}^{-1}\underline{x},$$

amiből

$$\widehat{\underline{\xi}}^* = \underline{\Gamma}^{-1}\underline{\Lambda}'\underline{\Theta}^{-1}\underline{x}. \tag{68}$$

Ez a becslés nem függ attól, hogy a faktorok korrelálnak-e egymással vagy nem.

Ha a $\widehat{\underline{\xi}}^*$ becslést összehasonlítjuk a $\widehat{\underline{\xi}}$ becsléssel (először a faktorok korrelálatlanságát feltételezve, (68) és (59) egyenletek), akkor a különbség az, hogy $(\underline{I}+\underline{\Gamma})^{-1}$-t az (59) egyenletben kicseréltük $\underline{\Gamma}^{-1}$-nel, vagyis a kétfajta becslés csak egy skálafaktorban különbözik.

A faktorok korreláltságát feltételezve az összefüggés a két becslés között:

$$\widehat{\underline{\xi}}^* = (\underline{I}+\underline{\Gamma}^{-1}\underline{\Phi}^{-1})\widehat{\underline{\xi}}, \tag{69}$$

vagy

$$\widehat{\underline{\xi}} = \underline{\Phi}\,\underline{\Gamma}(\underline{I}+\underline{\Phi}\,\underline{\Gamma})^{-1}\widehat{\underline{\xi}}^*.$$

A $\widehat{\underline{\xi}}^*$ kovariancia mátrixai

$$E(\widehat{\underline{\xi}}^*\widehat{\underline{\xi}}^{*\prime}) = \underline{\Phi}+\underline{\Gamma}^{-1}$$
$$E(\widehat{\underline{\xi}}^*\underline{\xi}') = \underline{\Phi}$$
$$E[(\widehat{\underline{\xi}}^* - \underline{\xi})(\widehat{\underline{\xi}}^* - \underline{\xi})'] = \underline{\Gamma}^{-1}.$$

A $\widehat{\underline{\xi}}^*$ becslés torzítatlan becslése a $\underline{\xi}$ elméleti értéknek:

$$\begin{aligned} E(\widehat{\underline{\xi}}^*|\underline{\xi}) &= E(\underline{\Gamma}^{-1}\underline{\Lambda}'\underline{\Theta}^{-1}\underline{x}|\underline{\xi}) \\ &= \underline{\Gamma}^{-1}\underline{\Lambda}'\underline{\Theta}^{-1}\underline{\xi} = \underline{\xi}. \end{aligned} \tag{70}$$

Annak, hogy a $\widehat{\underline{\xi}} = \underline{A}'\underline{x}$ becslés torzítatlan legyen, szükséges feltétele, hogy

$$\underline{A}'\underline{\Lambda} = \underline{\Lambda}'\underline{A} = \underline{I}, \tag{71}$$

mivel

$$E(\widehat{\underline{\xi}}|\underline{\xi}) = \underline{A}'\underline{\Lambda}\,\underline{\xi}.$$

L

4.

—
ele
kap
a n
fak
egy
a v
kva
as
láte

tor
láte
kva
nün
mér

disı
láte
oszt
kate
meg
fiká

(vál
ciác

kate

szuł

tent
láte
a di
lata

meg
első

4. Látens struktúraelemzés

A szociológiai kutatások — elsősorban a survey adatokon végzett kutatások — nagyobbrészt kvalitatív (nominális vagy ordinális) változók közötti kapcsolatok elemzésére irányulnak. Gyakori az a feltételezés, hogy a változók közötti megfigyelt kapcsolatok magyarázhatók más kvalitatív változóval, az ún. teszt faktorral, vagyis a megfigyelt változók közötti kapcsolat megszűnik, ha kontrolláljuk őket a teszt faktorral. Ez azt jelenti, hogy a vizsgált manifeszt változók feltételesen függetlenek egymástól, vagyis a teszt faktor minden kategóriájához rendelt feltételes táblában a változók kölcsönösen függetlenek egymástól. Ha a teszt faktor is megfigyelt és kvalitatív, akkor a hagyományos loglineáris, logit, probit modelleket alkalmazhatjuk az elemzésnél. Ha a teszt faktor nem megfigyelt, látens kvalitatív változó, akkor a látens struktúraelemzés módszereit kell alkalmaznunk.

A látens struktúraelemzés hasonlít a faktorelemzéshez, azonban míg a faktorelemzés folytonos megfigyelt változókat magyaráz folytonos látens változókkal, a látens struktúraelemzésnél mind a manifeszt, mind a látens változók között lehetnek kvalitatív változók is. Ezenkívül a látens struktúraelemzésnél nem kell feltételeznünk sem a gyakorlatban legtöbbször nem teljesülő normalitási feltételt, sem a mérési skála folytonosságát.

A látens struktúraelemzés kitüntetett esete, amikor a megfigyelt változók diszkrét, kategorikus változók (sokszor dichotómok), és egy vagy több kategorikus látens változónk van (T lehetséges kategóriával). Ezt az esetet hívjuk *látens osztályelemzés*nek. A modell alapfeltételezése, hogy a látens változó bármelyik kategóriájában a megfigyelt változók függetlenek egymástól. A manifeszt változók megfigyelt kapcsolatait az adatoknak a látens változó kategóriája szerinti klasszifikációja eredményezi.

A látens osztályelemzés célja, hogy jellemezze azt a kategorikus látens változót (változókat), amely magyarázza a megfigyelt kategorikus változók közötti asszociációkat:

– hogy becsülje a látens változó relatív gyakoriságeloszlását,

– hogy becsülje a megfigyelt változók relatív gyakoriságait a látens változó kategóriáiban,

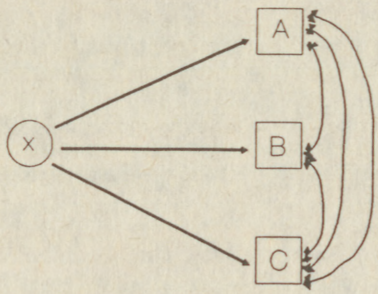
– hogy az előző gyakoriságeloszlások alapján következtessen a látens változó szubsztantív lényegére.

A látens struktúraelemzés speciális módszere a *látens tulajdonságelemzés* (latent trait analysis), amelyben a diszkrét megfigyelt változók asszociációit folytonos látens változókkal magyarázzuk. A harmadik módszer a *látens profilelemzés*, ahol a diszkrét látens változókkal magyarázzuk a folytonos megfigyelt változók kapcsolatait.

A látens struktúraelemzésről az első publikációk az 50-es években jelentek meg (lásd GREEN (1951), ANDERSON (1954) és GIBSON (1959) munkáit). Az első alapos és kimerítő kifejtését LAZARSFELD és HENRY (1968) végezte el. Az, hogy mégsem alkalmazták széles körben, ezt az egyébként elegáns módszert, a számítási és statisztikai nehézségeknek köszönhető. A modell becslésére alkalmazott eljárások (elsősorban az ún. determináns módszer (lásd MADANSKY 1960, LAZARSFELD és HENRY, 1968, LASY számítógépes program) okozták a nehézséget, mivel gyakran adtak lehetetlen eredményeket (pl. a valószínűségekre [0, 1] közé nem eső becslést, vagy egyszerűen azonos adathalmazra különböző eredményt. A hatékony, a maximum likelihood elv szerinti becslési eljárást GOODMAN (1974, 1975, 1976) és HABERMAN (1974, 1976, 1977, 1978) dolgozták ki. A paraméterek becslésére kidolgozott "iterative proportional scaling" eljárást alkalmazta CLOGG (1977) is az MLLSA (Maximum Likelihood Latent Structure Analysis) programban. Ez a program — helyes alkalmazás esetén — már elkerülte az előzőekben említett számítási problémákat. Az LCAG (Latent Class Models and Other Loglinear Models with Latent Variables) számítógépes program (JACQUES HAGENAARS munkája) is a GOODMAN (1974) által javasolt EM algoritmus alapján dolgozik és a paraméterekre maximum likelihood becslést ad.

4.1 A látens osztályelemzés

Tekintsünk egy háromdimenziós kereszttáblát, három megfigyelt változó, A, B, C együttes gyakoriságeloszlását. Jelentse π_{ijk} annak elméleti valószínűségét, hogy egy megfigyelés az (i, j, k) cellába esik $(i = 1, \ldots, I;\ j = 1, \ldots J;\ k = 1, \ldots, K)$. Tegyük fel, hogy a manifeszt változók kapcsolatát egyetlen látens változó (X) magyarázza, amelynek T kategóriája van és jelölje π_{ijkt}^{ABCX} a várható valószínűségét annak, hogy egy megfigyelés az (i, j, k, t) cellába esik. Amíg a π_{ijk} közvetlenül (direkt) megfigyelhető, addig a π_{ijkt}^{ABCX} közvetetten (indirekt) figyelhető csak meg az ABC táblából). Ezt az egyszerű esetet mutatja a következő ábra.



1. Ábra Három megfigyelt és egy látens változó kapcsolata

Ennek alapján a következő kifejezések írhatók fel:

$$\pi_{ijk} = \sum_{t}^{T} \pi_{ijkt}^{ABCX}, \tag{1}$$

és

$$\pi_{ijk} = \sum_{t}^{T} \pi_{ijkt}^{ABCX} = \sum_{t}^{T} \pi_t^X \pi_{ijk}^{\overline{ABC}X}, \tag{2}$$

ahol a

$\pi_{ijk}^{\overline{ABC}X}$ feltételes valószínűsége annak, hogy egy megfigyelés az (i, j, k) cellába esik, feltéve hogy a megfigyelés az X látens változó t-edik cellájában van,

π_t^X jelöli annak a valószínűségét, hogy a megfigyelés az X látens változó t-edik osztályába tartozik.

Az (1) egyenlet azt a pótlólagos feltételezést is tartalmazza, hogy a megfigyelések egyetlen látens változó osztályaiba teljes egészében szétoszthatók (ez az egyedek látens osztályba sorolhatósága).

Ha feltételezzük, hogy az X látens változó minden kategóriájában (osztályában) az A, B, C manifeszt változók kölcsönösen függetlenek egymástól, akkor a $\pi_{ijkt}^{\overline{ABC}X}$ a következőképpen írható:

$$\pi_{ijkt}^{\overline{ABC}X} = \pi_{it}^{\overline{A}X}\pi_{jt}^{\overline{B}X}\pi_{kt}^{\overline{C}X}, \tag{3}$$

ahol

$\pi_{it}^{\overline{A}X}$ annak a feltételes valószínűsége, hogy egy megfigyelés az A változó i-edik kategóriájában esik, feltéve hogy a megfigyelés az X látens változó t-edik osztályába tartozik.

A $\pi_{jt}^{\overline{B}X}$, $\pi_{kt}^{\overline{C}X}$ hasonlóan értelmezhető, mint a $\pi_{it}^{\overline{A}X}$.

A π_{ijkt}^{ABCX} a fentiek alapján:

$$\pi_{ijkt}^{ABCX} = \pi_t^X\pi_{it}^{\overline{A}X}\pi_{jt}^{\overline{B}X}\pi_{kt}^{\overline{C}X}. \tag{4}$$

Ennek alapján a (2) egyenletet a következőképpen írhatjuk:

$$\pi_{ijk} = \sum_t^T \pi_t^X\pi_{it}^{\overline{A}X}\pi_{jt}^{\overline{B}X}\pi_{kt}^{\overline{C}X}. \tag{5}$$

Az (5) egyenletet a látens osztályelemzés alapegyenletének nevezzük. Ez fejezi ki azt, hogy az A, B, C megfigyelt változók *feltételesen függetlenek* az adott látens változó kategóriáiban, vagyis az X látens változó magyarázza az A, B, C változók között megfigyelt asszociációt. Eszerint az A, B, C változót a látens X változó indikátorainak tekintjük.

A látens osztály modell paraméterei:

a) a látens osztály valószínűségei (π_t^X) a látens változó eloszlását írják le,

b) a feltételes valószínűségek ($\pi_{it}^{\overline{A}X}$, $\pi_{jt}^{\overline{B}X}$, $\pi_{kt}^{\overline{C}X}$), amelyek az egyes változók valószínűségeloszlásait írják le a látens változó egyes osztályaiban.

Ezek a feltételes valószínűségek hasonlatosak a faktorelemzés faktorsúlyaihoz. A feltételes valószínűségek segítségével jellemezhetjük a látens osztályokat. Egy adott látens osztályon belül a feltételes valószínűségek mutatják, hogy az adott látens osztály a megfigyelt változók milyen jellemzőit tükrözi, így mi az adott osztályok típusjegye.

Mivel a látens osztály modell paraméterei valószínűségek, nem lehetnek negatívak, és ki kell hogy elégítsék a következő feltételezéseket:

$$\sum_t^T \pi_t^X = \sum_i^I \pi_{it}^{\overline{A}X} = \sum_j^J \pi_{jt}^{\overline{B}X} = \sum_k^K \pi_{kt}^{\overline{C}X} = 1\,. \tag{6}$$

A (4) egyenlet alapján így

$$\pi_t^X = \sum_{i,j,k} \pi_{ijkt}^{ABCX}\,, \tag{7}$$

és

$$\pi_t^X \pi_{it}^{\overline{A}X} = \sum_{j,k} \pi_{ijkt}^{ABCX}\,. \tag{8}$$

Jelölje $\pi_{ijkt}^{ABC\overline{X}}$ annak a feltételes valószínűségét, hogy egy megfigyelés a t-edik látens osztályba esik, feltéve hogy az (i,j,k) cellában van.

A feltételes valószínűség definiciója szerint:

$$\pi_{ijkt}^{ABC\overline{X}} = \frac{\pi_{ijkt}^{ABCX}}{\pi_{ijk}}\,. \tag{9}$$

A (9) azonosságot felhasználva a (7) és (8) egyenleteket a következőképpen írhatjuk:

$$\pi_t^X = \sum_{i,j,k} \pi_{ijk}\pi_{ijkt}^{ABC\overline{X}}\,, \tag{10}$$

és a feltételes valószínűségek:

$$\pi_{it}^{\overline{A}X} = \frac{1}{\pi_t^X}\sum_{j,k} \pi_{ijk}\pi_{ijkt}^{ABC\overline{X}}\,, \tag{11}$$

$$\pi_{jt}^{\overline{B}X} = \frac{1}{\pi_t^X}\sum_{i,k} \pi_{ijk}\pi_{ijkt}^{ABC\overline{X}}\,, \tag{12}$$

$$\pi_{kt}^{\overline{C}X} = \frac{1}{\pi_t^X}\sum_{i,j} \pi_{ijk}\pi_{ijkt}^{ABC\overline{X}}\,. \tag{13}$$

A fenti öt egyenletet likelihood egyenletnek nevezzük és ezek segítségével becsüljük a paraméterek vektorát:

$$\underline{\pi} = \left[\pi_t^X,\, \pi_{it}^{\overline{A}X},\, \pi_{jt}^{\overline{B}X},\, \pi_{kt}^{\overline{C}X}\right].$$

A (9), (11)–(13) egyenleteknél az általánosság megszorítása nélkül feltételezhető, hogy a $\pi_t^X > 0$ és $\pi_{ijk} > 0$.

4.1.1 A paraméterek becslése

A látens osztályelemzésnél elsőként alkalmazott determináns becslési módszer (ANDERSON, 1954, LAZARSFELD és HENRY (1968)) nem biztosított mindig megfelelő megoldást, mint ezt már korábban említettük.

MCHUGH (1954) javasolt egy efficiens becslési eljárást, majd GOODMAN (1974, 1979) közölt egy általánosabb és egyszerűbb megoldást, amely maximum likelihood becslést adott a látens osztályvalószínűségre és a feltételes valószínűségre. Ezt az eljárást alkalmazták az MLLSA (Maximum Likelihood Latent Structure Analysis, CLOGG, 1977) számítógépes programban.

A becslést az "iteratív proportional scaling" eljárással végezzük. Jelölje a $\wedge$ szimbólum a maximum likelihood becslést.

Jelölje p_{ijk} a megfigyelt valószínűséget $(p_{ijk} = f_{ijk}/n)$, a megfigyelt relatív gyakoriságot az (i, j, k) cellában.

Megmutatható, hogy a maximum likelihood becslés kielégíti a következő egyenleteket (lásd pl. HABERMAN, 1974, 1979):

$$\hat{\pi}_t^X = \sum_{i,j,k} p_{ijk} \hat{\pi}_{ijkt}^{ABC\overline{X}} \tag{14}$$

$$\hat{\pi}_{it}^{\overline{A}X} = \frac{1}{\hat{\pi}_t^X} \sum_{j,k} p_{ijk} \hat{\pi}_{ijkt}^{ABC\overline{X}} \tag{15}$$

$$\hat{\pi}_{jt}^{\overline{B}X} = \frac{1}{\hat{\pi}_t^X} \sum_{i,k} p_{ijk} \hat{\pi}_{ijkt}^{ABC\overline{X}} \tag{16}$$

$$\hat{\pi}_{kt}^{\overline{C}X} = \frac{1}{\hat{\pi}_t^X} \sum_{i,j} p_{ijk} \hat{\pi}_{ijkt}^{ABC\overline{X}}. \tag{17}$$

Kiindulunk a paramétervektor egy kezdeti becsléséből, amely kielégíti a (6) egyenletet:

$$\hat{\underline{\pi}}(0) = \left[\hat{\pi}_t^X(0),\, \hat{\pi}_{it}^{\overline{A}X}(0),\, \hat{\pi}_{jt}^{\overline{B}X}(0),\, \hat{\pi}_{kt}^{\overline{C}X}(0)\right].$$

A kezdeti becslés és a (4) egyenlet alapján meghatározzuk a π_{ijkt}^{ABCX} kezdeti értékét $(\hat{\pi}_{ijkt}^{ABCX}(0))$.

Ezt behelyettesítve az (1) egyenletbe jutunk a π_{ijk} kezdeti becsléséhez $(\hat{\pi}_{ijk}(0))$.

Ezután a (9) egyenlet alapján kapjuk a $\pi_{ijkt}^{ABC\overline{X}}$ kezdeti becslését $(\hat{\pi}_{ijkt}^{ABC\overline{X}}(0))$.

Behelyettesítve ezeket a kezdeti becsléseket a (14)–(17) egyenletbe a paraméterek újabb becsléséhez jutunk $\hat{\underline{\pi}}(1)$. Miután kerekítési hibák miatt a (6) feltétel nem biztos hogy teljesül, a $\hat{\underline{\pi}}(1)$ paramétervektor elemeit a (6) a feltételnek megfelelően újraskálázzuk. Ezután az eljárást addig folytatjuk, amíg *vagy* el nem érünk egy megadott számú iterációig, *vagy* ha π_{ijkt}^{ABCX} ((4) egyenlet), egymást követő becslése közötti különbség valamilyen küszöbértéknél (tolerancia–szint) kisebb nem lesz.

4.1.2 A modell identifikálása

Láttuk, hogy a paramétervektor $\underline{\pi}$ (elemei a látens osztályok valószínűségei és a feltételes valószinűségek) a (4) és az (1) transzformációval meghatározza a $\{\pi_{ijk}\}$ valószínűségek vektorát:

$$\pi_{ijk} = f_{ijk}(\underline{\pi}), \tag{18}$$

ami azt fejezi ki, hogy a megfigyelt változók várható valószínűségei a modell paramétereinek a függvényei. Amennyiben a paramétertér bármely két $\underline{\pi}_1 \neq \underline{\pi}_2$ vektorához különböző várható valószínűségek tartoznak, vagyis a várható valószínűségeket egy és csak egy paramétervektor generálja, akkor a modellt identifikálhatónak nevezzük. A paraméterek identifikálhatóságának vizsgálatára MCTUGH (1956) korábbi eredményeinek általánosításával GOODMAN (1974) adott módszert. Az eljárás a transzformáció Jacobi mátrixának vizsgálatán alapul. Az (1) és (4) egyenletekkel definiált transzformáció akkor és csak akkor lesz nemszinguláris, ha az első deriváltak

$$\underline{H} = \left\{ \frac{\partial \{\pi_{ijk}\}}{\partial \underline{\pi}} \right\} \tag{19}$$

mátrixának a rangja egyenlő a paramétervektor nem redundáns elemeinek a számával (az oszlopok számával). A $\underline{H}$ elemei az (1) és (4) egyenletek alapján közvetlenül meghatározhatók. Pl. az π_t^X szerinti deriváltakat tartalmazó oszlop elemei

$$\frac{\partial \pi_{ijk}}{\partial \pi_t^X} = \pi_{it}^{\overline{A}X} \pi_{jt}^{\overline{B}X} \pi_{kt}^{\overline{C}X} - \pi_{iT}^{\overline{A}X} \pi_{jT}^{\overline{B}X} \pi_{kT}^{\overline{C}X}. \tag{20}$$

A minta becslései alapján becsülhetjük a parciális deriváltakat is, így juthatunk a $\hat{\underline{H}}$ becsléshez, majd a $\hat{\underline{H}}$ mátrix rangjának vizsgálatával határozhatjuk meg, hogy a modell identifikálható–e.

A megfigyelt gyakoriság–táblázat szabadságfoka

$$IJK - 1.$$

A becsült paraméterek száma

$$(T-1)+T(I-1)+T(J-1)+T(K-1)=(I+J+K-2)T-1\,.$$

Ha a paraméterek száma nagyobb, mint a várható valószínűségek (π_{ijk}) száma, akkor a modell nem identifikálható.

A modell szabadságfoka egyenlő $IJK-1$ minusz a $\underline{H}$ rangja, vagyis

$$\text{szf}=IJK-1-[(I+J+K-2)T-1]\,,$$

vagy

$$\text{szf}=IJK-1-[(I+J+K-M+1)T-1]\,, \tag{21}$$

ahol

M a megfigyelt változók száma.

Vagy másképpen:

szf = [a nem redundáns várható valószínűségek száma]

− [a nem redundáns paraméterek száma].

A modell identifikálhatóságának szükséges feltétele, hogy a modell szabadságfoka ne legyen negatív. Ha a modell nem identifikálható, akkor a paraméterekre tett pótlólagos feltételezésekkel azzá tehető. GOODMAN (1974) mutatott be eljárást arra, hogyan tehetünk egy nem–identifikálható modellt identifikálhatóvá.

4.1.3 A modell illeszkedése

A látens osztály modell illeszkedésének jóságát mérhetjük úgy, hogy a megfigyelt gyakoriságokat összehasonlítjuk a modell által becsült gyakoriságokkal a khi–négyzet statisztika segítségével.

A Pearson–féle χ^2 statisztika:

$$\chi^2=n\sum_{i,j,k}\frac{(p_{ijk}-\widehat{\pi}_{ijk})^2}{\widehat{\pi}_{ijk}}\,, \tag{22}$$

ahol

$p_{ijk}=f_{ijk}/n$ és f_{ijk} a megfigyelt gyakoriság az (ijk) cellában,

n a megfigyelések száma,

$\text{szf}=IJK-1-[(I+J+K-M+1)T-1]$ szabadságfokkal.

Egy másik, különösen nagy minták esetében előnyösebb mutató, a log–likelihood hányados (L^2):

$$L^2=2n\sum p_{ijk}\ln(p_{ijk}/\widehat{\pi}_{ijk})\,, \tag{23}$$

ami szintén aszimptotikusan χ^2 eloszlást követ a fenti szabadságfokkal.

A log–likelihood hányados statisztika előnye abban is megmutatkozik, hogy a látens változó adott számu (T) osztálya mellett a látens osztály valószínűségeire és a feltételes valószínűségekre vonatkozó hipotéziseket is tesztelhetjük vele, mivel az L^2 particionálható. Ez különösen fontos a konfirmatív modell illesztése esetében.

A két statisztikának azonos az eloszlása, ha a modell igaz és a minta elemszáma nagy, azonban különböznek kis minták, vagy az adatok nagyon egyenetlen cellák közötti eloszlása esetén. Kis minták esetén így előfordul, hogy a két statisztika különböző következtetésre vezet.

A modell illeszkedését vizsgálhatjuk az egyes cellákban külön is, a standardizált reziduálisok segítségével

$$e_{ijk}^{ABC} = \left(f_{ijk} - \widehat{F}_{ijk}\right) / \sqrt{\widehat{F}_{ijk}},$$

ahol

$\widehat{F}_{ijk} = n\widehat{\pi}_{ijk}$ a modell által becsült gyakoriság.

Az egymást követő modelleket összevethetjük az alapmodellel, (általában a kölcsönös függetlenséget feltételező H_0 modellel), és a változás mérésére a következő indexet használhatjuk:

$$R_i^2 = \frac{L^2(H_0) - L^2(H_i)}{L^2(H_0)},$$

ahol

H_i az i-edik hipotézist jelöli és $L^2(H_i)$ az ennek megfelelő log–likelihood hányados.

Az R_i^2 a modell javulását méri, értéke 0 és 1 közé esik.

A modell kiértékelésénél alkalmazhatjuk még az $F_i = L^2(H_i)/\text{szfok}$ mértéket, amely megfelel a regresszióelemzés F statisztikájának (HABERMAN (1978)).

Az F_i statisztikák alapján az R_i^2-hez hasonlóan mérhetjük a modell javulását is $R_\omega^2 = (F_0 - F_1)/F_0$. A kettő között az a különbség, hogy az R_ω^2-nál figyelembe vesszük a szabadságfokok változását is.

4.1.4 A látens osztály modell paramétereinek értelmezése

A látens osztályvalószínűségek

A látens osztályvalószínűségek (π_t^X) a minta látens osztályokba kerülésének arányait fejezik ki. Két alapvető kérdés merül fel: az egyik, hogy hány kategóriája, osztálya legyen a látens változónak, a másik, hogy a minta milyen arányban oszlik szét a látens osztályok között.

Az exploratív elemzés során, miután nincs elméleti feltevésünk az osztályok számára, a "lépésenkénti" eljárást alkalmazhatjuk. Kiindulhatunk a legegyszerűbb modellből, hogy ti. csak egy osztályunk van ($T = 1$), vagyis a változók kölcsönösen függetlenek. Ha ezt a modellt elfogadjuk, akkor nincs is szükség a látens változóra. Ha elvetjük, akkor először próbálkozhatunk a kétosztályos modellel, s ha ezt is el kell vetnünk, akkor a háromosztályossal, és így tovább, amíg elfogadhatóan illeszkedő modellt nem kapunk. A látens osztályok száma az illeszkedő modellben azt fejezi ki, hogy a minta elemei hány típusba — osztályba sorolhatók. Az osztályvalószínűségek pedig a típusok súlyát, arányát fejezik ki a mintán belül.

A látens feltételes valószínűségek

A feltételes valószínűségek $\pi_{it}^{\overline{A}X}$, $\pi_{jt}^{\overline{B}X}$, $\pi_{kt}^{\overline{C}X}$ a látens osztály (t) és a megfigyelt változók (A, B, C) (i, j, k) kategóriáinak kapcsolódásait fejezik ki, így hasonlítanak a faktorsúlyokhoz. A feltételes valószínűségek jelzik annak a valószínűségét, hogy egy megfigyelés egy adott látens osztályban a megfigyelt változók mely kategóriáiba esne. Ennek alapján következtetni tudunk a látens osztály jellegére, típusjegyeket rendelhetünk hozzá, és értelmezhetjük, megnevezhetjük őket.

A megfigyelt és a látens változók kapcsolatát jellemezhetjük az esélyhányados logaritmusával (GOODMAN, 1974). Az asszociáció X és A változók között:

$$\phi_{it}^{A} = \ln\left[\pi_{it}^{\overline{A}X}\pi_{i+1,t+1}^{\overline{A}X}/\pi_{i,t+1}^{\overline{A}X}\pi_{i+1,t}^{\overline{A}X}\right].$$

BECKER és CLOGG (1986) megmutatta, hogy az esélyhányados logaritmusa kapcsolatba hozható a tetrakorikus korrelációval.

A látens osztály tagságának becslése

A látens osztályelemzés egyik célja lehet, hogy a megfigyelt változók számát redukáljuk, és meghatározzuk az X látens változó becslését ($\widehat{X}$) a megfigyelt változók alapján.

A $\pi_{ijkt}^{ABC\overline{X}}$ fejezi ki azt a feltételes valószínűséget, hogy egy megfigyelés a látens változó t-edik osztályába esik, feltéve, hogy az A, B, C, D változók (i, j, k) cellájában van. Lásd a (9) képletet $\left(\pi_{ijkt}^{ABC\overline{X}} = \pi_{ijkt}^{ABCX}/\pi_{ijk}^{ABC}\right)$.

Így egy adott (i, j, k) cellába tartozó megfigyelésekre kiszámíthatjuk a $\pi_{ijkt}^{ABC\overline{X}}$ feltételes valószínűségeket, és ha a t' látens osztályhoz tartozik a legnagyobb valószínűség, akkor az (i, j, k) cellába tartozó megfigyeléseket a t'-edik látens osztályba soroljuk. Ezzel az osztályozási szabállyal a téves besorolások becsült valószínűsége $\varepsilon_{ijk} = 1 - \pi_{ijkt}^{ABC\overline{X}}$, és az adott ($i, j, k$) cellában a várható hibás osztályozás száma:

$$E_{ijk} = n\varepsilon_{ijk}\pi_{ijk}. \tag{24}$$

Az egész táblában a hibás osztályozások (klasszifikációk) várható száma:

$$E_2 = \sum_{ijk} E_{ijk} . \tag{25}$$

A hibás osztályozások becsült számát megkapjuk, ha a (24) és (25) képletekbe behelyettesítjük a megfelelő $\hat{e}_{ijk}$ és $\hat{\pi}_{ijk}$ becsléseket.

Tételezzük fel, hogy a megfigyelések látens osztályba sorolását úgy végezzük el, hogy eltekintünk attól, hogy az adott megfigyelés a manifeszt változók melyik kategóriáiba esik. Ha ismerjük a látens osztályok valószínűségeit (π_t^X), akkor minden megfigyelést X látens változó t^*-edik osztályába sorolunk, ha $\pi_{t^*}^X = max\pi_t^X$.

A feltétel nélküli osztálybasorolás várható hibája:

$$\varepsilon_1 = 1 - \pi_{t^*}^X ,$$

és a hibák várható száma

$$E_1 = n\varepsilon_1 . \tag{26}$$

A látens változó (X) és a megfigyelt változók együttese (ABC) közötti asszociációt mérhetjük a λ mutatóval (GOODMAN and KRUSKAL, 1954):

$$\lambda_{X \cdot ABC} = (E_1 - E_2)/E_1 . \tag{27}$$

Ez a mutató hasznosan kiegészíti a modell jóságát mérő korábban ismertetett statisztikákat, különösen az összehasonlító munkáknál.

4.1.5 A konfirmatív látens osztályelemzés

Az eddig tárgyalt modellben nem volt előzetes hipotézisünk a modell paramétereire, kivéve a valószínűségek axiómáiból adódó korlátozó feltételezést ((6) képlet), ugyanúgy nem volt előzetes ismeretünk a látens változó osztályainak számáról sem (és a látens változók számáról sem), így az exploratív elemzés logikája szerint jártunk el. Kiindultunk egy adott modellből (általában egy egyszerű, vagy a legegyszerűbb modellből), és a modell illeszkedésének vizsgálata után, lépések sorozatával juthattunk el az elfogadhatóan illeszkedő modellhez, amelynek a paramétereit értelmezhettük. A konfirmatív elemzésnél már előzetes feltételezésünk van a modell egy vagy több paraméterére, és a modellt ezekkel a korlátozásokkal akarjuk illeszteni a megfigyelt adatokhoz.

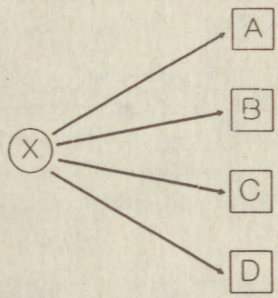
Általában a paraméterekre vonatkozó feltételezéseknek két fajtáját különböztetjük meg. Az egyikben azt írjuk elő, hogy a modell egy vagy több paramétere legyen egyenlő egymással, a másik feltételtípus, amikor *a priori* ismert egy vagy több paraméter értéke. A konfirmatív látens osztályelemzésnél a paraméterek lehetnek:

1) rögzítettek egy adott értékhez,

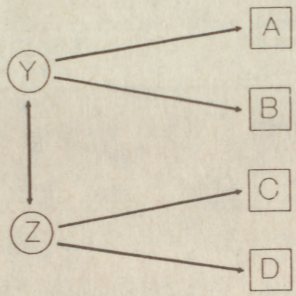
2) feltételes paraméterek, ismeretlenek, de egyenlőek egy vagy több paraméterrel,

3) szabad, ismeretlen paraméterek.

Tekintsük az A, B, C, D megfigyelt változókat és vizsgáljuk a következő, egy látens változós modellt:



Tegyük fel, hogy ez a modell nem illeszkedik elfogadhatóan az adatokhoz, így vizsgáljuk a következő ábrán látható modellt:



Ha Y és Z is dichotom látens változók, akkor a modellt kifejezhetjük úgy, mintha egy X látens változónk lenne négy kategóriával ($X \equiv YZ$).

Az X látens változó (1,2,3,4) osztályai (Y, Z) együttes látens változó (1,1), (1,2), (2,1) és (2,2) kategóriáinak felelnek meg.

A modell paraméterei:

$$\pi_t^X = \pi_{rs}^{YZ}, \pi_{it}^{\bar{A}X} = \pi_{irs}^{\bar{A}YZ}, \pi_{jt}^{\bar{B}X} = \pi_{jrs}^{\bar{B}YZ}, \ldots \text{stb}.$$

A modell feltételezi, hogy az A, B megfigyelt változók csak az Y, a C, D megfigyelt változók pedig csak a Z látens változótól függnek, valamint az A, B, C, D kölcsönösen függetlenek egymástól az adott Y, Z esetén, az Y és Z látens változók pedig kapcsolódnak egymáshoz. A következő pótlólagos feltételezések fejezik ki a fenti modellt:

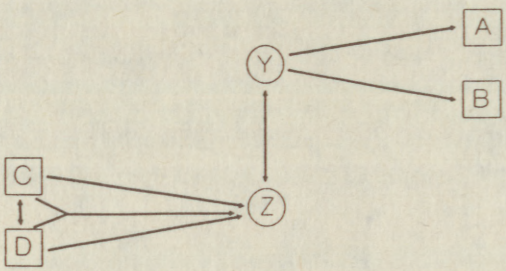
$$\pi_{i11}^{\bar{A}YZ} = \pi_{i12}^{\bar{A}YZ}, \quad \pi_{i21}^{\bar{A}YZ} = \pi_{i22}^{\bar{A}YZ}$$

$$\pi_{j11}^{\bar{B}YZ} = \pi_{j21}^{\bar{B}YZ} \qquad \text{stb.}$$

A klasszikus példa az ilyen modellre GOODMAN-tól (1974) származik: A és B dichotom változók a szavazás szándékát fejezik ki (1 = republikánus jelölt, 2 = más), C és D változók pedig a republikánus jelöltről kialakított véleményt tudakolta (1 = kedvező, 2 = nem kedvező).

A látens szavazás szándékát Y, a látens véleményt Z változók fejezik ki, és az utolsó ábrának megfelelő modell szerint ezek magyarázzák az $ABCD$ asszociációt. A maximum likelihood arány $L^2 = 7.32$ (szabadságfok = 4), ami jó illeszkedést mutat, bár az Y és Z közötti becsült kapcsolat negatív és kicsi, elhanyagolható volt (–0.08 logaritmikus skálán), és ez nem magyarázható.

CLOGG (1981) módosította az előző modellt a MIMC (Multiple–Indicator, Multiple–Cause) modellnek megfelelően:



A C és D változók nemcsak közvetlenül hatnak a látens változókra, hanem hat az interakciójuk is.

A CD együttes változó négy kategóriája (1,2,3,4) a (C, D) változók (1,1), (1,2), (2,1), (2,2) kategóriáival függ össze. A $\pi_{irs}^{\bar{A}YZ}$ az A változó feltételes valószínűségét jelöli, azt a valószínűséget, hogy egy megfigyelés az A változó i-edik kategóriájába esik, feltéve hogy az (Y, Z) változó (r, s) osztályába tartozik, és ehhez hasonlóan definiáljuk a többi feltételes valószínűségeket is.

A MIMC modell azt fejezi ki, hogy a $\pi_{irs}^{\bar{A}YZ}$ és a $\pi_{jrs}^{\bar{B}YZ}$ feltételes valószínűségek csak az Y látens változó r osztályától függnek, (és nem függnek a Z látens változó s kategóriájától), a $\pi_{n\ell rs}^{\overline{CD}YZ}$ feltételes valószínűségek pedig csak a Z változó s osztályától (és nem függnek az Y változó r kategóriájától).

A modell fenti korlátozó feltételezéseit a következő azonosságok fejezik ki:

$$\pi_{11}^{\bar{A}X} = \pi_{12}^{\bar{A}X}, \quad \pi_{13}^{\bar{A}X} = \pi_{14}^{\bar{A}X} \tag{28}$$

$$\pi_{11}^{\bar{B}X} = \pi_{12}^{\bar{B}X}, \quad \pi_{13}^{\bar{B}X} = \pi_{14}^{\bar{B}X} \tag{29}$$

$$\pi_{11}^{\overline{CD}X} = \pi_{13}^{\overline{CD}X}, \quad \pi_{21}^{\overline{CD}X} = \pi_{23}^{\overline{CD}X} \tag{30}$$

$$\pi_{31}^{\overline{CD}X} = \pi_{33}^{\overline{CD}X}, \quad \pi_{12}^{\overline{CD}X} = \pi_{14}^{\overline{CD}X} \tag{31}$$

$$\pi_{22}^{\overline{CD}X} = \pi_{24}^{\overline{CD}X}, \quad \pi_{32}^{\overline{CD}X} = \pi_{34}^{\overline{CD}X}. \tag{32}$$

A fenti (28)–(32) azonosságokkal definiálhatjuk a MIMC modellt a $(2 \times 2 \times 4)$-es $A \times B \times CD$ kereszttáblára. Az illeszkedés az említett példa esetén $L^2 = 7.04$ (szfok = 2), és a Z hatása Y-ra +.50, ami már közel van az *a priori* elvárásunkhoz.

Kvázi látens osztályok

Amikor kérdőíves vizsgálatoknál az emberek véleményét, attitűdjét tudakolják, gyakran előfordul, hogy egy–egy kérdéskörnél az emberek egy csoportja minden kérdésre vagy a kérdések egy részére azonosan válaszol, pl., mindegyikkel egyetért, vagy mindegyikkel nem ért egyet. Bármi is legyen ennek a válaszállandóságnak (response consistency) az oka, figyelembe kell venni a modellben extra látens osztály szerepeltetésével. Tegyük fel, hogy az A, B, C, D megfigyelt dichotom változó (1,1,1,1) és (2,2,2,2) cellája alábecsült az eredeti kétosztályos modellben. Bővítsük a modellt két kvázi osztállyal. Ha egy megfigyelés történetesen a harmadik látens osztályba esik, akkor az (1,1,1,1) cellába kell tartoznia, így $\pi_{13}^{\bar{A}X} = \pi_{13}^{\bar{B}X} = \pi_{13}^{\bar{C}X} = \pi_{13}^{\bar{D}X} = 1$ és természetesen $\pi_{23}^{\bar{A}X} = \pi_{23}^{\bar{B}X} = \pi_{23}^{\bar{C}X} = \pi_{23}^{\bar{D}X} = 0$.

Ehhez hasonlóan adhatjuk meg a másik kvázi osztályhoz tartozás feltételét is. Ilyen modell illesztésére ad példát HAGENAARS (1988).

4.1.6 Látens struktúralemzés ordinális változókra

A szociológiai kutatásokban a nominális mérés mellett az ordinális mérési skálát használjuk a leggyakrabban. Általánosan elterjedt gyakorlat, hogy a sorrendi értékeket mint numerikus mennyiségeket tekintjük és intervallum mérést feltételező statisztikai eljárásokban szerepeltetjük az egyébként ordinális mérési szintű változókat. Ehelyett helyesebb, ha a látens osztály modellt alkalmazzuk. A következőkben azt mutatjuk meg CLOGG (1979) példája segítségével, hogyan lehet a látens osztály modellt ordinális változókra alkalmazni.

Az elégedettség három indikátorát vizsgáljuk.

A. jelenti az elégedettséget a településsel (lakóhellyel),

B. jelenti az elégedettséget a hobbival (kevenc időtöltés),

C. jelenti az elégedettséget a családdal.

Az adatok az 1975–ös General Social Survey–ből (USA) származnak.

C. Clogg az eredeti 7 fokozatú elégedettség skálát átkódolta 3 kategóriába, ahol 1 jelentette a meglehetős (1=1,2), 2 a közepes (2=3,4), 3 az alacsony (3=5,6,7) elégedettségi szintet.

A megfigyelt háromdimenziós kereszttáblázatok:

1. Táblázat. Az elégedettség három trichotom indikátorának megfigyelt gyakorisága

B	A	C=1	C=2	C=3
1	1	466	27	16
1	2	191	38	14
1	3	64	18	5
2	1	126	31	5
2	2	117	58	12
2	3	45	23	3
3	1	54	12	7
3	2	49	26	11
3	3	23	16	15

$n = 1472$

2. Táblázat. Az illeszkedés jósága különböző látens osztálymodelleknél (az l. táblázat adatai alapján)

Modell	Szabadságfok	Khi négyzet (χ^2)	Likelihood arány L^2	$R^2 = \frac{1-L^2(H)}{L^2(H_0)}$
H_0	20	339.43	259.17	–
H_1	13	34.84	28.57	0.89
H_2	7*	2.32	2.36	0.99
H_3	14**	25.22	24.77	0.90
H_4	17	46.65	43.56	0.83

Megjegyzés:

* A modellben $\pi_{23}^{\overline{B}X} = 0$, de ezt feltételként kezelve a szfok 6 helyett 7

** A szabadságfok ténylegesen 12, de két becsült 0 paraméter miatt (azokat is feltételként kezelve) a szabadságfok 14.

Ahogyan általában az exploratív elemzésnél, most is először a legegyszerűbb modellel próbálkozunk. H_0 a kölcsönös függetlenség modellje (vagyis azt tételezzük fel, hogy csak egy látens osztályunk van). A H_1 hipotézisben azt feltételezzük, hogy két látens osztályunk van. Ez a modell elfogadhatóan illeszkedik: $L^2 = 28.57$ (szfok $= 13$). A H_3hipotézisben három látens osztályt tételezünk fel $L^2 = 2.36$ (szfok $= 7$). Ez a modell jól illeszkedik a megfigyelt adatokhoz, bár nem veszi figyelembe a változók rendezettségét. Ezért olyan feltételezéseket kell tennünk az illesztett modellre, hogy kielégítsük a következőket:

- az első látens osztály nem tartalmazhat alacsony elégedettségűeket (3. kategóriába esőket), de a közepes elégedettségűek, mint az 1. látens osztály hibája megengedett,
- a második látens osztályhoz tartozhatnak bármelyik elégedettségű kategóriákból,
- a harmadik látens osztályba nem tartozhatnak a meglehetősen elégedettek, a 2, és 3 kategóriák megengedettek, de a közepesen elégedettet válaszhibának tekintjük.

Ezzel a megkötöttséggel az X látens változó rendezett tulajdonságú lesz. Általánosabban azt mondhatjuk, hogy a t-edik látens osztályba a $(t-1, t, t+1)$ katógriákba tartozó egyedek kerülhetnek.

A H_3 hipotézis ezeket a korlátozásokat tartalmazza:

$$\pi_{31}^{\overline{A}X} = \pi_{31}^{\overline{B}X} = \pi_{31}^{\overline{C}X} = 0, \tag{33}$$

$$\pi_{13}^{\overline{A}X} = \pi_{13}^{\overline{B}X} = \pi_{13}^{\overline{C}X} = 0. \tag{34}$$

A likelihood–arány $h^2 = 24.77$ (szfok $= 14$) elfogadható illeszkedést mutat.

A paraméterek becsléseit a 3. táblázat tartalmazza.

3. Táblázat. A háromosztályos modellek (H_2, H_3) paraméterei

A látens osztályok valószínűségei

	H_2	H_3
1. Látens osztály	0.55	0.32
2. Látens osztály	0.41	0.66
3. Látens osztály	0.04	0.02

A feltételes valószínűségek

	1. látens osztály		2. látens osztály		3. látens osztály	
	H_2	H_3	H_2	H_3	H_2	H_3
$\pi_{1t}^{\bar{A}X}$	0.72	0.84	0.26	0.35	0.16	0.00*
$\pi_{2t}^{\bar{A}X}$	0.22	0.16	0.53	0.45	0.31	0.32
$\pi_{3t}^{\bar{A}X}$	0.06	0.00*	0.21	0.20	0.53	0.68
$\pi_{1t}^{\bar{B}X}$	0.79	0.92	0.32	0.42	0.15	0.00*
$\pi_{2t}^{\bar{B}X}$	0.14	0.08	0.50	0.39	0.00*	0.00**
$\pi_{3t}^{\bar{B}X}$	0.07	0.00*	0.18	0.19	0.85	1.00
$\pi_{1t}^{\bar{C}X}$	0.95	1.00	0.58	0.68	0.24	0.00*
$\pi_{2t}^{\bar{C}X}$	0.02	0.00**	0.36	0.25	0.28	0.39
$\pi_{3t}^{\bar{C}X}$	0.03	0.00*	0.06	0.07	0.48	0.61

* előre megadott, feltételezett érték

** becsült érték

Clogg javasolt még egy négyosztályos modellt is, amelyben az első három látens osztály csak az elégedettség azonos szintjeit tartalmazhatta (t-edik osztályba csak a (t, t, t) cella kerülhetett), a negyedik osztályba pedig a nem konzisztens megfigyelések kerültek. Ez a fajta feltételezés hasonlít HENRY (1975) relatív deprivációs modelljének feltételeihez.

4.1.7 Mobilitásvizsgálat látens osztályelemzéssel

A következőkben C. CLOGG (1981) vizsgálata alapján mutatjuk be a látens osztály modellek alkalmazását mobilitás táblákra.

A hipotézis az, hogy a megfigyelt foglalkozási kategóriák közötti mobilitási folyamatot a mögöttes, különböző, látens osztályok befolyásolják. A látens osztály fogalom nem azonos a weberi osztályfogalommal bár elég nagy a hasonlóság.

A weberi koncepció szerint a társadalmi osztályok az egyének olyan együttesei, amelyeknek közös a mobilitási esélyük, a látens osztály modellben a látens osztályok jellemzője, hogy az osztályokon belül statisztikusan független az eredeti és az elért (az apa és a fiú) státusza.

Tekintsük a $I \times I$ típusú mobilitás táblát, ahol a sorokban az apa foglalkozása (F), az oszlopokban pedig a fiú foglalkozása található (S).

A várható valószínűség az (i, j) cellában $(i = 1, \ldots, I,\ j = 1, \ldots, I)$ π_{ij}.

Az apa foglalkozásának (eredeti státusz) gyakoriságeloszlása (peremeloszlás) $\pi_i^F = \sum_j \pi_{ij}$, és a mobilitási ráta $r_{ij} = \pi_{ij}/\pi_i^F$.

Az eredeti státusz (apa foglalkozása) természetes módon megelőzi az elért, adott státuszt. Feltételezhetjük, hogy a kettő között létezik egy közvetítő, közbülső látens változó (X), amely magyarázza az F és S asszociációját.

Ha X megfigyelhető és diszkrét, akkor a loglineáris modellt illeszthetjük a mobilitás táblára($F \times S|X$). Ha az F, S és X változók intervallum mérési szintű manifeszt változók, akkor a lineáris regresszió módszerét alkalmazhatjuk, és a modell feltétele szerint a parciális regressziós együttható $\beta_{FS \cdot X} = 0$. Ha F, S és X is diszkrét és X látens változó, akkor a látens osztály modellt illeszthetjük a mobilitás táblára:

$$\pi_{ij} = \sum_{t=1}^{T} \pi_t^X \pi_{it}^{\overline{F}X} \pi_{jt}^{\overline{S}X} . \tag{35}$$

A $\pi_{it}^{F\overline{X}}$ jelöli annak a valószínűségét, hogy egy megfigyelés az i-edik eredeti státuszból a t-edik látens osztályba kerül.

A $\pi_{it}^{F\overline{X}}$ feltételes valószínűség a következőképpen írható fel:

$$\pi_{it}^{F\overline{X}} = \pi_{it}^{FX}/\pi_i^F = \pi_{it}^{\overline{F}X} \pi_t^X/\pi_i^F . \tag{36}$$

A mobilitási ráta a feltételes valószínűségek függvényében:

$$r_{ij} = \sum_{i=1}^{T} \pi_{it}^{F\overline{X}} \pi_{jt}^{\overline{S}X} . \tag{37}$$

A (37) egyenlet azt fejezi ki, hogy az i–státuszból a j–státuszba kerülésnek a valószínűsége egyenlő az i státuszból a t–látens osztályba és a t látens osztályból a j–státuszba kerülés valószínűségeinek szorzatösszegével (ahol feltételezzük, hogy a $(i \to t)$ és $(t \to j)$ függetlenek minden t–re $(t = 1, \ldots, T)$.

Tekintsünk most néhány speciális esetet.

A H_0 hipotézis szerint F és S függetlenek egymástól, $T = 1$ és így $\pi_{ij} = \pi_i^F \pi_j^S$. Ez a modell írja le a tökéletes vagy teljes mobilitást. A mobilitás ebben a modellben véletlenszerű, csak az eredeti és a tényleges peremeloszlástól függ.

Ha $T = 2$, akkor a két látens osztályon belül véletlenszerű a mobilitás, de a két osztály között létezik „osztálykorlát" (a mobilitás a $\pi_{it}^{\overline{F}X}$ és $\pi_{jt}^{\overline{S}X}$ feltételes valószínűségektől függ). Tekintsük GOODMAN (1969, 1972) kvázi–teljes mobilitási modelljét, amelyben $T = I + 1$ és ahol $\pi_{ii}^{\overline{F}X} = \pi_{ii}^{\overline{S}X} = 1$ $(i = 1, \ldots, I)$.

A (35) egyenlet ekkor a következő:

$$\pi_{ij} = \begin{cases} \pi_i^X + \pi_{I+1}^X \pi_{i,I+1}^{\overline{F}X} \pi_{j,I+1}^{\overline{S}X}, & \text{ha } i = j \\ \pi_{I+1}^X \pi_{i,I+1}^{\overline{F}X} \pi_{j,I+1}^{\overline{S}X}, & \text{ha } i \neq j. \end{cases} \tag{38}$$

A látens változó első I osztálya determinisztikus, mivel a feltétel szerint $\pi_{ii}^{\overline{F}X} = \pi_{ii}^{\overline{S}X} = 1$ és $\pi_{it}^{\overline{F}X} = \pi_{jt}^{\overline{S}X} = 0$ $i \neq j$ esetén.

Vezessük be a következő azonosságokat:

$$\alpha_i = \pi_{I+1}^X \pi_{i,I+1}^{\overline{F}X}$$

$$\beta_j = \pi_{j,I+1}^{\overline{S}X},$$

és

$$\gamma_i = 1 + \pi_i^X / \alpha_i \beta_i \quad i = 1, \ldots, I.$$

Ekkor a (38) egyenletet a következőképpen fejezhetjük ki:

$$\pi_{ij} = \begin{cases} \alpha_i \beta_i \gamma_i, & \text{ha } i = j \\ \alpha_i \beta_i, & \text{ha } i \neq j. \end{cases} \tag{39}$$

GOODMAN (1972) a γ_i-t immobilitási indexnek nevezte. Ez a kvázi–teljes mobilitás modellje, amelyben azt tételezzük fel, hogy mindegyik státuszkategóriában vannak olyanok, akik az adott státuszban maradnak (az immobilitás determinisztikus ezekben a látens osztályokban), és van egy olyan látens osztály, ahol a mobilitás véletlenszerű, az apa és a fiú peremeloszlástól függ csak a feltételes táblában. A kvázi–teljes mobilitást GOODMAN (1974) módosította úgy, hogy az I determinisztikus látens státuszt (ahol a státusz–örökség 1 valószínűségű esemény) két látens osztállyal bővítette a látens mobilok számára. Ezt a modellt Goodman kvázi látens struktúrának nevezte, amelyben

$$\pi_{ij} = \begin{cases} \pi_i^X + \sum_{t=I+1}^{I+2} \pi_t^X \pi_{it}^{\overline{F}X} \pi_{jt}^{\overline{S}X}, & \text{ha } i = j \\ \sum_{t=I+1}^{I+2} \pi_t^X \pi_{it}^{\overline{F}X} \pi_{jt}^{\overline{S}X}, & \text{ha } i \neq j. \end{cases} \tag{40}$$

Vezessük be a kvázi mobilitás modelljénél alkalmazott jelöléseket:

$$\alpha_{i1} = \pi^X_{I+1}\pi^{\overline{F}X}_{i,I+1}$$
$$\alpha_{i2} = \pi^X_{I+2}\pi^{\overline{F}X}_{i,I+2}$$
$$\beta_{j1} = \pi^{\overline{S}X}_{j,I+1}$$
$$\beta_{j2} = \pi^{\overline{S}X}_{j,I+2}.$$

Ezek alapján:

$$\pi_{ij} = \begin{cases} \pi^X_i + \sum_{k=1}^2 \alpha_{ik}\beta_{jk}\,, & \text{ha } i = j \\ \sum_{k=1}^2 \alpha_{ik}\beta_{jk}\,, & \text{ha } i \neq j\,. \end{cases} \tag{41}$$

Legyen $s_{ij} = \Sigma_k \alpha_{ik}\beta_{jk}$,

és $\gamma^*_i = 1 + \pi^X_i / s_{ij}$.

A fentiek alapján

$$\pi_{ij} = \begin{cases} s_{ij}\gamma^*_i\,, & \text{ha } i = j \\ s_{ij} & \text{ha } i \neq j\,. \end{cases} \tag{42}$$

A γ^*_i az i-edik státusz státusz–örökítésének a mértéke ($\gamma^*_i \leq 1$), az i-edik státuszban maradók azon mértéke, amely nem magyarázható a modell által várható immobilitással (két látens mobil osztályt feltételezve).

A fenti modelleket C. CLOGG (1981) illesztette angol és dán intergenerációs foglalkozási mobilitás táblázatokra. A dán adathalmaz forrása, SVALASTOGA (1959), az angol tábláé: GLASS (1954). Mindkét táblázat 5×5-ös, ahol a foglalkozási kategóriák: (1) vezető értelmiség, (2) magasan kvalifikált irányító, (3) adminisztratív, irányító, (4) szakmunkás, adminisztrátor, (5) betanított munkás.

4. Táblázat. Intergenerációs mobilitási táblázat

Dánia (n=2391)

Apa státusza*	Megkérdezett státusza 1	2	3	4	5
1	18	17	16	4	2
2	24	105	109	59	21
3	23	84	289	217	95
4	8	49	175	348	198
5	6	8	69	201	246

Anglia

	1	2	3	4	5
1	50	45	8	18	8
2	28	174	84	154	55
3	11	78	110	223	96
4	14	150	185	741	447
5	0	42	72	320	411

* ahol a kategóriák a következő nyolc foglalkozási kategória összevonásával keletkeztek: (1) szellemi irányító, (2) vezető, (gazdasági vezetők, döntéshozók), (3) magasan kvalifikált irányító, (4) irányító, (5) adminisztrátor, (6) szakmunkás, (7) betanított–munkás, (8) segédmunkás. Az összevonás (1), (2,3), (4), (5,6), (7,8).

A 4. táblázatra először a H_0, kölcsönös függetlenség modelljét illesztjük, majd egy két látens osztályos modellt (H_1), és egy három látens osztályos modellt (H_2), A H_3 hipotézis a kvázi teljes mobilitás hipotézisét fogalmazza meg.

A H_4 hipotézis a kvázi–teljes mobilitást fogalmazza meg arra az esetre, amikor az 1., 3. és 5. foglalkozási kategóriákra írjuk elő a státusz–örökséget (a 2. és 4. kátegóriákra nem), és két olyan látens osztályunk van, amelyekben a függetlenség, teljes mobilitás érvényes. A H_4 modell feltételei:

$$\pi_{11}^{\overline{F}X} = \pi_{11}^{\overline{S}X} = 1$$
$$\pi_{32}^{\overline{F}X} = \pi_{32}^{\overline{S}X} = 1$$
$$\pi_{53}^{\overline{F}X} = \pi_{53}^{\overline{S}X} = 1,$$

és a 4. és 5. látens osztályra nincs semmilyen korlátozó feltételünk.

A H_5 hipotézis a H_4 feltételeit kiegészíti a következőkkel:

$$\pi_{it}^{\overline{F}X} = \pi_{it}^{\overline{S}X} \quad \text{minden} \quad i = 1, \ldots, 5 \quad \text{és} \quad t = 4, 5\text{–re}.$$

Ezek a feltételek azt írják elő, hogy a 4. és 5. látens osztályban az apa és a fiú foglalkozási peremeloszlásai azonosak, vagyis mindkét látens osztályban a foglalkozási kategóriák eloszlásai az idővel nem változtak.

5. Táblázat. Látens osztály és kvázi látens struktúra modellek illesztése a mobilitási táblára

	Anglia			Dánia		
Modell	szfok	χ^2	L^2	szfok	χ^2	L^2
H_0 (függetlenség)	16	1225.59	821.89	16	754.10	654.20
H_1 (két látens osztály)	9	202.14	177.97	8	133.64	116.23
H_2 (három látens osztály)	2	50.55	50.85	2	33.28	32.67
H_3 (kvázi–teljes mobilitás $T = I + 1$)	11	327.3	249.50	11	270.30	248.70
H_3 (kétosztályos látens struktúra $T = I + 2$)	6	9.35	9.76	6	8.06	8.23
H_4 (korlátozott kétosztályos látens struktúra $T = I + 2$)	13	41.44	41.18	14	32.47	32.87

Ha az előző hipotéziseknek megfelelő modellek illeszkedését vizsgáljuk, az 5. táblázatból azt láthatjuk, hogy a teljes mobilitás modellje nem illeszkedik az adatokhoz.

A H_1 modell illeszkedése is rossz, a három látens osztályos modell viszont már elég jól illeszkedik.

BERGER (1982) az eredeti 8×8-as mátrixot vizsgálva arra az eredményre jutott, hogy a mobilitási tábla mögött egy háromosztályos társadalmi struktúra húzódik meg, mégpedig létezik egy felső osztály ehhez tartoznak az (1,2,3) kategóriák, egy középosztály (4,5) kategóriákkal, és egy alsóosztály (6,7,8) kategóriákkal. Láttuk, hogy az összevont 5×5–ös táblázat, amelyre a bemutatott modelleket C. Clogg illesztette, a foglalkozási kategóriákat ettől eltérően vonta össze. Ez a két modell nem tartalmazott külön feltételeket a paraméterekre, így minden látens osztályon belül a teljes mobilitás hipotézise érvényesül.

A H_3 modellnél a kvázi–látens hipotézisek alapján feltételeztük, hogy az 1., 3., és 5. státuszokban létezik a státuszátörökítés, és két olyan látens osztályt definiáltunk, amelyekben független az apa és a fiú foglalkozási státusza. Ez a modell, ahogy az 5. táblázatban látható, mindkét országban nagyon jól illeszkedik a mobilitási táblához.

A H_3 modell módosítja a H_4-et úgy, hogy a negyedik és ötödik látens osztályban az apa és a fiú foglalkozáselosztása azonos marad.

6. Táblázat. A H_5 hipotézis paramétereinek becslése

Látens osztályok valószínűségei	$\hat{\pi}_1^X$	$\hat{\pi}_2^X$	$\hat{\pi}_3^X$	$\hat{\pi}_4^X$	$\hat{\pi}_5^X$
Anglia	0.012	0.013	0.061	0.212	0.701
Dánia	0.006	0.058	0.059	0.335	0.541

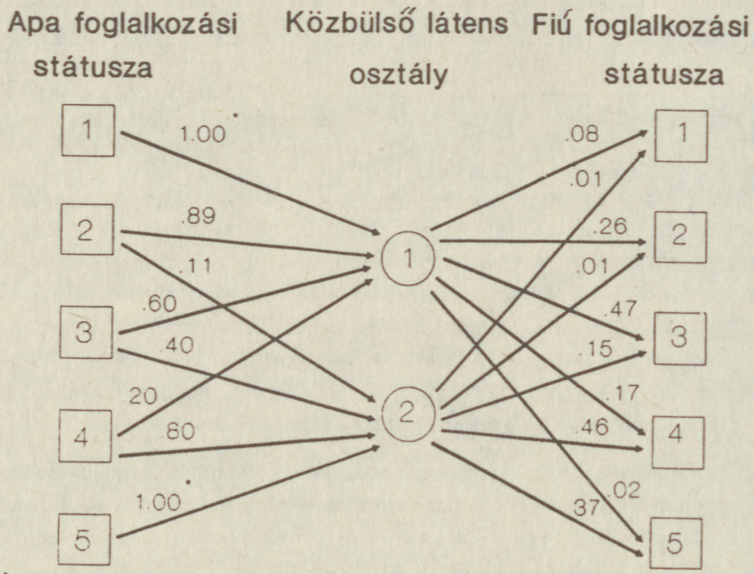
Feltételes valószínűségek a 4. látens osztályban

	1	2	3	4	5
Anglia	0.095	0.479	0.173	0.226	
Dánia	0.067	0.363	0.340	0.179	

Feltételes valószínűségek az 5. látens osztályban

	1	2	3	4	5
Anglia	0.0000	0.056	0.128	0.532	0.285
Dánia	0.0000	0.000	0.209	0.510	0.281

A (37) egyenletben a mobilitási ráta a feltételes valószínűségek szorzatainak összege. A kétosztályos modell dániai eredményeit láthatjuk a következő ábrán:



Megjegyzés:

A ⋆ a modell identifikálhatósága érdekében tett feltételezés (*a priori* érték).

4.2 Látens tulajdonságelemzés (Latent trait analysis)

A látens tulajdonság modellben (latent trait model) a kvalitatív (diszkrét értékekkel rendelkező) megfigyelt változók asszociációit kvantitatív, folytonos látens változóval magyarázzuk. A látens változót látens tulajdonságnak, látens jellemzőnek, látens trait–nek nevezzük. A trait a pszichológiában használatos kifejezés, olyan általános tulajdonságot, jellemző vonást jelent, amelynek segítségével a személyiségeket meg tudjuk különböztetni. Statisztikai értelemben a látens trait megfelel a közös faktor terminológiának.

A látens változónak (X) $T = n$ „osztálya" van (ahol n a minta elemszáma). Jelölje μ_i a látens változó értékét a t-edik megfigyelés esetén $(i = 1, \dots, n)$.

Tételezzük fel, hogy a megfigyelt változók dichotom változók, jelölje őket Y_j ($Y_j = 1$ vagy $Y_j = 0$). Jelölje Y_{ij} az i-edik megfigyelés válaszát a j-edik változóra, y_{ij} pedig a megfigyelt értéket.

Jelölje a feltételes valószínűségeket $\pi_{j|i} = P(Y_{ij} = 1|X = \mu_i)$.

Attól függően, hogy a megfigyelt és a látens változók közötti kapcsolatot milyen függvénytípussal írjuk le, különbözőképpen fejezzük ki a feltételes valószinűségeket. A logisztikus függvényt feltételezve :

$$\pi_{j|i} = P(Y_{ij} = 1 \mid X = \mu_i) = \frac{1}{1 + \exp(-z_{ij})},$$

ahol

$$z_{ij} = \alpha_j \mu_i + c_j,$$

ahol

c_j a j-edik változó „nehézségi fokát" jelzi (azt, hogy milyen arányban választották a mintából),

α_j a faktorsúlyokhoz hasonló értelmű, a változó diszkriminatív erejét fejezi ki.

Normális eloszlást feltételezve:

$$\pi_{j|i} = P(Y_{ij} = 1 \mid X = \mu_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{ij}} \exp(-\frac{1}{2} z^2) dz .$$

A Rasch–féle modell a feltételes valószínűségeket a következőképpen határozza meg:

$$\pi_{j|i} = \left(e^{\mu_i - c_j}\right) / \left(1 + e^{\mu_i - c_j}\right) , \tag{43}$$

ahol a nevező $\triangle_{ij} = (1 + e^{\mu_i - c_j})$ ún. normalizáló konstans.

Vegyük észre, hogy a Rasch–modell speciális faktorelemző modell, ahol a megfigyelt és a látens változók közötti kapcsolat nem–lineáris, és amelyben a faktorsúlyok azonosak (1).

A Rasch–modell likelihood függvénye (ANDERSEN, 1980)

$$L(\underline{\mu}, \underline{c} \mid \underline{y}) = D^{-1} \exp(\sum_i \mu_i y_{i+} - \sum_j c_j y_{tj}), \tag{44}$$

ahol

$D^{-1} = \prod_i \prod_j \triangle_{ij}$.

Az y_{i+} az i-edik megfigyelés „1" válaszainak a száma.

Összesen n y_{i+} értéket figyeltünk meg a J számú változóra, amelyek $J+1$ különböző értéket vehetnek fel. Legyen n_r, $(r = 0, 1, \ldots, J)$, akikre fennáll, hogy $y_{i+} = r$ és $\Sigma_r n_r = n$.

Jelölje R az így csoportosított változót. DUNCAN (1984) megmutatta, hogy a Rasch–féle modell feltételes formája felírható a következő módon:

$$\ln(F_u) = \lambda + \lambda_{R(r)} + \sum_j \lambda_{j(kj)}, \tag{45}$$

ahol

u a gyakoriságtáblázat egy cellája $\{k_1, k_2, \ldots, k_J\}$,

$r = \sum_j k_j$,

$\lambda_{R(r)}$ látens változó fő–hatása,

$\lambda_{j(kj)}$ a j-edik változó fő–hatása.

A modell ebben a formában túlparametrizált, ezért pótlólagos feltételezést kell tennünk (ez a szokásos feltételezés):

$$\sum_r \lambda_{R(r)} = 0,$$

$$\lambda_{j(1)} + \lambda_{j(2)} = 0,$$

és a redundancia elkerülése érdekében:

$$\lambda_{R(J)} = 0.$$

Az (45) loglineáris modell azt fejezi ki, hogy a megfigyelt változók függetlenek az adott R látens változó esetén.

Ez a modell megfelel a látens osztály modellnek, ahol $X = R$, $\pi_n^X = n_r/n$ és $T = J + 1$.

4.3 A GOM (Grade of Membership) modell

A GOM modell feloldja a többi módszernek azt a kötöttségét, hogy egy megfigyelés egy és csakis egy látens osztályba tartozhat, így lényegében egy fuzzy halmaz klasszifikációs eljárás.

Tegyük fel, hogy a megfigyelt változók binárisaku7i$) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{ij}} \exp(-\frac{1}{2}z^2)dz\,.$

A Rasch–féle modell a feltételes valószínűségeket a következőképpen határozza meg:

$$\pi_{j|i} = \left(e^{\mu_i - c_j}\right) / \left(1 + e^{\mu_i - c_j}\right)\,, \tag{43}$$

ahol a nevező $\triangle_{ij} = \left(1 + e^{\mu_i - c_j}\right)$ ún. normalizáló konstans.

Vegyük észre, hogy a Rasch–modell speciális faktorelemző modell, ahol a megfigyelt és a látens változók közötti kapcsolat nem–lineáris, és amelyben a faktorsúlyok azonosak (1).

Az $\underline{x}$ együttes sűrűségfüggvénye kifejezhető a következőképpen:

$$f(\underline{x}) = \int_{R_{\underline{y}}} h(\underline{y})q(\underline{x}|\underline{y})d\underline{y}, \tag{47}$$

ahol

$R_{\underline{y}}$ az $\underline{y}$ változó értéktere.

Miután $\underline{y}$ látens, nem megfigyelt változókat tartalmaz, ezért az $\underline{y}$ változókról az $\underline{x}$ manifeszt változók ismeretében tudunk csak valamit mondani. Ezt a feltételes–sűrűségfüggvénnyel fejezzük ki:

$$h(\underline{y}|\underline{x}) = h(\underline{y})g(\underline{x}|\underline{y})/f(\underline{x})\,. \tag{48}$$

Ahhoz azonban, hogy meghatározzuk $h(\underline{y}|\underline{x})$ feltételes sűrűségfüggvényt, ismernünk kell h és g sűrűségfüggvényeket, de csak az f-et tudjuk becsülni. Ebből látszik, hogy (47) és (48) alapján nem tudjuk egyértelműen definiálni a $h(\underline{y}|\underline{x})$-et, ezért további feltételeket kell tennünk h és g függvényekre.

A modellben csupán azt tételeztük fel, hogy az $\underline{x}$ és $\underline{y}$ változók függnek egymástól. Ha az x változók függnek az y változóktól, akkor az x változók korrelálnak egymással is. Ha az x_1 és x_2 függ az y_1 változótól, akkor az y_1 változó idézi elő az x_1 és x_2 közötti korrelációt. Ha x_1 és x_2 korrelálatlan változók, akkor nem tételezhetjük fel, hogy van valami közös bennük. Ha az x_1 és x_2 korrelálatlanok akkor, ha az y_1 változót konstansul tartjuk (ha a hatását kiszűrjük), akkor nincs szükség további y változókra. Általában ha az x változók közötti korrelációkat az y változók idézik elő, akkor az y változók hatását kiszűrve, az x változók korrelálatlanok lesznek. Ez a feltételes függetlenség axiómája. Eszerint

$$g(\underline{x}|\underline{y}) = \prod_i g_i(x_i\underline{y}) \tag{49}$$

adott r számú látens változó esetén. Feltételezzük tehát, hogy az r számú látens változó teljesen megmagyarázza az x változók kapcsolatait. Az $f(\underline{x})$ függvényt a következőképpen írhatjuk

$$f(\underline{x}) = \int_R h(\underline{y}) \prod_i g_i(x_i|\underline{y})d\underline{y}\,. \tag{50}$$

Könnyű belátni, hogy ha a megfigyelt változók binárisak, és a látens y változónak K kategóriája van, akkor

$$g_i(x_i|y_j) = P(x_i|y_j) = \pi_{ij}^{x_i}(1-\pi_{ij})^{1-x_i}\,,$$

és

$$g(\underline{x}|\underline{y}) = \prod_i \pi_{ij}^{x_i}(1-\pi_{ij})^{1-x_i} . \tag{51}$$

Ha annak a valószínűsége, hogy egy megfigyelés a j-edik látens kategóriába esik η_j $\left(\sum_j \eta_j = 1\right)$, akkor

$$f(\underline{x}) = \sum_j \eta_j \prod_i \pi_{ij}^{x_i}(1-\pi_{ij})^{1-x_i} , \tag{52}$$

és annak a valószínűsége, hogy az $\underline{x}$ megfigyelés a j-edik látens kategóriába esik:

$$h(j|\underline{x}) = \eta_j \prod_i \pi_{ij}^{x_i}(1-\pi_{ij})^{1-x_i}/f(\underline{x}) . \tag{53}$$

Vegyük észre, hogy ez megegyezik a látens osztály modellel bináris megfigyelt változók esetén.

Térjünk most vissza az eredeti problémára, arra, hogy a metrikus manifeszt változókat kategorikus látens változókkal akarjuk kifejezni. Ekkor a megfigyelt változók együttes sűrűségfüggvénye az (50) egyenlet szerint:

$$f(\underline{x}) = \sum_j^K \eta_j \prod_i^m g_i(x_i|j) , \tag{54}$$

ahol

$g_i(x_i|j)$ az x_i változó feltételes eloszlása a j-edik osztályban.

Az (54) modell illesztésére D. J. Bartholomew (1987) munkája alapján két eljárást mutatunk be.

4.4.1 Maximum likelihood becslés

Ez a becslési eljárás lényegét tekintve megegyezik a kategorikus manifeszt változókra kidolgozott eljárással, a különbség $g_i(x_i|j)$ eltérő megválasztásából adódik.

A likelihood–függvény:

$$L = \sum_{k=1}^n \ln f(\underline{x}) = \sum_{h=1}^n \ln \left\{ \sum_j^K \eta_j \prod_i^m g_i(x_i|j) \right\} . \tag{55}$$

Ezt a függvényt kell maximalizálni a $\Sigma\eta_j = 1$ feltétel mellett:

$$\phi = L + \lambda \sum_j \eta_j .$$

A parciális deriválásokat elvégezve:

$$\frac{\partial \phi}{\partial \eta_j} = \sum_h \{g(\underline{x}_h|j)/f(\underline{x}_h)\} + \lambda\,. \tag{56}$$

A Bayes–tétel szerint:

$$h(j|\underline{x}) = \eta_j g(\underline{x}_h|j)/f(\underline{x}_h)\,. \tag{57}$$

Ezt behelyettesítve az (56) egyenletbe:

$$\sum_h h(j|\underline{x}_h) = \lambda \eta_j\,.$$

Mindkét oldalt j szerint összegezve a $\lambda = n$ azonossághoz jutunk, így az első becslő egyenletünk:

$$\widehat{\eta}_j = \sum_h^n h(j|\underline{x}_h)/n\,. \tag{58}$$

Tételezzük fel, hogy

$$g_i(x_i|j) \equiv g(x_i|\theta_{ij})\,.$$

Ekkor

$$\frac{\partial \phi}{\partial \theta_{ij}} = \sum_h \eta_j \frac{\partial g}{\partial \theta_{ij}} / g(x_{ih}|\theta_{ij})\,. \tag{59}$$

Tételezzük fel továbbá, hogy a $g(x_i|\theta_{ij})$ θ_{ij} várható értékű és egységnyi varianciájú normális eloszlást követ. Így:

$$\frac{\partial g}{\partial \theta_{ij}} = (x_{ih} - \theta_{ij}) g(x_i|\theta_{ij})\,.$$

Az (59) egyenletből ezek alapján:

$$\sum_h (h_j|\underline{x}_h)(x_{ih} - \theta_{ij}) = 0\,,$$

amiből a második becslő egyenletet kaphatjuk:

$$\widehat{\theta}_{ij} = \sum_h x_{ih} h(j|\underline{x}_h) / \sum_h h(j|\underline{x}_h)\,, \tag{60}$$

vagy

$$\widehat{\theta}_{ij} = \sum_h x_{ih} h(j|\underline{x}_h)/n\widehat{\eta}_j\,.$$

A $h(j|\underline{x}_h)$ a $\{\eta_j\}$ és $\{\theta_{ij}\}$ paramétereknek a függvénye.

A becslést az (58) és (60) egyenletek alapján végezzük:

(I) kiindulunk az *a posteriori* valószínűségek kezdeti becsléseiből:

$$\{h(j|\underline{x}_h)\},$$

(II) az (58) és (60) egyenletek alapján megkapjuk a $\{\eta_j\}$ és $\{\widehat{\theta}_{ij}\}$ első becsléseit,

(III) ezeket behelyettesítve az (57) egyenletbe a $\{h(j|\underline{x}_n)\}$ újabb becsléseihez jutunk,

(IV) visszatérve a (II) lépéshez, addig folytatjuk a számításokat, amíg az eljárás nem konvergál.

A látens osztály tagságának becslése (Látens osztálybasorolás)

A megfigyelési egységeket a látens osztálytagság *a posteriori* valószínűségei alapján soroljuk be a látens osztályokba.

Ha több mint két osztályunk van, két osztály (j és k) relatív *a posteriori* valószínűsége:

$$h(j|\underline{x})/h(k|\underline{x}).$$

Mivel $h(j|\underline{x}) = \eta_j f(\underline{x}|j)/f(\underline{x})$, az $\underline{x}$ megfigyelés inkább a j-edik osztályba tartozik, ha

$$\frac{h(j|\underline{x})}{h(k|\underline{x})} > 1.$$

Ha az x_i változók együttes eloszlása normális minden osztályban:

$$f(\underline{x}|j) = (2\pi)^{\frac{1}{2}m} \exp\left\{-\frac{1}{2}\sum_i^m (x_i - \mu_i(j))^2\right\}.$$

Így

$$f(\underline{x}|j)/f(\underline{x}|k) = \exp\left\{\sum_i^m x_i\mu_i(j) - \frac{1}{2}\sum_i^m \mu_i^2(j)\right.$$
$$\left. - \sum_i^m x_i\mu_i(h) + \frac{1}{2}\sum_i^m \mu_i^2(j)\right\}.$$

Az osztálybasorolást a

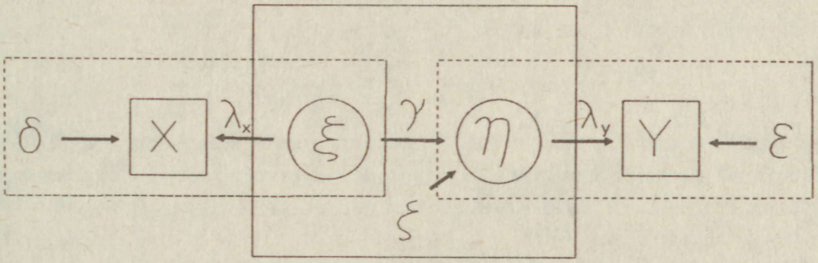
$$\sum_i^m x_i\mu_i(j) - \frac{1}{2}\sum_i^m \mu_i^2(j) + \log\eta_j \tag{61}$$

értékek alapján végezzük. A legvalószínűbb látens osztály az, amely esetén (61) maximális értéket ad. Láthatjuk, hogy a normális eloszlás feltételezésével az osztálybasorolás az adatok lineáris függvénye.

Látens változós modellek

5. A LISREL modell (Analysis of Linear Structural Relationships by the Method of Maximum Likelihood)

Tekintsük legelőször az általános modell szkémáját:



ahol

X: a megfigyelt, manifeszt független (exogén) változók halmaza,

Y: a megfigyelt, manifeszt függő (endogén) változók halmaza,

ξ: nem megfigyelt, látens független (exogén) változó(k),

η: nem megfigyelt, látens függő (endogén) változók(k),

λ_x: a független változók faktorsúlya(i),

λ_y: a függő változók faktorsúlya(i),

δ: a megfigyelt független változók mérési hibája,

ε: a megfigyelt függő változók mérési hibája,

ς: a látens függő változó(k) sztochasztikus reziduális tagja(i).

Az általános modellen belül megkülönböztetjük:

1. *a mérési modelleket*, amelyek azt fejezik ki, hogy a manifeszt változók két komponensre bonthatók, az egyik a szisztematikus komponens, amit a modellben a látens változó fejez ki, a másik a mérési hiba.

A független és függő változókra külön-külön írjuk fel a mérési modellt. (Az ábrán szaggatott vonallal jelöljük.)

2. *a strukturális modellt*, amely a látens változók kauzális összefüggéseit írja le (az általános modellben a szaggatott vonallal körülhatárolt mérési modellek között helyezkedik el.)

Az általános modell feltételei kétfélék:

a) rendszerfeltételek, a modell becsléséhez szükséges általános feltételek,

b) modellfeltételek, amelyekkel speciális modelleket definiálhatunk.

Ezeket a feltételeket később részletezzük.

Az általános modell két látens változó halmaz kauzális struktúráját írja le azzal a feltételezéssel, hogy a látens változók nem mérhetők, hanem mögöttes kifejezői a közvetlenül mérhető manifeszt változóknak, vagyis az egymással kauzális kapcsolatban lévő látens változók mögöttes okai a (mérési modellekben hozzájuk rendelt) megfigyelt változóknak.

A strukturális egyenletekben játszott kauzális kapcsolódásuk alapján a változókat függőnek neveztük, ha az okozat szerepét, és függetlennek, ha az ok szerepét játszották. Szokás ehelyett az eredmény és a magyarázó változó elnevezések használata is.

A változók egy másik osztályozása az, amikor nem egy–egy strukturális egyenletben, hanem az egész modellben játszott kauzális szerepüket ítéljük meg. Eszerint a változók kétfélék, endogén és exogén változók. A strukturális egyenletekben az exogén változók csak az ok szerepét játszhatják, és az okozatok az endogén változók. Az endogén változók között azonban kölcsönös kauzális kapcsolatok is lehetnek, de az exogén változók között ezt nem engedhetjük meg. A modellt szimultánnak nevezzük, ha az endogén változók között kölcsönös kauzális kapcsolatok vannak (mind az ok, mind az okozat szerepét betöltik), a modell rekurzív, ha az endogén változók ugyan lehetnek okok és okozatok is, de közöttük kölcsönös kauzális összefüggés nincsen.

Legyen a látens függő változók vektora $\underline{\eta}' = [\eta_1, \eta_2, \dots, \eta_m]$ és a látens független változók vektora $\underline{\xi}' = [\xi_1, \xi_2, \dots, \xi_n]$.

A strukturális egyenletek modellje:

$$\underline{B}\,\underline{\eta} = \underline{\Gamma}\,\underline{\xi} + \underline{\zeta}, \tag{1}$$

ahol

$\underline{B}$: a függő látens változók közötti regressziós együtthatók mátrixa, $(m \times m)$ típusú,

$\underline{\Gamma}$: a független látens változók és a függő látens változók közötti regressziós együtthatók mátrixa, $(m \times n)$ típusú,

$\underline{\zeta}' = [\zeta_1, \zeta_2, \dots, \zeta_m]$ a függő látens változók reziduális komponense (sztochasztikus reziduális tag).

A modellben szereplő változókról feltételezzük, hogy az átlaguktól való eltéréseket tartalmazzák, és a véletlen sztochasztikus tagok korrelálatlanok egymással és a független változókkal, valamint hogy a $\underline{B}$ mátrix nemszinguláris:

$$E(\underline{\eta}) = E(\underline{\zeta}) = \underline{0} \quad \text{és} \quad E(\underline{\xi}) = \underline{0},$$
$$E(\underline{\xi}\,\underline{\zeta}') = \underline{0} \quad \text{és} \quad \exists\, \underline{B}^{-1}.$$

A mérési modellek

a) a függő megfigyelt, manifeszt (endogén) változók $\underline{y}' = [y_1, y_2, \ldots, y_p]$ mérési modellje:

$$\underline{y} = \underline{\Lambda}_y \underline{\eta} + \underline{\varepsilon}, \tag{2}$$

ahol

$\underline{\Lambda}_y$: $(p \times m)$ típusú mátrix a függő látens változók hatását fejezi ki a függő megfigyelt változóra (faktorsúlyok mátrixa),

$\underline{\varepsilon}$: a függő változók mérési hibája.

Feltételezzük a fentiekhez hasonlóan, hogy

$$E(\underline{y}) = E(\underline{\varepsilon}) = \underline{0} \quad \text{és} \quad E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0}.$$

b) a független megfigyelt, manifeszt (exogén) változók $\underline{x}' = [x_1, x_2, \ldots, x_q]$ mérési modellje:

$$\underline{x} = \underline{\Lambda}_x \underline{\xi} + \underline{\delta}, \tag{3}$$

ahol

$\underline{\Lambda}_x$: $(q \times n)$ típusú mátrix a független látens változók hatását fejezi ki a független megfigyelt változókra (faktorsúlyok mátrixa),

$\underline{\delta}$: a független változók mérési hibája.

Tekintsük át ezután a modell változóit és a változók kapcsolódásait:

1. táblázat

A változók különböző fajtái az általános modellben

Változók	függő (endogén)	független (exogén)	eltérés (reziduális)
Megfigyelt (manifeszt)	$\underline{y}$	$\underline{x}$	–
Nem megfigyelt (látens)	$\underline{\eta}$	$\underline{\xi}$	$\underline{\zeta}, \underline{\varepsilon}, \underline{\delta}$

2. táblázat

A közvetlen hatások lehetséges irányai az általános modellben

Változók	függő (endogén) $\underline{\eta}$	független (exogén) $\underline{\xi}$	eltérés (reziduális) $\underline{\varsigma}\ \underline{\varepsilon}\ \underline{\delta}$
Nem megfigyelt η	$\underline{\beta}$	$\underline{\Gamma}$	$\underline{I}\ \underline{0}\ \underline{0}$
Megfigyelt			
$\underline{y}$	$\underline{\Lambda}_y$	$\underline{0}$	$\underline{0}\ \underline{I}\ \underline{0}$
$\underline{x}$	$\underline{0}$	$\underline{\Lambda}_x$	$\underline{0}\ \underline{0}\ \underline{I}$

A 2. táblázat az általános modell közvetlen hatásait mutatja. Az első sora a következő egyenletet adja:

$$\underline{\eta} = \underline{\beta}\,\underline{\eta} + \underline{\Gamma}\,\underline{\xi} + \underline{I}\,\underline{\varsigma}. \tag{4}$$

Ezt az egyenletet átrendezhetjük:

$$(\underline{I} - \underline{\beta})\underline{\eta} = \underline{\Gamma}\,\underline{\xi} + \underline{\varsigma},$$

vagy az (1) egyenletben szereplő formában:

$$\underline{B}\,\underline{\eta} = \underline{\Gamma}\,\underline{\xi} + \underline{\varsigma},$$

ahol

$$\underline{B} = (\underline{I} - \underline{\beta}) \quad \text{vagy} \quad \underline{\beta} = (\underline{I} - \underline{B}).$$

A fentiek alapján az (1) egyenletben szereplő $\underline{B}$ együtthatómátrixból a közvetlen hatások mátrixát úgy kapjuk meg, hogy vesszük az együtthatókat ellenkező előjellel, a diagonális elemek pedig egyenlőek lesznek 0-val, (mivel a $\underline{B}$ mátrix diagonális elemei 1-gyel egyenlőek).

Az első három egyenlethez tartozó feltételek mellett a következőkben olyan feltételeket említünk, amelyeket minden modell esetén érvényesnek tekintünk, és ezért *rendszerfeltételeknek* nevezzük őket:

1) a mérési hibák és a sztochasztikus reziduális tag között a kovariancia (korreláció) nulla:

$$E(\underline{\varsigma}\,\underline{\varepsilon}') = \underline{0} \qquad E(\underline{\varsigma}\,\underline{\delta}') = \underline{0},$$

2) a két mérési modell hibája között a kovariancia nulla:

$$E(\underline{\varepsilon}\,\underline{\delta}') = \underline{0},$$

3) a látens változók és a manifeszt változók mérési hibája között a kovariancia (korreláció) nulla:

$$E(\underline{\xi}\,\underline{\delta}') = \underline{0} \qquad E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0}.$$

Ezen kívül minden modellre külön feltételeket specifikálhatunk, amelyeket modellfeltételeknek nevezünk. Ezeket majd a későbbiekben tárgyaljuk.

A modelleket akkor tekintjük adottnak, ha a következő nyolc paramétermátrixot specifikáltuk:

$\underline{B}$: a látens (endogén) függő változók közötti regressziós együtthatók mátrixa $(m \times m)$ típusú, általános eleme β_{ij} a j-edik endogén változónak az i-edik endogén változóra kifejtett közvetlen hatását mutatja az előjel megfordításával.

$\underline{\Gamma}$: a független (exogén) látens változónak a függő (endogén) látens változóra vonatkozó regressziós együtthatók mátrixa, $(m \times n)$ típusú, általános eleme γ_{ij} a j-edik exogén változónak az i-edik endogén változóra kifejtett közvetlen hatását mutatja.

$\underline{\Phi}$: a független (exogén) látens változók közötti kovariancia (korreláció) mátrix, $(n \times n)$ típusú, általános eleme Φ_{ij} az i-edik és j-edik exogén változók közötti kovarianciát (korrelációt) fejezi ki $(E(\xi_i\ \xi_j))$, a Φ_{ii} diagonális elem az i-edik exogén látens változó varianciája $E(\xi_i\ \xi_i)$.

$\underline{\Psi}$: az endogén változók sztochasztikus reziduális tagjai között számított kovariancia (korreláció) mátrix, $(m \times m)$ típusú, általános eleme Ψ_{ij} az i-edik és j-edik reziduális komponens között számolt kovariancia (korreláció) $E(\zeta_i\ \zeta_j)$, és a Ψ_{ii} jelöli az i-edik reziduum varianciáját $E(\zeta_i\ \zeta_i)$.

$\underline{\Lambda}_x$: a megfigyelt független (exogén) változók faktorsúly–mátrixa, $(q \times n)$ típusú, az általános eleme λ_{xij} a j-edik látens exogén változó közvetlen hatását fejezi ki az i-edik megfigyelt változóra.

$\underline{\Theta}_\delta$: a megfigyelt független (exogén) változók hiba tagjainak kovariancia mátrixa, $(q \times q)$ típusú, általános eleme $\Theta_{\delta ij}$ az i-edik és j-edik mérési exogén változók hiba tagjainak kovarianciája $E(\delta_i\ \delta_j)$, a $\Theta_{\delta ii}$ jelöli az i-edik változó hibájának a varianciáját $E(\delta_i\ \delta_i)$.

$\underline{\Lambda}_y$: a megfigyelt függő (endogén) változók faktorsúly–mátrixa, $(p \times m)$ típusú, általános eleme λ_{yij} a j-edik látens változó közvetlen hatását fejezi ki az i-edik megfigyelt endogén változóra.

$\underline{\Theta}_\varepsilon$: a megfigyelt függő (endogén) változók hiba tagjainak kovariancia mátrixa, $(p \times p)$ típusú, általános eleme $\Theta_{\varepsilon ij}$ az i-edik és j-edik mérési endogén változók hiba tagjainak kovarianciája $E(\varepsilon_i\ \varepsilon_i)$, a $\Theta_{\varepsilon ii}$ jelöli az i-edik változó hibájának a varianciáját $E(\varepsilon_i\ \varepsilon_i)$.

5.1 A strukturális egyenlet redukált formája

A következőkben először a strukturális egyenlet redukált formáját vizsgáljuk meg, majd a megfigyelt változók kovariancia mátrixát a paraméterek függvényeként határozzuk meg.

Az (1) strukturális egyenletet ($\underline{B}\,\underline{\eta} = \underline{\Gamma}\,\underline{\xi} + \underline{\varsigma}$) balról megszorozzuk $\underline{B}^{-1}$ inverz mátrixszal; így a strukturális egyenlet redukált formájához jutunk:

$$\underline{\eta} = \underline{D}\,\underline{\xi} + \underline{\varsigma}_r\,. \tag{5}$$

A $\underline{D}$ mátrix a redukált forma együtthatóit tartalmazza:

$$\underline{D} = \underline{B}^{-1}\underline{\Gamma} \qquad \text{és} \qquad \underline{\varsigma}_r = \underline{B}^{-1}\underline{\varsigma}\,. \tag{6}$$

A függő változók ($\underline{\eta}$) variancia–kovariancia mátrixa:

$$\underline{C} = E(\underline{\eta}\,\underline{\eta}') = E(\underline{B}^{-1}\underline{\Gamma}\,\underline{\xi} + \underline{B}^{-1}\underline{\varsigma})(\underline{B}^{-1}\underline{\Gamma}\,\underline{\xi} + \underline{B}^{-1}\underline{\varsigma})\,,$$

mivel $E(\underline{\xi}\,\underline{\xi}') = \underline{\Phi}$, $E(\underline{\varsigma}\,\underline{\varsigma}') = \underline{\Psi}$, $E(\underline{\xi}, \underline{\varsigma}) = \underline{0}$,
$E(\underline{\varsigma}\,\underline{\xi}') = \underline{0}$ és $E(\underline{\varsigma}_r\,\underline{\varsigma}_r) = \underline{B}^{-1}\underline{\Psi}\,\underline{B}'^{-1}$.

A fentiek alapján:

$$\underline{C} = \underline{D}\,\underline{\Phi}\,\underline{D}' + \underline{B}^{-1}\underline{\Psi}\,\underline{B}'^{-1}\,. \tag{7}$$

5.2 A megfigyelt változók variancia–kovariancia mátrixa

A függő és független megfigyelt változókat helyezzük a $\underline{z}$ vektorba, és tegyük fel, hogy a megfigyelt változókat az átlaguktól való eltérésekkel mértük: $\underline{z}' = [\underline{y}', \underline{x}']$.

Ekkor

$$\underline{\Sigma} = E(\underline{z}\,\underline{z}') = \begin{pmatrix} E(\underline{y}\,\underline{y}') & E(\underline{y}\,\underline{x}') \\ E(\underline{x}\,\underline{y}') & E(\underline{x}\,\underline{x}') \end{pmatrix} = \begin{pmatrix} \underline{\Sigma}_{yy} & \underline{\Sigma}_{yx} \\ \underline{\Sigma}_{xy} & \underline{\Sigma}_{xx} \end{pmatrix}. \tag{8}$$

A $\underline{\Sigma}$ különböző elemeit a modellel a következőképpen kaphatjuk meg:

$$\underline{\Sigma}_{yy} = E(\underline{y}\,\underline{y}') = E[(\underline{\Lambda}_y\underline{\eta} + \underline{\varepsilon})(\underline{\Lambda}_y\underline{\eta} + \underline{\varepsilon})']\,,$$

mivel $E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0}$, $E(\underline{\varepsilon}\,\underline{\eta}') = \underline{0}$, $E(\underline{\varepsilon}\,\underline{\varepsilon}') = \underline{\Theta}_\varepsilon$
és $E(\underline{\eta}\,\underline{\eta}') = \underline{C}$, így
$\underline{\Sigma}_{yy} = \underline{\Lambda}_y\underline{C}\,\underline{\Lambda}_y' + \underline{\Theta}_\varepsilon$.

Hasonló módon kapjuk a $\underline{\Sigma}_{yx}$ elemeit is:

$$\underline{\Sigma}_{yx} = E(\underline{y}\,\underline{x}') = E[(\underline{\Lambda}_y\underline{\eta} + \underline{\varepsilon})(\underline{\Lambda}_x\underline{\xi} + \underline{\delta})']\,,$$

$$E(\underline{\eta}\,\underline{\delta}') = \underline{0}\,, \quad E(\underline{\varepsilon}\,\underline{\xi}') = \underline{0}\,, \quad E(\underline{\varepsilon}\,\underline{\delta}') = \underline{0}\,,$$

mivel $E(\underline{\zeta}\,\underline{\xi}') = 0\,, \quad E(\underline{\eta}\,\underline{\xi}') = \underline{D}\,\underline{\Phi},$ így

$$\underline{\Sigma}_{yx} = \underline{\Lambda}_y\,\underline{D}\,\underline{\Phi}\,\underline{\Lambda}_x'\,.$$

Mivel $\underline{\Sigma}_{xy} = \underline{\Sigma}_{yx}'$, így ismerjük már a $\underline{\Sigma}_{xy}$–t is.

Az alsó diagonális blokk $\underline{\Sigma}_{xx}$:

$$\underline{\Sigma}_{xx} = E(\underline{x}\,\underline{x}') = E[(\underline{\Lambda}_x\underline{\xi} + \underline{\delta})(\underline{\Lambda}_x\underline{\xi} + \underline{\delta})'] \,,$$

mivel

$$E(\underline{\xi}\,\underline{\xi}') = \underline{\Phi}, \quad E(\underline{\xi}\,\underline{\delta}') = \underline{0}, \quad E(\underline{\delta}\,\underline{\xi}') = \underline{0}, \quad \text{és}$$

$$E(\underline{\delta}\,\underline{\delta}') = \underline{\Theta}_\delta\,, \quad \text{így}$$

$$\underline{\Sigma}_{xx} = \underline{\Lambda}_x\underline{\Phi}\,\underline{\Lambda}_x' + \underline{\Theta}_\delta\,.$$

A fentiek alapján a megfigyelt változók variancia–kovariancia mátrixát a paraméterek mátrixaival kifejezve:

$$\underline{\Sigma} = \begin{pmatrix} \underline{\Lambda}_y\underline{C}\,\underline{\Lambda}_y' + \underline{\Theta}_\varepsilon & \underline{\Lambda}_y\underline{D}\,\underline{\Phi}\,\underline{\Lambda}_x' \\ \underline{\Lambda}_x\underline{\Phi}\,\underline{D}'\underline{\Lambda}_y' & \underline{\Lambda}_x\underline{\Phi}\,\underline{\Lambda}_x' + \underline{\Theta}_\delta \end{pmatrix},$$

vagy a $\underline{C}$ és $\underline{D}$ mátrixokat a (6) és (7) azonosságokkal helyettesítve:

$$\underline{\Sigma} = \begin{pmatrix} \underline{\Lambda}_y(\underline{B}^{-1}\underline{\Gamma}\,\underline{\Phi}\,\underline{\Gamma}'\underline{B}'^{-1} + \underline{B}^{-1}\underline{\Psi}\,\underline{B}'^{-1})\,\underline{\Lambda}_y + \underline{\Theta}_\varepsilon & \underline{\Lambda}_y\underline{B}^{-1}\underline{\Gamma}\,\underline{\Phi}\,\underline{\Lambda}_x' \\ \underline{\Lambda}_x\underline{\Phi}\,\underline{\Gamma}'\underline{B}'^{-1}\underline{\Lambda}_y' & \underline{\Lambda}_x\underline{\Phi}\,\underline{\Lambda}_x' + \underline{\Theta}_\delta \end{pmatrix}. \tag{9}$$

A (8) egyenlet a megfigyelt változók variancia–kovariancia mátrixa ($\underline{\Sigma}$) és a modell $\underline{B}\,,\underline{\Gamma}\,,\underline{\Phi}\,,\underline{\Psi}\,,\underline{\Lambda}_y\,,\underline{\Lambda}_x\,,\underline{\Theta}_\varepsilon\,,\underline{\Theta}_\delta$ paraméterei közötti kapcsolatot írja le. A (8) egyenlet alapján az általános modell vizsgálatának fontossága abban látszik, hogy a $\underline{\Sigma}$ mátrix és a modell paraméterei közötti kapcsolatot nem kell minden speciális modell esetén megvizsgálni, hanem az általános modell eredményei alkalmazhatók a speciális modellekre is.

5.3 Standardizálás

Az eddigiekben feltételeztük, hogy a megfigyelt változókat a várható értéküktől való eltéréseikkel mértük. Ismeretes, hogy ha a várható értéküktől való elérésekkel mért változókat elosztjuk a szórásukkal, standardizált változókhoz jutunk, amelyekre az a jellemző, hogy várható értékük nulla, szórásuk pedig egy.

Nézzük először azt az esetet, amikor a látens változók standardizáltak. Az η és ξ várható értéke feltételezésünk szerint nulla.

Az η és ξ változókat úgy standardizáljuk, hogy az $\underline{A}_\eta^{-1}$ és az $\underline{A}_\xi^{-1}$ mátrixokkal beszorozva újraskálázzuk őket, ahol $\underline{A}_\eta = (\text{diag}\,.\underline{C})^{1/2}$ és $\underline{A}_\xi = (\text{diag}\,.\underline{\Phi})^{1/2}$.

A $\underline{\Lambda}_y, \underline{\Lambda}_x, \underline{B}, \underline{\Gamma}, \underline{\Phi}, \underline{\Psi}$ paramétermátrixokat a standardizált látens változók esetén a következőképpen transzformáljuk:

$$
\begin{aligned}
\underline{\Lambda}_y^* &= \underline{\Lambda}_y \underline{A}_\eta \\
\underline{\Lambda}_x^* &= \underline{\Lambda}_x \underline{A}_\xi \\
\underline{B}^* &= \underline{A}_\eta^{-1} \underline{A}_\eta \\
\underline{\Gamma}^* &= \underline{A}_\eta^{-1} \underline{\Gamma}\, \underline{A}_\eta \\
\underline{\Phi}^* &= \underline{A}_\xi^{-1} \underline{\Phi}\, \underline{A}_\xi^{-1} \\
\underline{\Psi}^* &= \underline{A}_\eta^{-1} \underline{\Psi}\, \underline{A}_\eta^{-1}\,.
\end{aligned}
\tag{10}
$$

A látens változók a standardizálás után is előállítják a megfigyelt változók eredeti értékeit és így a kovariancia mátrixot is:

$$
\begin{aligned}
\underline{y} &= \underline{\Lambda}_y \underline{A}_\eta\, \underline{A}_\eta^{-1} \underline{\eta} + \underline{\varepsilon} = \underline{\Lambda}_y \underline{\eta} + \underline{\varepsilon} \\
\underline{x} &= \underline{\Lambda}_x \underline{A}_\xi\, \underline{A}_\xi^{-1} \underline{\xi} + \underline{\delta} = \underline{\Lambda}_x \underline{\xi} + \underline{\delta}\,.
\end{aligned}
$$

Ha $\underline{y}$ és $\underline{x}$ megfigyelt változókat standardizáljuk, megszorozva megfigyelt értékeiket az $\underline{A}_y^{-1}$ és $\underline{A}_x^{-1}$ mátrixokkal, ahol

$$
\begin{aligned}
\underline{A}_y &= (\text{diag}\,\underline{\Sigma}_{yy})^{1/2} \\
\underline{A}_x &= (\text{diag}\,\underline{\Sigma}_{xx})^{1/2}\,.
\end{aligned}
$$

$\underline{\Lambda}_y$, $\underline{\Lambda}_x$, valamint $\underline{\Theta}_\varepsilon$ és $\underline{\Theta}_\delta$ együtthatók a következőképpen transzformálódnak:

$$
\begin{array}{ll}
\underline{A}_y^{-1} \underline{\Lambda}_y^* & \qquad \underline{A}_y^{-1} \underline{\Theta}_\varepsilon \underline{A}_y^{-1} \\
\underline{A}_x^{-1} \underline{\Lambda}_y^* & \qquad \underline{A}_x^{-1} \Theta_\delta \underline{A}_x^{-1}\,.
\end{array}
\tag{11}
$$

Megjegyezzük, hogy standardizált változók esetén a variancia–kovariancia mátrix egyenlő a korrelációs mátrixszal, és hogy a standardizált együtthatókat a szociológiai irodalomban út–együtthatóknak nevezik.

5.4 Identifikáció

Mielőtt a paraméterek becslését megadnánk, az identifikáció problémáját kell megvizsgálnunk. Feltételezzük, hogy a megfigyelt változók valószínűségeloszlása jól jellemezhető az első és második momentumaikkal és így a magasabbrendű momentumok elhanyagolhatók. Gyakorlatilag ezzel feltételezzük, hogy a megfigyelt változók együttes eloszlása normális. Mivel a változók várható értékeiről feltételeztük, hogy egyenlők nullával, a $\underline{z}' = [\underline{y}', \underline{x}']$ eloszlását a $\underline{\Lambda}_y, \underline{\Lambda}_x, \underline{B}, \underline{\Gamma}, \underline{\Phi}, \underline{\Psi}, \underline{\Theta}_\epsilon, \Theta_\delta$, paraméterektől függő variancia–kovariancia mátrixszal jellemezzük. Jelölje $\underline{\pi}$ vektor az összes paramétert, és legyen t a vektor elemeinek a száma. A variancia–kovariancia mátrix $\underline{\Sigma}$ általános elemére felírhatjuk a következő összefüggést:

$$\sigma_{ij} = f_{ij}(\underline{\pi}), \tag{12}$$

ami azt fejezi ki, hogy a megfigyelt változók varianciái és kovarianciái a $\pi_1, \pi_2, \ldots, \pi_t$ paraméterek függvényei. A $\sigma_{ij} = f_{ij}(\underline{\pi})$ függvényekről feltételezzük, hogy folytonosak és folytonosak az első deriváltjai is. A $\underline{\Sigma}$ variancia–kovariancia mátrixról feltételezzük, hogy a lehetséges paraméter–tér minden $\underline{\pi}$ pontjában pozitív definit.

Egy specifikált modell, egy adott struktúra ($\underline{\Lambda}_y, \underline{\Lambda}_x, \underline{B}, \underline{\Gamma}, \underline{\Phi}, \underline{\Psi}, \underline{\Theta}_\epsilon, \underline{\Theta}_\delta$, vagy egyszerűen $\underline{\pi}$) egy és csak egy $\underline{\Sigma}$-t generál. Ugyanakkor különböző struktúrák is generálhatják ugyanazt a $\underline{\Sigma}$-t. Ha két vagy több struktúra ugyanazt a $\underline{\Sigma}$-t generálja, akkor ezeket empirikusan ekvivalens struktúráknak nevezzük. Ha egy paraméter az összes ekvivalens struktúrában azonos értéket vesz fel, akkor a kérdéses paraméter identifikálható. Ha a modell valamennyi paramétere identifikálható, akkor a modellt identifikálhatónak nevezzük. Más szóval a modellt identifikálhatónak nevezzük, ha a paraméter–tér bármely két $\underline{\pi}_1 \neq \underline{\pi}_2$ vektorára igaz, hogy $\underline{\Sigma}(\underline{\pi}_1) \neq \underline{\Sigma}(\underline{\pi}_2)$, vagyis a variancia–kovariancia mátrixot ($\underline{\Sigma}$) egy és csak egy $\underline{\pi}$ generálja. Ha a modell nem identifikálható, abból még nem következik, hogy a modellnek nincs identifikálható paramétere.

Az identifikálhatóság a modell megválasztásától függ. Az identifikálhatóságot vizsgálva induljunk ki a (12) egyenletből:

$$\sigma_{ij} = f_{ij}(\underline{\pi}) \qquad i \leq j\,.$$

Összesen $(1/2)(p+q)(p+q+1)$ egyenletünk van t ismeretlennel (a $\underline{\pi}$ paraméter–vektor elemeinek száma). Eszerint a modell identifikálhatóságának szükséges feltétele, hogy a

$$t \leq (1/2)(p+q)(p+q+1)$$

reláció teljesüljön. Ha egy paramétert a $\underline{\Sigma}$ mátrixból meg tudunk határozni, a kérdéses paraméter identifikálható, különben nem . Azonban előfordulhat, hogy néhány paramétert többféleképpen is meghatározhatunk, (a (12)-ből különböző egyenleteket felhasználva). Ez a túlidentifikáltság feltétele. Mivel a (12) egyenlet gyakran nem lineáris, az egyenletek megoldása bonyolult, és a $\underline{\pi}$ explicit megoldása ritkán létezik, az identifikáció vizsgálatára egy ún. információs mátrixot konstruálunk, és az információs mátrix vizsgálata alapján döntjük el, hogy a modell, illetve annak paraméterei identifikálhatók–e. Ha az információs mátrix pozitív definit, akkor a modell identifikálható. Ha az információs mátrix szinguláris, akkor a modell nem identifikálható, és az információs mátrix rangja alapján ítélhető meg, hogy mely paraméterek nem identifikálhatók.

5.5 *A paraméterek becslése*

Feltételezzük, hogy a megfigyelt változók eloszlása a várható értékekkel és a kovariancia mátrixszal jellemezhető.

A kovariancia mátrix $\underline{\Sigma}$ a $\underline{\pi}$ paraméter függvénye. A paraméterek a gyakorlaban ismeretlenek, és az n elemű független minta alapján becsüljük őket. Az $(\underline{y}', \underline{x}')$ megfigyelt változók mintabeli kovariancia mátrixát $\underline{S}$ mátrix jelölje. A gyakorlatban a változók mérésénél a mérési skála egysége gyakran önkényes, vagy indifferens, ezért a kovariancia mátrix helyett szoktuk a korrelációs mátrixot is használni $(\underline{R})$.

Mivel a változókat a várható értéküktől való eltérésükkel jellemezzük, a becslés problémája tulajdonképpen az, hogy a $\underline{\pi}$ paraméterektől függő $\underline{\Sigma}$ mátrixot illesszük a mintabeli $\underline{S}$ kovariancia mátrixhoz. Az illesztésre három eljárást tekintsünk át.

Az első a legegyszerűbb. Minimalizáljuk a megfigyelt és a modell által reprodukált varianciák és kovarianciák különbségeinek négyzetösszegét:

$$U = \sum_{i}^{p+q} \sum_{j}^{j<i} (S_{ij} - \sigma_{ij})^2 = \sum_{i}^{p+q} \sum_{j}^{j<i} (S_{ij} - f_{ij}(\underline{\pi}))^2 \,. \tag{13}$$

Vagy mátrix jelöléssel:

$$U = \frac{1}{2}\mathrm{tr}(\underline{S} - \underline{\Sigma})^2 \,. \tag{14}$$

Ez a közismert legkisebb négyzetek módszere (unweighted least squares, ULS), amellyel a reziduális variancia–kovarianciákat minimalizáljuk.

A második illesztő eljárás az általánosított legkisebb négyzetek módszere (generalized least squares, GLS). Ennek lényege, hogy az $(\underline{S} - \underline{\Sigma})$ reziduális kovariancia mátrixot egy $\underline{V}$ súlymátrixszal megszorozzuk, és a súlyozott reziduális négyzetösszeget minimalizáljuk:

$$G = \frac{1}{2}\mathrm{tr}[\underline{V}(\underline{S} - \underline{\Sigma})]^2 \,. \tag{15}$$

Ha a súlymátrix egységmátrix, akkor visszakapjuk a súlyozatlan legkisebb négyzetek módszerét. A gyakorlatban súlynak a $\underline{V} = \underline{S}^{-1}$ mátrixot szokták megadni, akkor

$$G = \frac{1}{2}\mathrm{tr}(\underline{I} - \underline{S}^{-1}\underline{\Sigma})^2 \,. \tag{16}$$

Mindkét eljárás nagy előnye, hogy a változók eloszlására nem tesz feltételt.

A harmadik eljárás a maximum likelihood–módszer. Ennél az eljárásnál a $\underline{\Sigma}$ adott értékeire az $\underline{S}$ elemeinek sűrűségfüggvényét ismernünk kell. Ha ismerjük az $\underline{y}$ és $\underline{x}$ változók eloszlását, akkor $\underline{S}$ sűrűségfüggvényét is meghatározhatjuk. Jelöljük ezt a sűrűségfüggvényt $f(\underline{S}|\underline{\Sigma})$-val. Ez a függvény $\underline{\Sigma}$-n keresztül függ a $\underline{\pi}$ paramétertől. A mintából számított $\underline{S}$ kovariancia mátrix elemeit behelyettesítve a fenti függvénybe jutunk a likelihood–függvényhez (L). A $\underline{\pi}$ maximum likelihood $\hat{\underline{\pi}}$ becslésén a $\underline{\pi}$-nek azt az értékét tekintjük, amely mellett a likelihood–függvény felveszi a maximumát. Más szóval, adott minta esetén $\underline{\pi}$ becsléséül azt fogadjuk el, amely mellett éppen ennek a mintának a bekövetkezése a legvalószínűbb.

A likelihood–függvény (L) helyett a gyakorlatban annak logaritmusát maximalizáljuk $(\log L)$. A $\hat{\underline{\pi}}$ értékeit szintén nem változtatja, ha $\log L$ helyett a $(c_1 \log L + c_2)$-t maximalizáljuk, vagy az $F = -(c_1 \log L + c_2)$ függvényt minimalizáljuk, ahol c_1 és c_2 konstansok.

Feltételezve, hogy a megfigyelt változók eloszlása normális, a likelihood–függvény (ANDERSON, L. (1958) p. 159) logaritmusa egy konstans elhagyásával a következő:

$$\log L = -\frac{1}{2}(N-1)[\log|\underline{\Sigma}| + \mathrm{tr}(\underline{S}\,\underline{\Sigma}^{-1})], \tag{17}$$

ahol

$|\underline{\Sigma}|$: a $\underline{\Sigma}$ determinánsát jelöli,

$\mathrm{tr}(\underline{S}\,\underline{\Sigma}^{-1})$: az $(\underline{S}\,\underline{\Sigma}^{-1})$ mátrix szorzat diagonális elemeinek összegét jelöli,

N: a megfigyelési egységek száma.

Jöreskog javasolta az F függvény minimalizálását $\underline{\pi}$ paraméter becslésére

$$F = \log|\underline{\Sigma}| + \mathrm{tr}(\underline{S}\,\underline{\Sigma}^{-1}) - \log|\underline{S}| - (p+q)\,. \tag{18}$$

Az F függvény minimalizálása ugyanazt a $\hat{\underline{\pi}}$ értéket adja, mint L maximalizálása, mivel

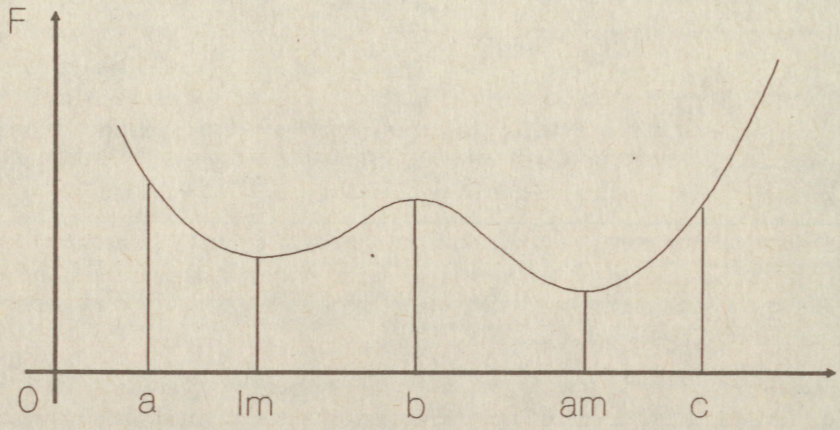
$$F = -(c_1 \log L + c_2)\,,$$

ahol

$$c_1 = \frac{2}{N-1}$$
$$c_2 = \log|\underline{S}| + (p+q) \quad \text{konstansok}\,.$$

Az F függvény minimalizálását a FLETCHER–POWELL (1963)–féle eljárással végezzük. Az eljárás F első és másodrendű deriváltjait használja, és egy önkényes kezdőpontból elindulva gyors konvergenciát biztosít egy lokális minimumhoz. Ha az F függvénynek több minimumhelye létezik, akkor nincs garancia, hogy a globális optimumot megtaláljuk. A következő ábrából látszik, hogy ha a kezdőpontunk a és

b közé esik, akkor az eljárás lokális optimumhoz konvergál, míg ha a kezdőpont a b és c értékek közé esik, akkor a program az abszolút minimumot fogja megtalálni.



A paraméterek standard hibáját az ún. információs mátrixból számítjuk, invertáljuk az információs mátrixot, és az inverz diagonális elemei adják a paraméterek standard hibáit. A standard hibák ismeretében a paraméterekre konfidencia intervallumot szerkeszthetünk. Ha $\widehat{\pi}_i$ a π_i paraméter becslését jelöli, és $\widehat{\sigma}_{\widehat{\pi}_i}$ jelöli a $\widehat{\pi}_i$ standard hibáját (a $\widehat{\pi}_i$ mintabeli szórását), akkor a becslés valódi értéktől való eltérése osztva a standard hibával nagy minta esetén standard normális eloszlású:

$$z = \frac{\pi_i - \widehat{\pi}_i}{\widehat{\sigma}_{\widehat{\pi}_i}}.$$

Ennek alapján a π_i paraméter értéke $1 - \alpha$ valószínűséggel a következő intervallumba esik:

$$P(\widehat{\pi}_i - z_\alpha \widehat{\sigma}_{\widehat{\pi}_i} \leq \pi_i \leq \widehat{\pi}_i + z_\alpha \widehat{\sigma}_{\widehat{\pi}_i}) = 1 - \alpha. \qquad (19)$$

Ha például $\widehat{\pi}_i = 3$ és $\widehat{\sigma}_{\widehat{\pi}_i} = 2$ és a paraméter 95%-os konfidencia intervallumát szeretnénk tudni, akkor $z = 1,96$ és így a 95%-os konfidencia intervallum (3-1,96·0,2, 3-1,96·0,2), vagyis a kérdéses paraméter az (2.608, 3.392) intervallumba esik 95%-os valószínűséggel.

5.6 A modell tesztelése

A paraméterek maximum likelihood becslése után a modell illeszkedésének a jóságát a likelihood–hányados technikával tesztelhetjük. Legyen H_0 az adott modell nullhipotézise.

Tekintsük először azt az alternatív hipotézist H_1, hogy a $\underline{\Sigma}$ mátrix tetszőleges pozitív definit mátrix. Ekkor minusz kétszer a likelihood–hányados logaritmusa egyenlő lesz $(N/2)F_0$-val, ahol F_0 az F függvény minimuma, és N a minta elemeinek száma. Nagy minta esetén ez χ^2 eloszláshoz vezet

$$d = \frac{1}{2}(p+q)(p+q+1) - t \quad \text{szabadságfokkal},$$

ahol:

t: a paraméterek száma a H_0 hipotézisben.

Az adott modellben (amelyre a H_0 hipotézist fogalmaztuk meg) a paraméterek száma legyen egyenlő u-val. Az adott modell paraméterei ($\underline{\gamma}$) részhalmazát adja $\underline{\pi}$-nak, és $u < t$.

Legyen F_0 az F függvény minimuma a H_0 hipotézis mellett, és a H_1 hipotézisnek megfelelő modell F függvényének minimumhelyét jelölje F_1. Ekkor $(N/2)(F_0 - F_1)$ közelítőleg χ^2 eloszlású $(t-u)$ szabadságfokkal.

Az illeszkedés jóságát kifejező χ^2 értékét a következőképpen értelmezzük. Az adott modellre kapott χ^2 értéket a megfelelő szabadságfok mellett összehasonlítjuk az elméleti (táblázatból kiolvasható) χ^2 értékekkel, és megkeressük a szignifikanciaszinteknek azt a tartományát, amely mellett a táblázatbeli (elméleti) χ^2 érték nagyobb az adott modell mintából számított χ^2 értékénél. Ezen tartományon belül a modellt elfogadjuk, ezen kívül pedig elvetjük. Általában a $p = .05$ szignifikanciaszinthez (ez 95%-os valószínűségi állítást jelent) meghatározzuk a χ^2 kritikus (elméleti) értékét, és ha az adott modell számított értéke ennél nagyobb, a hipotézist (az adott modellt) elvetjük, és megvizsgáljuk a mintából számított kovarianciák ($\underline{S}$) és a modell által becsült $\underline{\Sigma}$ kovarianciák eltéréseit, valamint az F függvény első deriváltjait. Ezek alapján módosítjuk a modellt, általában úgy, hogy több paramétert veszünk, engedünk meg a modellben, és újra elvégezzük az illesztést. Így általában kisebb χ^2 értékhez jutunk. Ha ez a javulás (a két χ^2 érték különbsége) a megfelelő szabadságfok (a két szabadságfok különbsége) mellett szignifikáns, akkor a modell változtatása jelentős, helyes volt, egyébként a modell illeszkedését jelentősen nem javítottuk.

5.7 Az általános modell speciális esetei

A következőkben az általános modell speciális eseteit vesszük sorra.

5.7.1 A klasszikus mérési modell

A klasszikus mérési modell szerint a megfigyelt (manifeszt) változók mérési értékeit a valódi érték (ξ) és a mérési hiba (δ) összegeként kapjuk meg, más szóval a mérés két hatás eredője, az egyik hatás a valódi (a hipotetikus hatás), a másik a hiba hatás. Formalizálva ezt a következőképpen írhatjuk fel:

$$x = \xi + \delta\,, \tag{20}$$

ahol

$$E(\delta) = 0\,, \quad E(x) = 0\,, \quad E(\xi) = 0\,, \quad E(\xi\delta) = 0\,.$$

Mivel a regressziós együttható egyenlő eggyel, a valódi értéket a megfigyelt értékkel azonos skálán fejezzük ki.

Ha a valódi és a megfigyelt értéket standardizáljuk, a fenti mérési modellt a következőképpen írhatjuk fel:

$$x^* = \lambda^* \xi^* + \delta^*\,, \tag{21}$$

ahol

$$x^* = \frac{x}{\sigma_x}\,, \qquad \lambda^* = \frac{\sigma_\xi}{\sigma_x}\,,$$
$$\delta^* = \frac{1}{\sigma_x}\delta\,, \qquad \xi^* = \frac{\xi}{\sigma_\xi}\,,$$

és

$$E(x^*) = 0\,, \quad E(\xi^*) = 0\,, \quad E(\delta^*) = 0\,, \quad E(\xi^*\delta^*) = 0\,,$$
$$E(\xi^*\xi^*) = 1\,, \quad E(x^*x^*) = 1\,.$$

A valódi és a megfigyelt értékek közötti korreláció négyzete a mérés megbízhatóságát fejezi ki, és megbízhatósági együtthatónak nevezzük (ρ^2):

$$\rho_{x\xi} = E(x^*\xi^*) = E[(\lambda^*\xi^* + \delta^*)(\xi^*)] = \lambda^* = \frac{\sigma_\xi}{\sigma_x}\,,$$

és

$$\rho^2_{x\xi} = \lambda^{*2} = \frac{\sigma^2_\xi}{\sigma^2_x} = \frac{\Phi}{\sigma^2_x}\,. \tag{22}$$

A megbízhatósági együttható egyenlő a standardizált regressziós együttható négyzetével, vagy egyenlő a valódi értékek varianciája és a megfigyelt értékek varianciája hányadosával. Ha legalább két manifeszt változó hipotetikus komponense megegyezik (a valódi értékük ugyanaz), és csak a hiba komponensben különböznek, a következő mérési modellt írhatjuk fel:

$$x_i = \xi + \delta_i \qquad \forall i\text{–re}, \tag{23}$$

ahol

$$\begin{array}{llll} E(x_i) = 0, & E(\xi) = 0, & E(\delta_i) = 0, & \forall i\text{–re}, \\ E(\xi\delta_i) = 0, & & & \forall i\text{–re}, \\ E(\delta_i\delta_j) = \Theta_{\delta_{ij}} = 0, & & & \forall i\text{–re}, \\ \Theta_{\delta_{ii}} = \Theta_{\delta_{jj}}, & & & \forall i, j\text{–re}. \end{array}$$

Azokat a mérőeszközöket, amelyek eleget tesznek a fenti modell követelményeinek, „parallel instruments", „párhuzamos mérőeszközöknek" nevezzük.

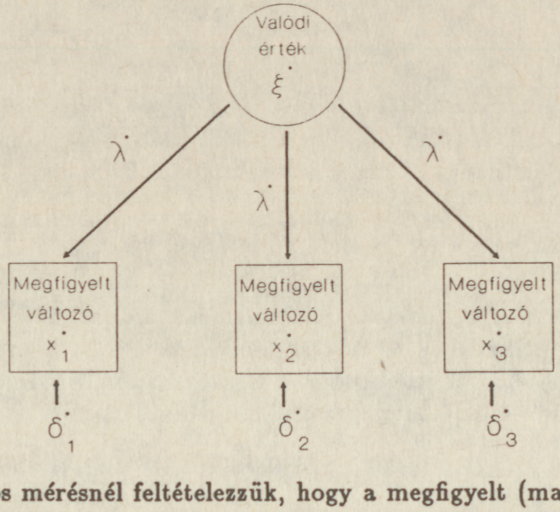
A modell standardizált változókkal:

$$x_i^* = \lambda_i^* \xi^* + \delta_i^*, \tag{24}$$

ahol

$$\begin{array}{llll} E(x_i^*) = 0, & E(\xi^*) = 0, & E(\delta_i^*) = 0, & \forall i\text{-re}, \\ \sigma^2_{x_i^*} = 1, & \sigma^2_{\xi^*} = \Phi = 1, & & \forall i\text{-re}, \\ E(\xi^*\delta_i^*) = 0, & & & \forall i\text{-re}, \\ E(\delta_i^*\delta_j^*) = \Theta_{\delta_{ij}} = 0, & & & \forall i, j\text{-re}, \\ \Theta_{\delta_{ii}} = \Theta_{\delta_{jj}}, & & & \forall i, j\text{-re}, \\ \lambda_i^* = \lambda_j^*, & & & \forall i, j\text{-re}. \end{array}$$

Az utolsó feltétel azt jelenti, hogy a megbízhatósági együttható (λ^{*2}) minden párhuzamos mérőeszköz esetén azonos. A „párhuzamos mérőeszközök" modell útdiagramja:



A párhuzamos mérésnél feltételezzük, hogy a megfigyelt (manifeszt) változók valódi értéke és a hiba varianciája minden változónál azonos. Gyakran előfordul, hogy a megfigyelt változók ugyanazt az elméleti változót, kategóriát mérik, de éppen az eltérő mérőeszköz (annak eltérő minősége) miatt különböző lesz a hiba varianciája. Ha a mérésnél az utóbbi feltételt (a hiba varianciája állandó) nem tekintjük érvényesnek, akkor a mérőeszközöket „tau ekvivalens"-nek nevezzük.

Más esetekben a különböző mérőeszközöknél a valódi értékek nem azonosak, de közöttük lineáris összefüggés van, valamint a mérőeszközök varianciája is különbözik. Ebben az esetben a mérőeszközöket rokonnak, „congeneric"-nek nevezzük, és a következőképpen írhajuk fel:

$$\begin{aligned} x_1 &= \xi_1 + \delta_1 \\ x_2 &= \xi_2 + \delta_2 \\ \xi_2 &= \lambda_{21}\xi_1 , \end{aligned} \tag{25}$$

vagy

$$\begin{aligned} x_1 &= \xi_1 + \delta_1 \\ x_2 &= \lambda_{21}\xi_1 + \delta_2 , \end{aligned}$$

ahol

$$\begin{array}{llll} E(x_i) = 0 , & E(\delta_i) = 0 , & E(\xi_i) = 0 , & \forall i\text{-re}, \\ E(\xi\,\delta_i) = 0 , & & & \forall i\text{-re}, \\ E(\delta_i\delta_j) = 0 . & & & \end{array}$$

A „rokonsági" modell általánosabb mint az előző két modell, mivel:

a) a „tau–ekvivalens" modelleknél:

$$\lambda_{ij} = 1 \qquad \forall i, j\text{–re} ,$$

b) a „párhuzamos" mérési modellnél:

$$\lambda_{ij} = 1 \quad \text{és} \quad \Theta_{\delta_{ii}} = \Theta_{\delta_{jj}} \qquad \forall i, j\text{–re} .$$

A megbízhatósági együttható nem lesz egyenlő a „tau–ekvivalens" és a „rokonsági" tesztnél. A korrelációs mátrixból azonban becsülhetjük a λ^*-t, és így ebből a megbízhatósági együtthatót is megkaphajuk $\lambda_i^{*2} = \rho^2_{x_i\xi}$.

5.7.2 A többjellemzős–többmódszeres modell (Mutlitrait– multimethod model)

Az előző részben a klasszikus méréselmélet szerinti mérési modelleket tárgyaltuk. A következőkben a mérési modell megbízhatósága helyett a mérés érvényességének (validity) problémáját a CAMPBELL és FISKE (1959) által bevezetett többjellemzős–többmódszeres modell alapján elemezzük.

A többjellemzős–többmódszeres modell egyenletei és feltételei:

$$\begin{aligned} x_i &= \tau_i + \varepsilon_i \\ \tau_i &= \xi_j + \xi_k , \end{aligned} \tag{26}$$

ahol

τ_i: az i-edik megfigyelt változó valódi értéke,

ε_i: az i-edik hiba tag,

ξ_j: a j-edik elméleti változó,

ξ_k: a k-adik módszer hatását kifejező változó,

valamint:

$$\begin{aligned} &E(\varepsilon_i) = 0\,, && \\ &E(\xi_j) = 0\,, && \forall j\text{-re}, \\ &E(\xi_j \xi_k) = 0\,, && \forall j, k\text{-ra}. \end{aligned}$$

A fentiekből következik, hogy a valódi érték varianciája az elméleti változó és a módszer varianciájának összege:

$$\sigma_{\tau_i \tau_i} = \Phi_{jj} + \Phi_{kk}\,,$$

ahol

$\Phi_{jj} = E(\xi_j^2)$ az elméleti változó varianciája,

Φ_{kk} a módszer hatását képviseli.

Ha valamennyi változó standardizált, a fenti modellt a következőképpen írhatjuk fel:

$$x_i^* = \lambda_{ij}^* \xi_j^* + \lambda_{ik}^* \xi_k^* + \varepsilon_i^*\,, \tag{27}$$

ahol

$$\begin{aligned} &E(x_i^*) = E(\xi_i^*) = E(\varepsilon_i^*) = 0\,, && && \forall i\text{-re}, \\ &E(x_i^{*2}) = E(\xi_i^{*2}) = 1\,, && && \forall i\text{-re}, \\ &E(\varepsilon_i^* \xi_j^*) = 0\,, && \text{és } E(\varepsilon_i^* \varepsilon_j^*) = 0\,, && \forall i, j\text{-ra}. \end{aligned}$$

A fenti modellből következik:

$$\sigma_{x_i^* x_i^*} = \lambda_{ij}^* \lambda_{ij}^* + \lambda_{ik}^* \lambda_{ik}^* + \Theta_{\varepsilon_{ii}}^*\,,$$

vagyis a megfigyelt változók varianciája egyenlő az elméleti változó (ξ_i^*) varianciája, a módszer (ξ_k^*) varianciája és a hiba varianciája $(\Theta_{\varepsilon_{ii}}^*)$ összegével.

A megfigyelt változók kovarianciái:

$$\sigma_{x_i^* x_j^*}^* = \lambda_{ij}^* \Phi_{if}^* \lambda_{if}^* + \lambda_{ik}^* \Phi_{k\ell}^* \lambda_{i\ell}^*\,,$$

ami azt jelenti, hogy a megfigyelt változók kapcsolódásaira nemcsak az elméleti változó, hanem a módszer is hat.

5.7.3 A variancia–kovariancia komponens modell

A variancia–kovariancia modell a megfigyelt változók mintabeli megfigyelt értékeit a valódi értékek és a hiba összegeként állítja elő:

$$\underline{y} = \underline{\Lambda}_y \underline{\eta} + \underline{\varepsilon}\,, \tag{28}$$

ahol

$E(\underline{y}) = \underline{0}, \qquad E(\underline{\eta}) = \underline{0}, \quad E(\underline{\varepsilon}) = \underline{0},$

$E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0},$

$\underline{\Lambda}_y$ diagonális mátrix.

A valódi értékek azonban bizonyos látens változók függvényei:

$$\underline{\eta} = \underline{\Gamma}\,\underline{\xi}\,, \tag{29}$$

ahol

$$E(\underline{\xi}\,\underline{\varepsilon}') = \underline{0}\,, \qquad E(\underline{\xi}) = \underline{0}\,.$$

A modellben a $\underline{\Lambda}_y$ és $\underline{\Gamma}$ paraméter–mátrixok speciális felépítésűek. A $\underline{\Lambda}_y$ mátrix diagonális, a diagonális elemei akkor különböznek 1-től, amikor a megfigyelt változók mértékegységei különbözőek. A $\underline{\Lambda}$ mátrix (amit design mátrixnak nevezünk) elemei 0 és 1 értéket vehetnek fel.

5.7.4 A faktorelemzés

A faktorelemzésben a megfigyelt változók (x) valódi értékeit (τ) a látens változók függvényeként fejezzük ki. A látens változók közül vannak, amelyek több megfigyelt változó előállításában is szerepelnek, ezeket közös faktoroknak (ξ) nevezzük, és vannak olyanok, amelyek csak egy–egy változó reprodukálásában játszanak szerepet, ezeket egyedi faktoroknak (u) nevezzük.

A faktorelemzés matematikai modellje:

$$x_i = \tau_i + e_i\,, \tag{3}$$

és

$$\tau_i = \lambda_{i1}\xi_1 + \lambda_{i2}\xi_2 + \ldots + \lambda_{im}\xi_m + u_i\,,$$

ahol

e_i: az i-edik megfigyelt változó mérési hibája,

τ_i: az i-edik megfigyelt változó valódi értéke,

ξ_j: a j-edik közös faktor (látens változó),

u_i: az i-edik egyedi faktor,

λ_{ij}: a j-edik közös faktor regressziós együtthatója az i-edik megfigyelt változóra,

$\delta_i = u_i + e_i$: az i-edik mérési hiba és az i-edik egyedi faktor összege.

Feltételezzük, hogy:

$E(x_i) = 0\,,\quad E(\xi_j) = 0\,,\quad E(e_i) = 0\,,$

$E(u_i) = 0\,,\qquad \forall i, j\text{-re},$

valamint

$E(\xi_j u_i) = 0\,,\quad E(\xi_j e_i) = 0\,,\quad \forall i, j\text{-re}.$

A faktorelemzés modellje mátrixaritmetikai jelölésekkel:

$$\underline{x} = \underline{\Lambda}_x \underline{\xi} + \underline{\delta}\,, \tag{31}$$

ahol

$\underline{\delta} = \underline{u} + \underline{e}\,,$

$E(\underline{x}) = \underline{0}\,,\quad E(\underline{\xi}) = \underline{0}\,,\quad E(\underline{\delta}) = \underline{0}\,,\quad E(\underline{\xi}\,\underline{\delta}') = \underline{0}\,,$

$E(\underline{\delta}\,\underline{\delta}') = \underline{\Theta}_\delta$ diagonális.

A modellben feltételezzük, hogy (1) a mérési hiba változói korrelálatlanok egymással $E(e_i e_j) = 0$, (2) az egyedi faktorok korrelálatlanok egymással $E(u_i u_j) = 0$, (3) a közös faktorok korrelálatlanok $\underline{\delta}$-val. $E(\underline{\xi}\,\underline{\delta}') = \underline{0}$.

A modell paramétermátrixai: $\underline{\Lambda}_x$, $\underline{\Phi}$ és $\underline{\Theta}_\delta$.

A variancia–kovariancia mátrix a paraméterek függvényében:

$$\underline{\Sigma} = \underline{\Lambda}_x \underline{\Phi}\,\underline{\Lambda}_x' + \underline{\Theta}_\delta\,. \tag{32}$$

A gyakorlatban legtöbbször az exploratív elemzésekben feltételezzük, hogy $\underline{\Phi} = \underline{I}$, vagyis hogy a közös faktorok korrelálatlanok. Az exploratív faktorelemzés során általában a következő kérdéseket kell megválaszolnunk:

a) mennyi közös faktor szükséges a megfigyelt változók közötti korreláció magyarázatához,

2) a faktorsúlyok között melyek szignifikánsak és melyek nem,

3) mi a faktorok elméleti tartalma.

Az első kérdés könnyen megválaszolható, mivel a különböző faktorszámú modellek szignifikanciáját χ^2 próbával tesztelhetjük, és így megtalálhatjuk a faktoroknak azt a számát, amelyet tovább növelve a modell illeszkedése szignifikánsan tovább nem javítható.

A második kérdés megválaszolására ARCHER és JENNRICH (1973) a rotált faktorsúlyok standard hibájának számítására adott eljárást. A módszer azonban elég bonyolult, így a gyakorlati alkalmazása korlátozott. Ennél egyszerűbb eljárást javasol JÖRESKOG (1978), amit „legjobban illeszkedő egyszerű struktúrá"-nak nevezett el.

A harmadik kérdés megválaszolásához a szignifikáns faktorsúlyok mellett a szakmai ismeretek kapnak nagy szerepet. A konfirmatív faktorelemzés alapvetően abban különbözik az exploratív faktorelemzéstől, hogy az előbbinél a faktor–struktúrát *a priori*–elméleti vagy korábbi empirikus vizsgálatok alapján ismertnek tételezzük fel, míg az utóbbinál a faktor–struktúrát az elemzés során, a vizsgált adatokból származtatjuk.

5.7.5 *A másodrendű faktorelemzés*

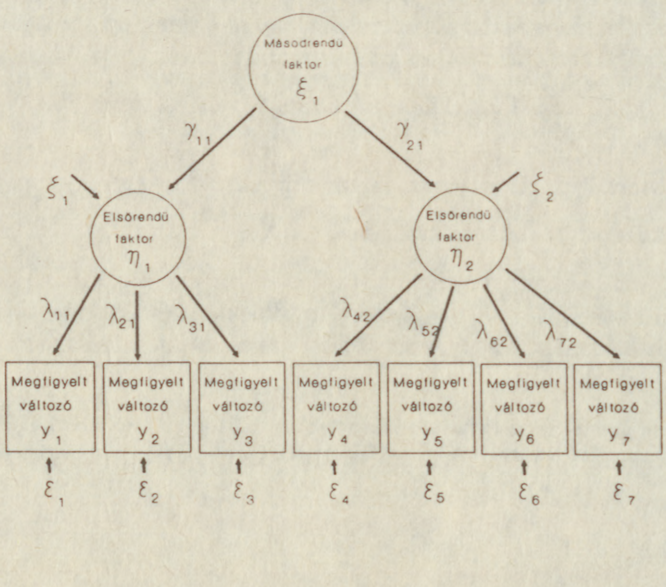
A másodrendű faktorelemzés hipotézise szerint a megfigyelt változók közös faktorai kifejezhetők további látens változók (faktorok) függvényeiként, amelyeket másodrendű faktoroknak nevezünk. A másodrendű faktorelemzés egyenletei:

$$\begin{aligned} \underline{y} &= \underline{\Lambda}_y \underline{\eta} + \underline{\varepsilon} \\ \underline{\eta} &= \underline{\Gamma}\, \underline{\xi} + \underline{\zeta}, \end{aligned} \tag{33}$$

ahol

$E(\underline{\eta}) = \underline{0}$, $E(\underline{\xi}) = \underline{0}$, $E(\underline{y}) = \underline{0}$,
$E(\underline{\zeta}) = \underline{0}$, $E(\underline{\varepsilon}) = \underline{0}$,
$E(\underline{\eta}\, \underline{\varepsilon}') = \underline{0}$,
$E(\underline{\xi}\, \underline{\zeta}') = \underline{0}$,
$E(\underline{\xi}\, \underline{\varepsilon}') = \underline{0}$,
$E(\underline{\zeta}\, \underline{\varepsilon}') = \underline{0}$.

Egy ilyen modellt mutat a következő ábra:



5.7.6 A regresszióelemzés

Az eddig tárgyalt speciális modellekben nem tettünk különbséget a megfigyelt változók között. A következő részben a megfigyelt változók között megkülönböztetjük a függő, okozati változókat (y) és a független, az ok szerepét betöltő változókat (x). A modell szerint előállítani, magyarázni vagy becsülni a függő változót akarjuk a független változók segítségével, amelyekről feltételezzük, hogy hatással vannak a függő változóra.

A független változók lehetnek rögzítettek, vagy véletlen hatást is tartalmazók.

A regressziós modell egyenlete:

$$y = \underline{\gamma}'\underline{x} + z\,, \tag{34}$$

ahol

$E(y) = 0\,, \quad E(\underline{x}) = \underline{0}\,, \quad E(z) = 0\,,$
$E(\underline{x}\,z) = \underline{0}\,.$

A modellben feltételezzük, hogy a reziduális (vagy véletlen) tag korrelálatlan az $\underline{x}$ vektorváltozóval.

A modell kovariancia–mátrixa:

$$\underline{\Sigma} = \begin{pmatrix} \underline{\gamma}'\underline{\Sigma}_{xx}\underline{\gamma} + \sigma_z^2 & \\ \underline{\Sigma}_{xx}\underline{\gamma} & \underline{\Sigma}_{xx} \end{pmatrix}. \tag{35}$$

Ha $\underline{x}$ értékei rögzítettek, akkor $\underline{\Sigma}_{xx}$ egyenlő $\underline{S}_{xx}$-vel, a kovariancia mátrix közvetlenül $\underline{x}$ rögzített értékeiből számítható.

Ha $\underline{x}$ valószínűségi változó, véletlen hatásokat is tartalmaz, $\underline{\Sigma}_{xx}$ becslése az $\underline{S}_{xx}$ (mintából számított) kovariancia mátrix lesz.

Mindkét esetben az együtthatókat ($\underline{\gamma}$) az $\underline{S}$ mátrix első sora alapján (az y és az $\underline{x}$ változók közötti kovarianciákból) becsüljük.

A reziduális tag varianciájának (σ_z^2) becslése $s_{yy} - \underline{\gamma}'\underline{S}_{xx}\underline{\gamma}$ lesz.

Ha egy vagy több független változó ($\underline{x}$) és a függő változó megfigyelt értékei mérési hibát is tartalmaznak, akkor a (34) egyenlet helyett a valódi regressziót keressük:

$$\eta = \underline{\gamma}'\underline{\xi} + \zeta\,, \tag{36}$$

ahol

ζ: a függő látens változó reziduális komponense, amiről feltételezzük, hogy korrelálatlan $\underline{\xi}$-vel:

$E(\underline{\xi}\,\zeta) = \underline{0}\,.$

Az η és $\underline{\xi}$ a függő és független változók valódi értékei, vagy más szóval közös faktorai:

$$y = \lambda\eta + \varepsilon\,, \tag{37}$$

és

$$\underline{x} = \underline{\Lambda}\,\underline{\xi} + \underline{\delta}\,, \tag{38}$$

ahol feltételezzük, hogy

$$E(y) = 0\,, \quad E(\eta) = 0\,, \quad E(\varepsilon) = 0\,, \quad E(\eta\,\varepsilon) = 0\,,$$

és

$$E(\underline{x}) = \underline{0}\,, \quad E(\underline{\xi}) = \underline{0}\,, \quad E(\underline{\delta}) = \underline{0}\,, \quad E(\underline{\xi}\,\underline{\delta}') = \underline{0}\,.$$

A modell kovariancia mátrixa:

$$\underline{\Sigma} = \begin{pmatrix} \lambda(\underline{\gamma}'\underline{\Phi}\,\underline{\gamma} + \Psi^2)\lambda + \Theta_\varepsilon & \\ \underline{\Lambda}\,\underline{\Phi}\,\underline{\gamma}\,\lambda & \underline{\Lambda}\,\underline{\Phi}\,\underline{\Lambda}' + \underline{\Theta}_\delta \end{pmatrix},$$

ahol

$$\underline{\Phi} = \operatorname{cov}(\underline{\xi})\,, \qquad \Psi^2 = \operatorname{cov}(\varsigma)\,,$$

$$\Theta_\varepsilon = \operatorname{cov}(\varepsilon)\,, \qquad \underline{\Theta}_\delta = \operatorname{cov}(\underline{\delta})\,.$$

5.7.7 Az útelemzés

Az útelemzés modelljében az ok szerepét játszó változók közvetlen hatásait vizsgáljuk az okozat szerepét betöltő változókra. Ezeket a közvetlen hatásokat lineáris struktúrális egyenletekkel fejezzük ki, és a feladat a hipotetikusan felállított rekurzív modell paramétereinek a becslése.

Tegyük fel például, hogy két változót két időpontban figyeltünk meg, és a két változó ugyanazon látens változó függvénye. Az y_1 és y_2 az η_1 látens változót méri az első időpontban, az y_3 és y_4 változók pedig az η_2 látens változót mérik a második időpontban. A mérési egyenletek ekkor:

$$\begin{aligned} y_1 &= \eta_1 + \varepsilon_1 \\ y_2 &= \lambda_1\eta_1 + \varepsilon_2 \\ y_3 &= \eta_2 + \varepsilon_3 \\ y_4 &= \lambda_2\eta_2 + \varepsilon_4\,. \end{aligned}$$

A fő kérdés az η látens változó időbeli stabilitása, amit a következő regressziós egyenlet fejez ki:

$$\eta_2 = \beta\,\eta_1 + \varsigma\,.$$

Ha $\underline{\Phi}$ az (η_1, η_2) látens változók kovariancia mátrixa, a $\underline{\Theta}$ a $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ mérési hibák kovariancia mátrixa (amelyről feltételezzük, hogy diagonális), akkor a megfigyelt változók kovariancia mátrixa:

$$\underline{\Sigma} = \begin{pmatrix} \Phi_{11} + \Theta_{11} & & & \\ \lambda_1 \Phi_{11} & \lambda_1^2 \Phi_{11} + \Theta_{22} & & \\ \Phi_{21} & \lambda_1 \Phi_{21} & \Phi_{22} + \Theta_{33} & \\ \lambda_2 \Phi_{21} & \lambda_1 \lambda_2 \Phi_{21} & \lambda_2 \Phi_{22} & \lambda_2^2 \Phi_{22} + \Theta_{44} \end{pmatrix}.$$

5.7.8 *A MIMIC modell (Multiple Indicators Multiple Causes Model)*

A MIMIC modell két egyenletből áll. A strukturális egyenletből, amelyben az ok szerepét megfigyelt változók ($\underline{x}$), az okozat szerepét pedig látens változók (η) töltik be:

$$\underline{\beta}\,\underline{\eta} = \underline{\Gamma}\,\underline{x} + \underline{\varsigma}, \tag{39}$$

ahol

$$E(\underline{\eta}) = \underline{0}, \qquad E(\underline{x}) = \underline{0}, \qquad E(\underline{\varsigma}) = \underline{0},$$
$$E(\underline{x}\,\underline{\varsigma}') = \underline{0},$$

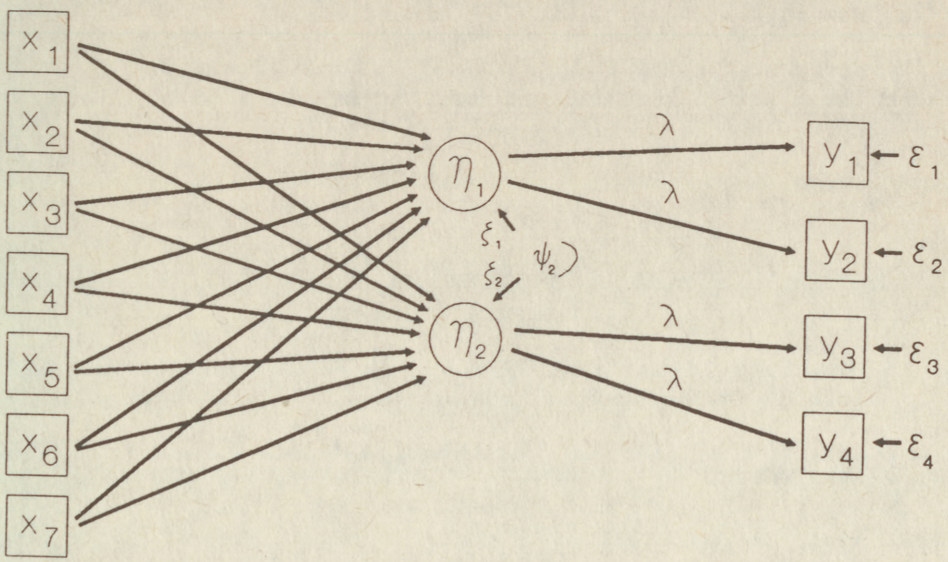
és a mérési hibát tartalmazó egyenletből:

$$\underline{y} = \underline{\Lambda}\,\underline{\eta} + \underline{\varepsilon}, \tag{40}$$

ahol

$$E(\underline{y}) = \underline{0}, \qquad E(\underline{\varepsilon}) = \underline{0}, \qquad E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0},$$
$$E(\underline{x}\,\underline{\varepsilon}') = \underline{0}, \qquad E(\underline{\varepsilon}\,\underline{\varsigma}') = \underline{0}.$$

A MIMIC modell diagramja:



5.8 Szimultán elemzés több csoportban

Tekintsük egy vizsgált sokaság g számú csoportját. Ezek a csoportok lehetnek nemzetek egy nemzetközi vizsgálatban, lehetnek kultúrálisan vagy szocioökonomikusan elkülönűlő csoportok egy nemzeti vizsgálatban, vagy bármely, megfigyelési egységnek bizonyos változók alapján elkülönített csoportjai.

Feltételezzük, hogy minden csoportban a vizsgálat változóit véletlen mintán mértük. Feltételezzük továbbá, hogy minden csoportban a látens változók strukturális modelljét, és a megfigyelt változók mérési modelljeit az (1), (2) és (3) egyenletek formájában adjuk meg, vagyis:

$$\underline{B}^{(k)}\underline{\eta} = \underline{\Gamma}^{(k)}\underline{\xi} + \underline{\zeta}^{(k)}\,, \tag{41}$$

$$\underline{y} = \underline{\Lambda}_y^{(k)}\underline{\eta} + \underline{\varepsilon}^{(k)}\,, \tag{42}$$

$$\underline{x} = \underline{\Lambda}_x^{(k)}\underline{\xi} + \underline{\delta}^{(k)}\,, \tag{43}$$

formában, ahol k a csoportot jelölő index

$$k = 1, 2, \ldots, g\,.$$

A k-adik csoportban a modellt a nyolc paramétermátrix megadásával definiáljuk $(\underline{\Lambda}_y^{(k)}, \underline{\Lambda}_x^{(k)}, \underline{B}^{(k)}, \underline{\Gamma}^{(k)}, \underline{\Phi}^{(k)}, \underline{\Psi}^{(k)}, \underline{\Theta}_\varepsilon^{(k)}, \underline{\Theta}_\delta^{(k)})$. A paramétermátrixok elemei lehetnek rögzítettek (fixed), szabadok (free), vagy kötöttek (constrained). Ha a csoportok között a paraméterek egyezőségére nem írunk elő feltételt, akkor a csoportokat egymástól függetlenül elemezhetjük. Azonban ha a csoportok paraméterei között feltételezünk egyezőséget, akkor a csoportokat csak szimultán elemezhetjük, csak így kaphatunk efficiens becsléseket.

Ha például a megfigyelt változók mérési tulajdonságai azonosak a különböző csoportokban, akkor a következő paraméterek egyezőségét tételezhetjük fel:

$$\underline{\Lambda}_y^{(1)} = \underline{\Lambda}_y^{(2)} = \ldots = \underline{\Lambda}_y^{(g)}$$
$$\underline{\Lambda}_x^{(1)} = \underline{\Lambda}_x^{(2)} = \ldots = \underline{\Lambda}_x^{(g)}$$
$$\underline{\Theta}_\varepsilon^{(1)} = \underline{\Theta}_\varepsilon^{(2)} = \ldots = \underline{\Theta}_\varepsilon^{(g)}$$
$$\underline{\Theta}_\delta^{(1)} = \underline{\Theta}_\delta^{(2)} = \ldots = \underline{\Theta}_\delta^{(g)}\,.$$

Ekkor a csoportok a látens változók eloszlásában külöbözhetnek egymástól.

A strukturális kapcsolódások invarianciáját tesztelhetjük a következő azonosságok feltételezésével:

$$\underline{B}^{(1)} = \underline{B}^{(2)} = \ldots = \underline{B}^{(g)}$$
$$\underline{\Gamma}^{(1)} = \underline{\Gamma}^{(2)} = \ldots = \underline{\Gamma}^{(g)}\,.$$

A csoportok függetlenségétől a paraméterek teljes azonosságáig terjedően az invariancia bármely fokát tesztelhetjük. A szimultán modellt a következő függvény minimalizálásával becsüljük:

$$F = \sum_{k=1}^{g} (N_k/N)[\log|\underline{\Sigma}^{(k)}| + \mathrm{tr}(\underline{S}^{(k)} \cdot \underline{\Sigma}^{(k)-1}) - \log|\underline{S}^{(k)}| - (p+q)], \tag{44}$$

ahol

N_k: a megfigyelések száma a k-adik csoportban,

$N = N_1 + N_2 + \ldots + N_g$.

Ha a megfigyelt változók együttes eloszlása normális, az F függvény a likelihood–függvény logaritmusának minusz $(N/2)$-szeresével lesz egyenlő. A modell illeszkedését a χ^2 próbával végezzük hasonlóan ahhoz, amikor csak egy csoportot vizsgáltunk.

A szabadságfok:

$$d = (1/2)g(p+q)(p+q+1) - t,$$

ahol

t a független paraméterek száma az összes csoportban.

5.9 Példa a LISREL modellre

A következő példában azt vizsgáltuk, hogy az emberek társadalmi háttere, társadalmi státusza mennyire befolyásolja az emberek társadalmi tudatát, azt ahogyan vélekednek a társadalom különböző dolgairól (a példában két társadalmi problémát ragadtunk ki, az egyik a társadalmi intézményekbe vetett bizalom, és az életvitelváltozásról alkotott vélemény), és hogyan módosítja, erősíti, csökkenti ezt a hatást az emberek értékrendje, világnézete. Azt a hipotézist vizsgáltuk, hogy az emberek értékrendszere szűrő szerepet játszik–e a társadalomban elfoglalt hely és a szociális tudat között. A modellezés lépéseit most nem követjük végig, csupán a végső modellt prezentáljuk; méghozzá az Életvitelváltozás és Bizalom endogén változó–blokkot tartalmazó modelleket. (A vizsgált adathalmaz megegyezik a 2. fejezetben elemzett adathalmazzal.)

Hogy a két modell ábráit olvasni tudjuk, a Gyermeknevelés, Tízparancsolat, Vallás blokkok indikátor változóinak az értelmezéséhez — amelyek a manifeszt változók közös faktorai — közöljük a megfelelő faktorstruktúrákat.

Gyermeknevelés

A 17 gyermeknevelési elv három közös faktorát az exploratív faktorelemzés fejezetében már vizsgáltuk, most csak a könnyebb olvasat miatt ismételjük meg őket.

Az első rotálatlan faktor, a legfőbb értékpolaritást fejezi ki:

Modern személyiségértékek ⟷ Hagyományos közösségi értékek

Ezt az alapvető értékdimenziót figyeltünk meg a Roeach– értékteszt elemzése során 1980 és 1982-ben is (Intellektuális, autonómia értékei álltak szemben a hagyományos közösségi és öröm–értékekkel). Lásd Hankiss, Manchin, Füstös, Szakolczai: Kényszerpályán. MTA Szociológiai Intézet, 1983. Ezen a dimenzión a „belülről irányított ember" típusa áll szemben a „tradícióktól irányított" ember tulajdonságaival.

A második rotálatlan faktor:

Munka, biztonság, hit ⟷ Emberi kapcsolatok, autonómia

A pozitív oldalon kétfajta értékek keverednek; a munkával kapcsolatos és a vallásos értékek, ezen belül az elsőknek van nagyobb, meghatározóbb súlya. Ebben az összefüggésben a vallásos hit a kötelesség, felelősség, becsület fogalmaihoz kapcsolódik. A munka értéke bizonyos fajta biztonságra való törekvéssel fonódik össze: anyagi szempontból a takarékossággal, lelki szempontból a vallásos hittel és a hűséggel. A másik, negatív oldalon az emberi kapcsolatokat szabályozó tulajdonságok szerepelnek. Olyan értékek, amelyek a személyes autonómia fenntartása mellett a másik ember számára is könnyebbé, zökkenőmentesebbé teszik az együttélést.

A harmadik rotálatlan faktor dichotómiája:

Hagyományos keresztény értékek ⟷ Protestáns értékek

A pozitív póluson a hagyományos keresztény értékek szerepelnek, középpontban mások tisztelete, az alkalmazkodás értékei, mindössze a határozottság keveredik közéjük. Ennek a dimenziónak az ellentétes oldala az evilági élet, ahol az anyagi, protestáns értékek a fontosak.

Tízparancsolat

A Tízparancsolat látens változó manifeszt változói és a faktorelemzés eredményei, a következők:

Kérem, mondja meg mindegyik parancsolatról, hogy mennyire érvényes Önre: teljesen, bizonyos mértékig vagy nem érvényes!

1. Én vagyok a te Istened, ne legyenek néked idegen isteneid előttem!
2. A te Istenednek a nevét hiába fel ne vedd!
3. Megemlékezzél az Úr napjáról, hogy megszenteljed azt!
4. Tiszteld atyádat és anyádat!
5. Ne ölj!
6. Ne paráználkodj!
7. Ne lopj!
8. Ne tégy felebarátod ellen hamis tanubizonyságot!
9. Ne kívánd a te felebarátod feleségét!
10. Se házát, se mezejét, se másféle jószágát!

Életvitelváltozás

A kérdés, amit a vizsgált személyeknek feltettünk a következő volt:

Felsorolok Önnek olyan életvitelváltozásokat, amelyek a közeli jövőben bekövetkezhetnek. Kérem mondja meg mindegyikről, hogy jó, rossz dolognak tartaná vagy nem törődne vele!

1. visszaszorulna a pénz és anyagi javak fontossága
2. csökkenne a munka fontossága életükben
3. a technológia fejlesztése nagyobb hangsúlyt kapna
4. fontosabbá válna az emberi személyiség fejlesztése
5. növekedne a tekintélytisztelet
6. nagyobb hangsúlyt kapna a családi élet
7. egyszerűbb és természetesebb lenne az életvitel

Az első faktor pozitív oldalán a *belső boldogságra törekvés*, a családi egyensúly, a személyiség fejlesztése és a tekintély növekedése, a negatív — bár gyenge oldalon — a *külső boldogságra* törekvés, a pénz, a technológia, az anyagi, hatalmi megelégedettség változói jelennek meg. Az első faktorban megjelenik — a máshol is tapasztalt — belső kontra külső irányultság dichotómiája.

A második faktor lényegében egy posztmodern vágyódás faktorának tekinthető, az egyszerűbb életvitel, a személyiség fejlesztésének a pozitív súlya áll szemben a materiális változókkal.

Intézményekbe vetett bizalom

Felsorolok néhány intézményt. Kérem mondja meg mindegyikről, hogy menynyire van bizalma benne!

1. az egyházban
2. az oktatási rendszerben
3. a jogrendszerben
4. a sajtóban, televízióban, rádióban
5. a szakszervezetben

6. a vállalatokban
7. az országgyűlésben
8. a tanácsokban

Az intézményekbe vetett bizalom első faktora főfaktornak tekinthető, azzal a lényeges megjegyzéssel, hogy a politikai intézményrendszer elemeihez képest az egyház intézménye nem bipolarizálódik, hanem, egy az első civil társadalom intézményeitől független, második faktorban jelenik meg.

Tízparancsolat
első rotált faktor
Magyarország

(sajátérték = 4.62, 46.2 %)

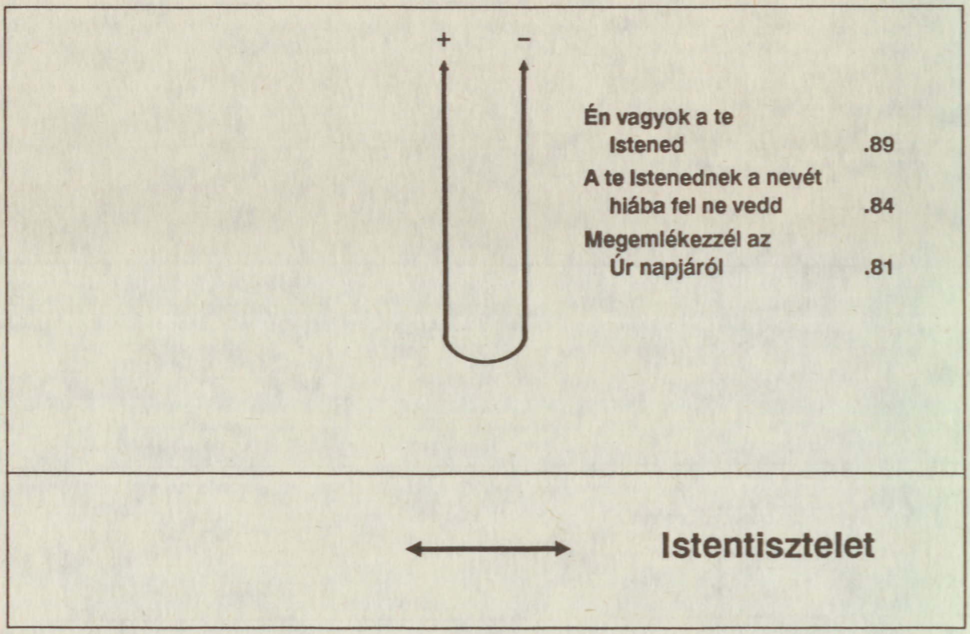
Ne lopj	.87
Ne tégy hamis tanúbizonyságot	.86
Ne ölj	.82
Ne kívánd más házát	.81
Ne kívánd más feleségét	.74
Tiszteld atyádat	.68
Ne paráználkodj	.58

+ –

Magatartás-értékek

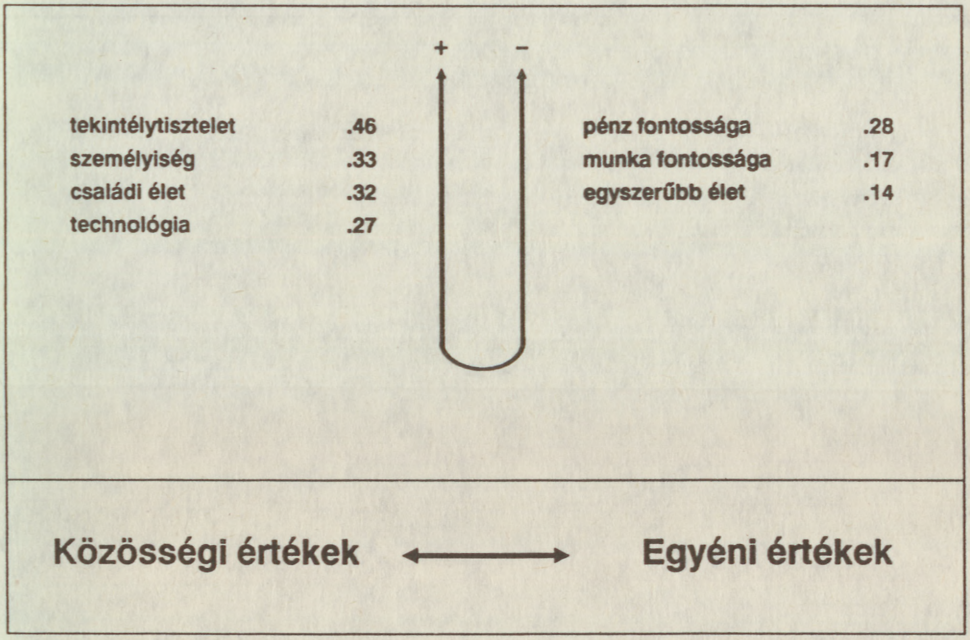
Tízparancsolat
második rotált faktor
Magyarország

(sajátérték = 1.92, 19.3 %)



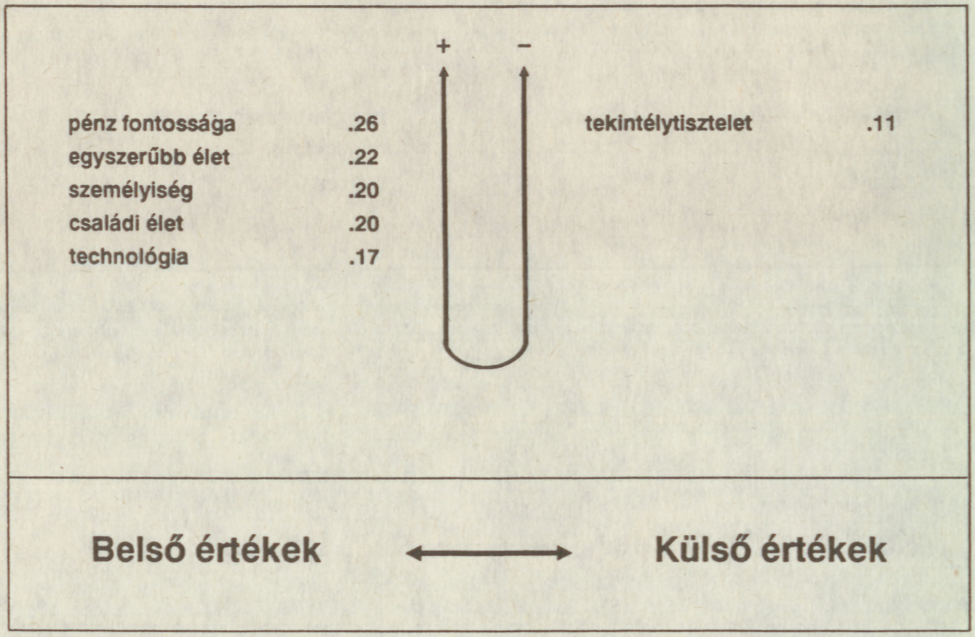
Életvitelváltozás első, rotálatlan faktora

(sajátérték = 1.46, 20.9%)



Életvitelváltozás második, rotálatlan faktora

(sajátérték = 1.25, 17.9%)



Intézményekbe vetett bizalom első, rotálatlan faktora

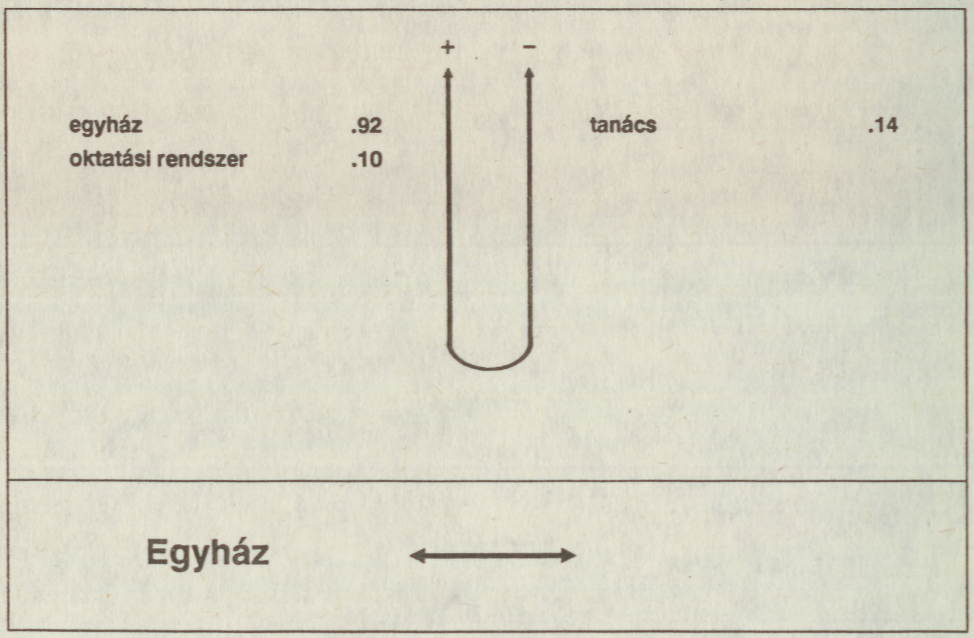
(sajátérték = 2.50, 35.7%)

országgyűlés	.71
tanács	.70
szakszervezet	.68
sajtó, TV	.64
jogrendszer	.56
oktatási rendszer	.52

Civil társadalom intézményei

Intézményekbe vetett bizalom második, rotálatlan faktora

(sajátérték = .89, 12.7%)



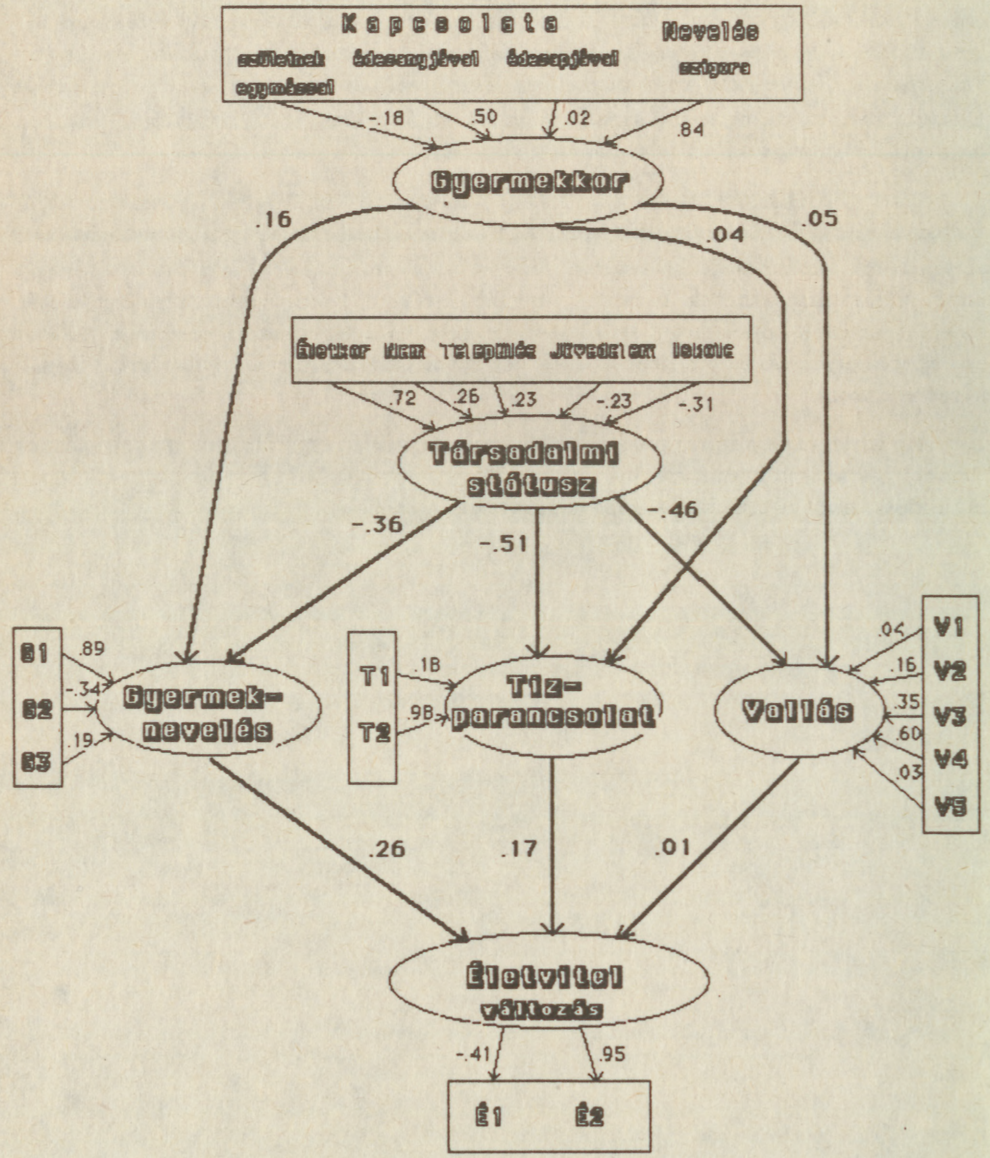
A komplex modell vizsgálata

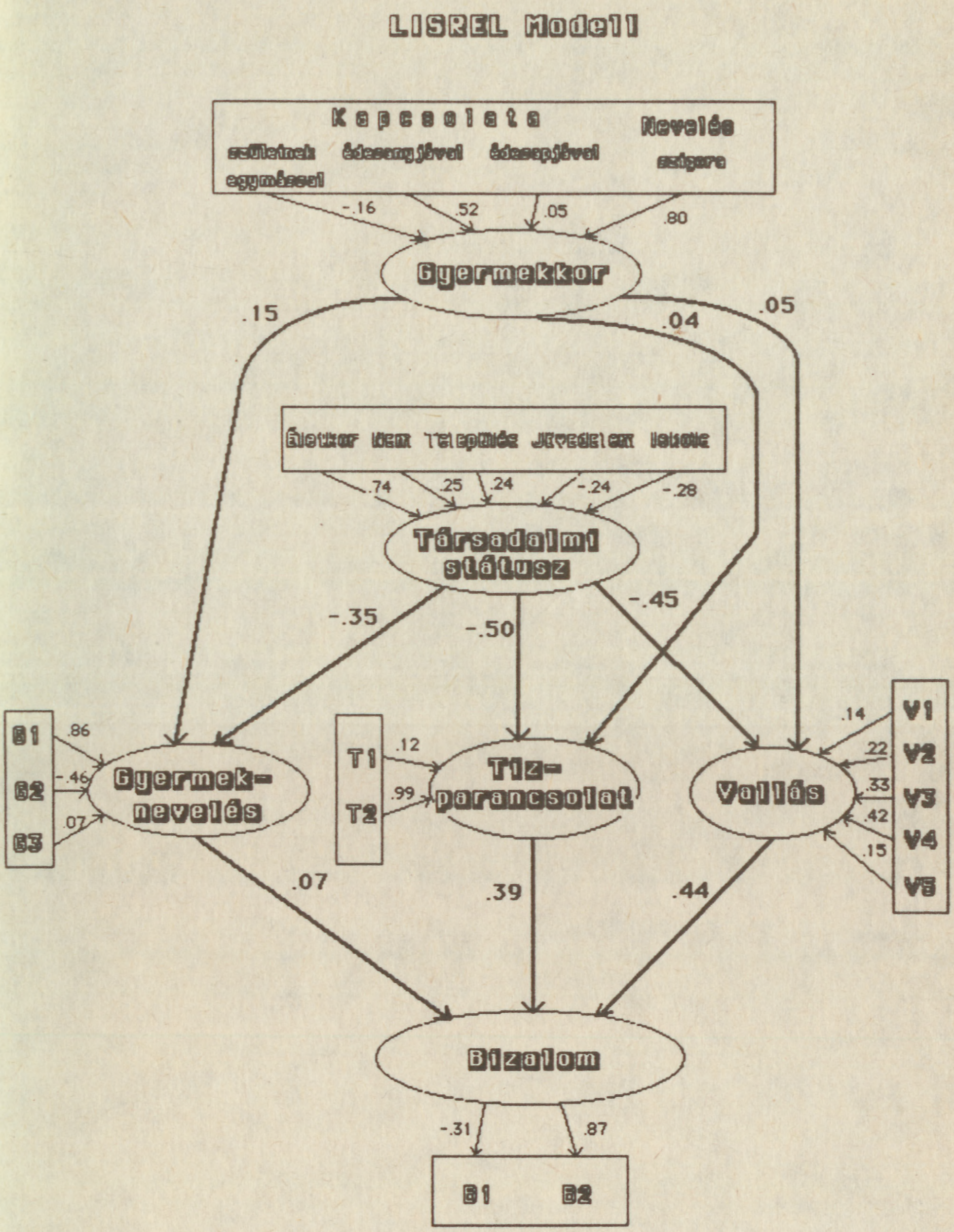
Itt és most csak a végső, statisztikailag illeszkedő (.01 elsőfajú hibával) modellt ismertetjük, és nem térünk ki a modellváltozatokra. A vizsgálat során azt a determinisztikus modellt elemeztük, hogy a multbeli társadalmi pozíció és a jelenlegi pozíció hogyan befolyásolja az emberek értékrendjét, és ez az értékrend milyen hatással van az emberek nézetrendszerére, a társadalmi problémákra adott válaszaikra. Ehelyütt csupán két endogén változócsoportot emeltünk ki, az Életvitelváltozásra, és az értékrendnek az Intézményekbe vetett bizalomra gyakorolt befolyásoló és/vagy szűrő szerepét modelleztük.

A két LISREL modell ábráján látható útegyütthatók alapján az exogén látens változók szerepét, hatását érthetjük meg. A multbeli társadalmi pozíció hatását az emberek értékrendje jelentősen felerősítette amikor az Életvitelváltozásokra adott válaszokat akartuk becsülni, míg az Intézményekbe vetett bizalom esetén az értékrendnek éppen hogy szűrő szerepe volt, és a Gyerekkori pozíció a vallásos és hagyományos hit és erkölcsi elveken keresztül vezető úton befolyásolta a vizsgált nézetrendszert.

A jelenlegi társadalmi pozíció hatása mindegyik értékrendi blokkra jelentősnek ítélhető, és az értékrend erősítő, fokozó, vagy éppen szűrő szerepe hasonlóképpen működött, mint a multbeli társadalmi pozíció esetében.

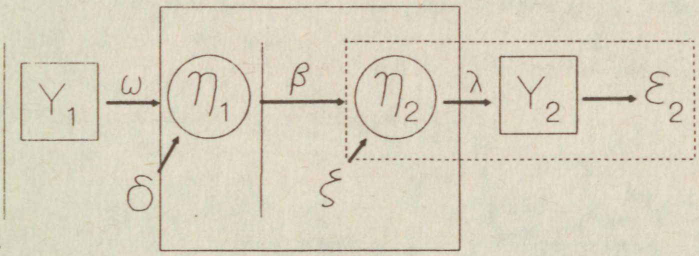
LISREL Modell




LISREL Modell
Kapcsolata
szüleinek egymással
édesanyjával
édesapjával
Nevelés
-.16
.52
.05
.80
Gyermekkor
.15
.04
.05
.74
.25
.24
-.24
-.28
Társadalmi státusz
-.35
-.50
-.45
G1
G2
G3
.86
-.46
.07
Gyermek-nevelés
T1
T2
.12
.99
Tíz-parancsolat
V1
V2
V3
V4
V5
.14
.22
.33
.42
.15
Vallás
.07
.39
.44
Bizalom
-.31
.87
B1
B2

6. Az LVPLS modell (Latent Variables Path Analysis with Partial Least–Squares Estimation)

A látens változók útmodelljének szkémája:



ahol

Y_1, Y_2: a megfigyelt, manifeszt változók halmaza,

η_1, η_2: nem megfigyelt, látens változók halmaza,

β: útegyüttható(k),

λ: az endogén manifeszt változók faktorsúlya(i),

ω: az exogén manifeszt változók regressziós sulya(i),

ζ: a látens endogén változó sztochasztikus reziduális tagja(i),

δ: az exogén látens változó reziduális tagja(i),

ε: az endogén manifeszt változó mérési hibája(i).

A modell három egyenletből áll.

1. Az első a látens változók közötti utakat írja le, ez a modell strukturális egyenlete:

$$\underline{\eta} = \underline{B}\,\underline{\eta} + \underline{\zeta}, \tag{1}$$

ahol

$$E(\underline{\eta}) = \underline{0}, \quad E(\underline{\zeta}) = \underline{0}, \quad E(\underline{\eta}\,\underline{\zeta}') = \underline{0}.$$

Az i-edik látens változót előállító strukturális egyenlet:

$$\eta_i = \sum_{j<i} (\beta_{ij}\eta_j) + \zeta_i . \tag{2}$$

Feltételezzük, hogy az endogén látens változó (η_i) az ok szerepét betöltő látens változók feltételes várható értéke:

$$E(\eta_i|\underline{\eta}_j) = \sum_{j<i} (\beta_{ij}\eta_j) . \tag{3}$$

2. A második egyenlet a manifeszt változók mérési modellje:

$$\underline{y} = \underline{\Lambda}\,\underline{\eta} + \underline{\varepsilon}, \tag{4}$$

ahol

$\underline{\Lambda}$: általános eleme λ_{kj} a j-edik látens változó (η_j) regressziós együtthatója a k-adik megfigyelt változó (y_k) előállításában, vagy másképpen a k-adik megfigyelt változónak a j-edik látens változóra vonatkozó faktorsúlya,

$\underline{\varepsilon}$: a megfigyelt változók mérési hibáit, a megfigyelt változók reziduális tagjait tartalmazza.

Feltételezzük, hogy

$$E(\underline{y}) = \underline{0}, \quad E(\underline{\eta}) = \underline{0}, \quad E(\underline{\varepsilon}) = \underline{0}, \quad E(\underline{\eta}\,\underline{\varepsilon}') = \underline{0}.$$

Az utolsó feltétel azt jelenti, hogy a mérési hibák korrelálatlanok a látens változókkal.

Az egyenlet mérési hiba nélküli része feltételezés szerint a megfigyelt változók látens változóra vonatkozó feltételes várható értékét adja. A k-adik megfigyelt változó feltételes várható értéke:

$$E(y_k|\underline{\eta}) = \Sigma(\lambda_{kj}\eta_j). \tag{5}$$

3. A harmadik egyenlet a súlymodell. Ebben a látens változókat a megfigyelt változók regressziós függvényeként állítjuk elő. A súlymodell egyenlete a következő:

$$\underline{\eta} = \underline{\Omega}\,\underline{y} + \underline{\delta}, \tag{6}$$

ahol

$\underline{\Omega}$: a súlyegyütthatókat tartalmazó mátrix,

$\underline{\delta}$: a látens változók reziduális tagjait tartalmazza.

A modellben feltételezzük, hogy

$$E(\underline{\eta}) = \underline{0}, \quad E(\underline{y}) = \underline{0}, \quad E(\underline{\delta}) = \underline{0}, \quad E(\underline{y}\,\underline{\delta}') = \underline{0}.$$

A j-edik látens változó feltételes várható értéke:

$$E(\eta_j|\underline{y}) = \Sigma(\omega_{jk}y_k). \tag{7}$$

6.1 A strukturális egyenlet redukált formája

A strukturális egyenlet a látens változók kapcsolódásait írja le:

$$\underline{\eta} = \underline{B}\,\underline{\eta} + \underline{\zeta}. \tag{8}$$

A $\underline{B}$ mátrix látens változók egymásra gyakorolt közvetlen hatásait, a közvetlen kapcsolódások útegyütthatóit tartalmazza. Ha ezt az egyenletet átrendezzük úgy, hogy a bal oldalra kerüljenek a látens változók, akkor a strukturális egyenlet redukált formájához jutunk:

$$\underline{\eta} = (I - \underline{B})^{-1}\underline{\varsigma} = \underline{B}^*\underline{\varsigma}. \tag{9}$$

A redukált forma $\underline{B}^*$ mátrixa a látens változók egymásra gyakorolt teljes hatásait tartalmazza. A teljes és a közvetlen hatások különbsége a közvetett hatást mutatja, azt a hatást, amelyet a direkt utak mellett az indirekt utakon az egyes látens változók más változókra kifejtenek:

$$\underline{B}^+ = \underline{B}^* - \underline{B}. \tag{10}$$

Ehhez hasonlóan kifejezhetjük a látens változónak a megfigyelt változókra vonatkozó teljes és közvetett hatásait is:

$$\underline{y} = \underline{\Lambda}\underline{\eta} + \underline{\varepsilon} = \underline{\Lambda}\,\underline{B}^*\underline{\varsigma} + \underline{\varepsilon} = \underline{\Lambda}^*\underline{\varsigma} + \underline{\varepsilon}, \tag{11}$$

ahol

$\underline{\Lambda}$: a közvetlen hatásokat,

$\underline{\Lambda}^*$: a teljes hatásokat tartalmazza.

Így

$$\underline{\Lambda}^+ = \underline{\Lambda}^* - \underline{\Lambda} \tag{12}$$

a látens változók manifeszt változókra vonatkozó közvetett hatásait tartalmazza, amelyek egyenlőek nullával az exogén változók esetén.

6.2 A modellben szereplő változók és paraméterek

$\underline{y}' = [y_1, y_2, \ldots, y_p]$: a megfigyelt, manifeszt változók p elemű vektora,

$\underline{\eta}' = [\eta_1, \eta_2, \ldots, \eta_m]$: a látens változók m elemű vektora,

$\underline{\varsigma}' = [\varsigma_1, \varsigma_2, \ldots, \varsigma_m]$: a látens változók reziduális komponense (sztochasztikus reziduális tag),

$\underline{\varepsilon}' = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p]$: a manifeszt változók reziduális komponense (mérési hiba),

$\underline{\delta}' = [\delta_1, \delta_2, \ldots, \delta_m]$: a látens változók reziduális komponense (a súlyegyenletek reziduális tagja),

$\underline{B}$: a látens változók útegyütthatóinak mátrixa, $(m \times m)$ típusú, általános eleme β_{ij} a j-edik látens változónak az i-edik látens változóra kifejtett közvetlen hatását mutatja,

$\underline{\Lambda}$: a megfigyelt (manifeszt) változók faktorsúly mátrixa, $(m \times p)$ típusú, általános eleme λ_{kj} a j-edik látens változó közvetlen hatását fejezi ki a k-adik megfigyelt változóra,

$\underline{\Omega}$: a látens változók súlyegyütthatói (regressziós együtthatók), $(m \times p)$ típusú, általános eleme ω_{jk} a k-adik megfigyelt változó közvetlen hatását fejezi ki a j-edik látens változóra,

$\underline{\Psi}$: a látens változók sztochasztikus reziduális tagjai között számított variancia–kovariancia mátrix, $(m \times m)$ típusú, általános eleme ψ_{ij} az i-edik és j-edik reziduális komponens kovarianciája $(E(\varsigma_i\, \varsigma_j)$ és ψ_{ii} jelöli az i-edik reziduális komponens varianciáját $(E(\varsigma_i\, \varsigma_i))$,

$\underline{\Theta}$: a megfigyelt változók mérési hibáinak variancia–kovariancia mátrixa, $(p \times p)$ típusú, általános eleme $\theta_{k\ell}$ a k-adik és ℓ-edik megfigyelt változó mérési hibája között számított kovarianciát, a θ_{kk} a k-adik mérési hiba (reziduális tag) varianciáját jelöli.

6.3 A modellben szereplő változók variancia–kovariancia mátrixai

Mivel a megfigyelt változókat a mérési modellben a látens változók függvényeiként fejeztük ki, nézzük először a látens változók kovariancia mátrixát:

$$\underline{C} = E(\underline{\eta}\,\underline{\eta}') = \mathrm{cov}(\underline{\eta}) = (\underline{I} - \underline{B})^{-1}\underline{\Psi}\,(\underline{I} - \underline{B}')^{-1} = \underline{B}^{*}\underline{\Psi}\,\underline{B}^{*\prime}\,, \tag{13}$$

ahol

$\underline{\Psi}$: a sztochasztikus reziduális tag variancia–kovariancia mátrixa.

A paraméter becslésénél feltételezzük, hogy a látens változók standardizáltak, így a kovariancia egyenlő lesz a korrelációval. A fentieket felhasználva a megfigyelt változók variancia–kovariancia mátrixa:

$$\underline{\Sigma} = \mathrm{cov}(\underline{y}) = \underline{\Lambda}\,\underline{C}\,\underline{\Lambda}' + \underline{\Theta}_{\varepsilon}\,. \tag{14}$$

A megfigyelt változók variancia–kovariancia mátrixa a $\underline{\Lambda}$, $\underline{C}$ és $\underline{\Theta}_{\varepsilon}$ paraméterek függvénye.

Számíthatjuk még a megfigyelt változók és a látens változók közötti páronkénti kovarianciákat (standardizált változók esetén korrelációkat), a látens változók struktúráját:

$$\underline{T} = \mathrm{cov}(\underline{y}\,\underline{\eta}') = \underline{\Lambda}\,\underline{C} = \underline{\Lambda}(\underline{I} - \underline{B})^{-1}\underline{\Psi}\,(\underline{I} - \underline{B}')^{-1}\,. \tag{15}$$

A sztochasztikus reziduális tag (ς) variancia–kovariancia mátrixa:

$$\underline{\Psi} = (\underline{I} - \underline{B})\underline{C}(\underline{I} - \underline{B}')\,. \tag{16}$$

A mérési hiba variancia–kovariancia mátrixa:

$$\underline{\Theta}_\varepsilon = \underline{S} + \underline{\Lambda}\,\underline{C}\,\underline{\Lambda}' - \underline{\Lambda}\,\underline{T}' - \underline{T}\,\underline{\Lambda}' . \tag{17}$$

A mérési hiba és a megfigyelt változók variancia–kovariancia mátrixa:

$$\operatorname{cov}(\underline{\varepsilon}\,\underline{y}) = \underline{T} - \underline{\Lambda}\,\underline{C} . \tag{18}$$

A mérési hiba és a sztochasztikus reziduális tag variancia–kovariancia mátrixa:

$$\operatorname{cov}(\underline{\varepsilon}\,\underline{\rho}) = (\underline{T} - \underline{\Lambda}\,\underline{C})(\underline{I} - \underline{B}') . \tag{19}$$

A látens változó ($\hat{\eta}$) magyarázott részének ($\underline{B}\hat{\underline{\eta}}$) kovarianciái:

$$\underline{C}^* = \operatorname{cov}(\underline{B}\,\hat{\underline{\eta}}) = \underline{B} \cdot \underline{C} \cdot \underline{B}' . \tag{20}$$

A $\underline{C}^*$ mátrix diagonálisában az endogén látens változók többszörös korrelációi négyzetét tartalmazza, ami a látens változók varianciáinak az a része, amit a becslő látens változók magyarázni tudnak.

A $\underline{C}^*$ diagonális elemei:

$$\underline{R}^2 = \underline{I} \times (\underline{B}\,\underline{C}\,\underline{B}') , \tag{21}$$

ahol

$\times$ az ún. Hadamard–szorzatot jelöli[*]

A megfigyelt változók mérési modellje a megfigyelt változókat két részre bontja, a közös faktorok hatására ($\underline{\Lambda}\,\underline{\eta}$) és a mérési hibára ($\underline{\varepsilon}$). A közös, a látens változók (faktorok) által magyarázott rész kovarianciái:

$$\underline{H} = \operatorname{cov}(\underline{\Lambda}\,\hat{\underline{\eta}}) = \underline{\Lambda}\,\underline{C}\,\underline{\Lambda}' . \tag{22}$$

A $\underline{H}$ diagonális elemei a megfigyelt változók varianciáival standardizálva:

$$\underline{H}^2 = (\underline{I}\,(\underline{\Lambda}\,\underline{C}\,\underline{\Lambda}')) \cdot (\underline{I}\,\underline{S})^{-1} \tag{23}$$

a megfigyelt változók kommunalitásait adják. A h_k^2 a k-adik megfigyelt változó (y_k) varianciájának az a része, amelyet a megfigyelt változóhoz közvetlenül kapcsolódó látens változók reprodukálnak.

---

[*] A Hadamard–szorzatot a következőképpen definiáljuk: $\underline{A} = \underline{B} \times \underline{C}$ ha $a_{jk} = b_{jk} c_{jk}$ bármely j, k-ra.

Kapcsoljuk össze a mérési modellt és a struktúrális egyenletet

$$\underline{y} = \underline{\Lambda}\,\underline{\beta}\,\underline{\eta} + \underline{\Lambda}\,\underline{\varsigma} + \underline{\varepsilon}\,. \tag{24}$$

Azt látjuk, hogy a manifeszt változók ($\underline{y}$) közös része két részre bontható: az egyedi részre $\underline{\Lambda}\,\underline{\varsigma}$ és a redundáns részre $\underline{\Lambda}\,\underline{\beta}\,\underline{\eta}$.

Az egyedi rész a megfigyelt változónak az a része, amit a vele közvetlenül összekötött látens változók önmagukban reprodukálnak. Az y_k redundáns részét akár a vele közvetlenül összekötött látens változók, akár ezek prediktorai reprodukálhatják.

A megfigyelt változók redundáns részeinek kovariancia mátrixa:

$$\underline{F} = \mathrm{cov}(\underline{\Lambda}\,\underline{B}\,\underline{\eta}) = \underline{\Lambda}\,\underline{B}\,\underline{C}\,\underline{B}'\underline{\Lambda}'\,. \tag{25}$$

Az $\underline{F}$ mátrix diagonális elemeit standardizáljuk a megfigyelt változók varianciáival:

$$\underline{F}^2 = (\underline{I}(\underline{\Lambda}\,\underline{B}\,\underline{C}\,\underline{B}'\underline{\Lambda}')) \cdot (\underline{I}\,\underline{S})^{-1}\,. \tag{26}$$

Az $\underline{F}^2$ mátrix diagonális elemei a megfigyelt változók redundancia–együtthatói. A redundancia–együttható a megfigyelt változó varianciájának az a része, amit a megfigyelt változóval közvetetten kapcsolódó látens változók magyaráznak. Az exogén megfigyelt változók redundanciája nulla.

A (24) egyenletbe behelyettesítjük a látens változók becsléseit ($\underline{\hat{\eta}} = \underline{\Omega}\,\underline{y}$):

$$\underline{y} = \underline{\Lambda}\,\underline{B}\,\underline{\Omega}\,\underline{y} + \underline{\Lambda}\,\underline{\varsigma} + \underline{\varepsilon}\,, \tag{27}$$

ahol a megfigyelt változók redundáns része:

$$\underline{\hat{y}} = \underline{\Lambda}\,\underline{B}\,\underline{\Omega}\,\underline{y} = \underline{M}\,\underline{y}\,. \tag{28}$$

A reziduális rész kovariancia mátrixa:

$$\begin{aligned}\mathrm{cov}(\underline{\hat{y}}) &= \underline{\Lambda}\,\underline{B}\,\underline{\Omega}\,\underline{S}\,\underline{\Omega}'\underline{B}'\underline{\Lambda}' \\ &= \underline{\Lambda}\,\underline{B}\,\underline{C}\,\underline{B}'\underline{\Lambda}' \\ &= \mathrm{cov}(\underline{\Lambda}\,\underline{B}\,\underline{\eta}) = (25) \quad \text{egyenlet}\,.\end{aligned}$$

6.4 A paraméterek becslése a parciális legkisebb négyzetek módszerével

A legkisebb négyzetek módszere a statisztikában jól ismert eljárás. A parciális jelzővel ellátott becslési módszer a klasszikus kritérium új kiterjesztését jelöli. A parciális legkisebb négyzetek módszer lényege, hogy a paramétereket részhalmazokra bontja, particionálja, az egyes partíciókat a legkisebb négyzetek módszere kritériuma szerint becsüli, miközben a többi paraméter kötött értékekkel szerepel. Az eljárás iteratív, az iterációs ciklus mindegyikében a paraméterek egy–egy részhalmazát becsüli a többi paraméter értékét ismertnek feltételezve.

A parciális legkisebb négyzetek módszerénél a következő előfeltételezéseket tesszük:

1) A manifeszt változók egymást át nem fedő (diszjunkt) blokkokra vannak particionálva.

2) A látens változók is egymást át nem fedő blokkokra vannak particionálva, és egy látens változó–blokk csak egy manifeszt változó–blokkhoz kapcsolódhat, vagyis a manifeszt és látens változók kapcsolódásait egy blokk–diagonális mátrix írja le.

3) Az útmodell (a strukturális egyenletrendszer) rekurzív.

A becslés technikai feltételei:

1) A látens változókat a manifeszt változók lineáris kombinációjával becsüljük:

$$\hat{\eta}_j = \Sigma(w_{jk} y_k)\,, \tag{29}$$

ahol

w_{jk} a ω_{jk} becslése.

2) A látens változók strukturális egyenleteit (útegyenleteket) a reziduális tag varianciájának minimalizálásával becsüljük:

$$\mathrm{var}(\zeta_k) = \Psi_k^2 \rightarrow \min\,. \tag{30}$$

3) A modell strukturális egyenleteken kívüli részét blokkonként, vagy

a) a súlymodell szerint $E(\eta_j|\underline{y}) = \Sigma(\omega_{jk} y_k)$, a súlyegyütthatókat a reziduális tag varianciájának minimalizálásával becsüljük:

$$\mathrm{var}(\delta_j) \rightarrow \min\,, \tag{31}$$

vagy

b) a mérési modell szerint $E(y_k|\underline{\eta}) = \Sigma(\lambda_{kj}\eta_j)$, a faktorsúly együtthatókat a reziduális tag varianciájának minimalizálásával becsüljük:

$$\mathrm{var}(\varepsilon_k) \rightarrow \min\,. \tag{32}$$

A parciális legkisebb négyzetek módszer algoritmusa iteratív eljárással ad becslést a paraméterekre. Minden iterációs ciklus három lépésből áll.

Az iterációs eljárás lépései:

1. lépés: Becsüljük a látens változókat vagy a súlymodell (18) kritériuma, vagy a mérési modell (19) kritériuma szerint: $\underline{\hat{\eta}} = \underline{W}\,\underline{y}$.

Kiszámítjuk a látens változók kovarianciamátrixának becslését:

$$\underline{\hat{C}} = \underline{W}\,\underline{S}\,\underline{W}' , \tag{33}$$

ahol

$\underline{S}$ a manifeszt változók variancia–kovariancia mátrixa, valamint a látens változók és a megfigyelt változók páronkénti kovarianciáit tartalmazó struktúra mátrix becslése:

$$\underline{\hat{T}} = \underline{S}\,\underline{W}' . \tag{34}$$

Az eljárás legelső lépésében a súlymátrix ($\underline{W}$) elemeit pszeudo–véletlen számokkal feleltetjük meg.

2. lépés: Amennyiben a látens változók egymásközti kapcsolódásaira feltételekkel élünk (közöttük korrelálatlanságot írunk elő), a második lépésben a látens változókat transzformáljuk a „patterned orthonormalizing rotation" eljárás szerint.

3. lépés: Becsüljük a látens változók belső súlyegyütthatóit:

$$\underline{\eta}^* = \underline{V}\,\underline{\hat{\eta}} .$$

A belső súlyegyütthatókat a következő módokon adhatjuk meg:

a) *útegyütthatóként*, amikor a belső strukturális modell útmodell . Ebben az esetben az exogén látens változó optimális becslő, az endogén látens változó optimális becsült, és a predeterminált látens változó optimális közvetítő változó.

A belső súlyegyütthatók ekkor a következők:

$$v_{hh'} = \begin{cases} b_{hh'} & \text{ha } \eta_{h'} \text{ megelőzi } \eta_h\text{-t} \\ r_{hh'} & \text{ha } \eta_{h'} \text{ követi } \eta_h\text{-t} \\ 0 & \text{ha } \eta_{h'} \text{ és } \eta_h \text{ nincs közvetlenül összekötve.} \end{cases} \tag{35}$$

b) *centroid együtthatóként*, amikor azt figyeljük, hogy a látens változókhoz mely látens változók kapcsolódnak közvetlenül. A látens változót az útdiagramban szomszédos látens változók súlyozatlan összegével becsüljük. Hogy az egymással negatívan korreláló szomszédos látens változók ne semlegesítsék egymás hatását, az előjelfüggvényt alkalmazzuk:

$$v_{hh'} = \begin{cases} \operatorname{sign}(r_{hh'}) & \text{ha } \eta_h \text{ és } \eta_{h'} \text{ közvetlenül össze van kötve} \\ 0 & \text{egyébként.} \end{cases} \tag{36}$$

Ezzel a feltétellel a látens változók úgy korrelálnak egymással, hogy a centroid faktornak (amit másodrendű faktornak tekinthetünk) maximális legyen a varianciája. (A centroid faktor faktorsúlyai 1, 0, -1 értéket vehetnek fel.)

c) faktor–súlyként, amikor az a célunk, hogy a látens változók főkomponensei varianciái maximálisak legyenek. Ekkor az együtthatók:

$$v_{hh'} = \begin{cases} r_{hh'} & \text{ha } \eta_h \text{ és } \eta_{h'} \text{ közvetlenül össze van kötve} \\ 0 & \text{egyébként.} \end{cases} \tag{37}$$

4. lépés: Becsüljük a belső súlyegyütthatókkal a látens változókat:

$$\underline{\eta}^* = \underline{V}\,\underline{\hat{\eta}}\,.$$

5. lépés: Ortogonálisan transzformáljuk a becsült látens változókat $(\underline{\eta}^*)$, ha a látens változók páronkénti kapcsolódásaira feltételezésekkel élünk.

6. lépés: Becsüljük a külső súlyokat $(\underline{W})$ a reziduális tag minimalizálásával:

$$\underline{\hat{\eta}} = \underline{W}\,\underline{y}\,.$$

Az iteráció az 1-től 6-ig lépés ismétléséből áll, és akkor áll le, ha a két egymás utáni sorozatban a súlyegyütthatókra a következő teljesül:

$$|w^{(s)} - w^{(s+1)}| < 10^{-5}\,,$$

vagy

$$|(w^{(s)} - w^{(s+1)})/w^{(s)}| < 10^{-5}\,.$$

6.5 A becslés illeszkedésének mérése

A becslés illeszkedésére nincs egyetlen indexünk, de a becslés jóságát megítélhetjük a következő mennyiségekkel, és így dönthetünk az érvényességéről:

$$h^2 = \mathrm{trace}(\underline{H}^2)/p \quad \text{(kommunalitás)}, \tag{38}$$

$$f^2 = \mathrm{trace}(\underline{F}^2)/p \quad \text{(redundancia)}, \tag{39}$$

$$c^2 = \mathrm{trace}(\underline{C}^2)/m\,, \tag{40}$$

$$s^2 = \mathrm{trace}(\underline{S})/p\,, \tag{41}$$

$$r^2 = \mathrm{trace}(\underline{C})/m\,, \tag{42}$$

$$\Theta^2 = \mathrm{trace}(\underline{\Theta})/m\,, \tag{43}$$

$$\Psi^2 = \mathrm{trace}(\underline{\Psi})/m\,, \tag{44}$$

$$s = \mathrm{rms}\underline{S}\,, \tag{45}$$

$$c = \text{rms}\underline{C}\,, \tag{46}$$

$$\Theta = \text{rms}\underline{\Theta}\,, \tag{47}$$

$$\underline{\Psi} = \text{rms}\underline{\Psi}\,, \tag{48}$$

$$\text{ceh} = \text{rmscov}(\underline{\varepsilon}\,\underline{\eta})\,, \tag{49}$$

$$\text{ceh} = \text{rmscov}(\underline{\varepsilon}\,\underline{\varsigma})\,, \tag{50}$$

$$\bar{c} = \sum_{h}^{-} \sum_{h<h'}^{-} |c_{hh'}|\,, \tag{51}$$

$$t = \text{rms}\underline{T}\,. \tag{52}$$

A fenti mennyiségekből h^2, f^2 és c^2 arányszámok, a variancia arányait jelölik. Ha a megfigyelt változók standardizáltak, akkor $\Theta^2 = s^2 - h^2 = 1 - h^2$ is arányszám. Ha a látens változók is standardizáltak, akkor a $\Psi^2 = 1 - c^2$ szintén arányszám. A paraméterek maximum likelihood becslésének veszteségfüggvényét

$$L = \log|\underline{\Sigma}| + \text{tr}(\underline{S}\,\underline{\Sigma}^{-1}) - \log|\underline{S}| - p \tag{53}$$

használhatjuk a modell tesztelésére. Ha a megfigyelt változók együttes eloszlása normális, az $(n-1)L$ eloszlása közelítőleg χ^2 eloszlású (n a megfigyelések száma)

$$f_t = (p^2 + p)/2 - t \tag{54}$$

szabadságfokkal, ahol t a modell független paramétereinek a számát jelöli.

Tekintsük a legegyszerűbb modellt:

$$H_d : \underline{\Sigma} = \underline{S}\,\underline{I}\,, \tag{55}$$

ahol a megfigyelt változókról feltételezzük, hogy páronként korrelálatlanok. A veszteségfüggvény egyszerűsödik, mivel $\underline{S}\,\underline{\Sigma}^{-1}$ mátrix egységmátrix, aminek a nyoma (trace) egyenlő p-vel, így

$$L_d = \log|\underline{I}\,\underline{S}| - \log|\underline{S}|\,. \tag{56}$$

Megjegyzés: Az rms$\underline{A}$ az $\underline{A}$ mátrix diagonális nélküli elemeinek négyzetes átlagát jelöli (root mean squared (rms)

$$\text{rms}\underline{A} = \left[\sum_{i}^{-} \sum_{\neq j}^{-} a_{ij}^2\right]^{1/2}$$

Ha a változók standardizáltak is, akkor $(\underline{I}\,\underline{S}) = \underline{I}$, és így $\log|\underline{I}\,\underline{S}| = 0$, és $L_d = -\log|\underline{S}|$. A szabadságfok $f_d = (p \cdot p - p)/2$. Tekintsük továbbá a korlátozott faktoranalitikus modellt:

$$H_r : \underline{\Sigma} = \Sigma(\underline{C}, \underline{\Lambda}, \underline{\Theta} = \underline{I}\,\underline{\Theta}), \tag{57}$$

az L_r veszteségfüggvénnyel, ahol a látens változók korrelációi és a faktorsúlyok közül rögzíthetünk értékeket (feltételezhetjük, hogy egyenlőek nullával), és feltételezzük, hogy a reziduális kovariancia mátrix diagonális.

A H_r és a H_s hiptézissel megfogalmazott modellek illeszkedésének különbözőségére TUCKER és LEWIS (1973) a következő megbízhatósági indexet definiálta:

$$r_{TL} = \frac{L_s/f_s - L_r/f_r}{L_d/f_d - 1/N}. \tag{58}$$

Vagy a BENTLER és BONNET (1980)-féle megbízhatósági index:

$$r_{BB} = \frac{L_s - L_r}{L_d}. \tag{59}$$

Mindkét megbízhatósági indexnél a modellek különbségeit a legegyszerűbb (a legkorlátozottabb) modellhez viszonyítottuk.

A látens út-modell mérési és súlymodell részét (a modell külső részét, egyenleteit) a mérési adatokhoz kielégítően illeszkedőnek akkor tekintjük,

– ha a manifeszt változók varianciáinak nem megmagyarázott része $(1 - h^2)$ elég kicsi,

– ha a manifeszt változók kovarianciáinak nem megmagyarázott része (Θ/s) elég kicsi,

– ha a h_k^2 (az y_k változó kommunalitása) elég nagy és a Θ_k^2 reziduális kovarianciák elég kicsik a mérési modellben,

– ha a blokkok közötti reziduális kovarianciák $(\Theta_{kk'})$ a nullához közelítenek,

– ha a mérési hiba (a reziduális tag) és a látens változók közötti kovarianciák $\mathrm{cov}(\underline{\varepsilon}\,\underline{\eta})$ közelítőleg nullával egyenlőek,

– ha egy blokkban csak egy manifeszt változó van, akkor a reziduális tag varianciája egyenlő nullával,

– ha egy blokkban csak két manifeszt változó van, akkor a reziduális változók tökéletesen korreláltak lesznek,

– ha egy blokkban csak kevés látens változó szerepel, a reziduális változók negatívan korrelálnak.

A modell strukturális (belső) részének illeszkedése kielégítő,

– ha a látens változók nem megmagyarázott része $(1-r^2)$ elég kicsi,

– ha a látens változók kovarianciáinak nem megmagyarázott része (Ψ/c) elég kicsi.

A teljes modell illeszkedését kielégítőnek tekintjük,

– ha a redundancia–együttható (f^2) elég nagy,

– ha a belső és külső reziduális tagok közötti kovarianciák $(\mathrm{cov}(\underline{\varepsilon}\,\underline{\zeta}))$ elég kicsik.

6.6 Kategorikus változók

Ismeretes, hogy a dichotom változó két kategóriájához bármely értéket (természetesen két különbözőt) hozzárendelhetünk. Ha a dichotom változó két lehetséges értékéhez a 0 és az 1 értékeket rendeljük, akkor a változót Boole–változónak nevezzük. A Boole–változót formálisan a következőképpen definiáljuk:

$$P(x=1) = 1 - P(x=0) = \mu\,. \tag{60}$$

Ebből a definicióból következik, hogy

$$E(x) = \Sigma P(x=i)\; i = \mu\; 1 + (1-\mu)\; 0 = \mu\,, \tag{61}$$

$$E(x^2) = \Sigma P(x=i)\; i^2 = \mu\; 1^2 + (1-\mu)\; 0^2 = \mu\,, \tag{62}$$

és általában

$$E(x^r) = \Sigma P(x=i)\; i^r = \mu\; 1^r + (1-\mu)\; 0^r = \mu\,, \tag{63}$$

$$\text{ha} \qquad r > 0\,.$$

Az x változó varianciája:

$$\sigma^2 = \mu(1-\mu) = \mu - \mu^2\,. \tag{64}$$

Egy kategorikus változó kategóriái legfeljebb megszámlálhatóan végtelen számosságú halmazt alkotnak. A k-adik kategória bekövetkezésének valószínűsége:

$$P(x=k) = \mu_k\,. \tag{65}$$

Egy kategorikus változót helyettesíthetünk Boole–változók egy halmazával a következő előírás szerint:

$$x = k \Longleftrightarrow x_k = 1\,. \tag{66}$$

A fentiekből következik, hogy

$$E(x_k) = E(x_k^2) = E(x_k^\nu) = P(x=k) = \mu\,. \tag{67}$$

vagyis a várható értéket mint valószínűséget értelmezhetjük. Boole–változók bármely párjának, hármasának, stb várható értéke:

$$E(x_k\, x_{k'}\, x_{k''} \ldots) = \begin{cases} \mu & \text{ha } k = k' = k'' = \ldots \\ 0 & \text{különben.} \end{cases} \tag{68}$$

Ha az $\underline{x}$ vektor tartalmazza a Boole–változókat, akkor a (67) a következőképpen írható fel:

$$E(\underline{x}) = \underline{\mu}\,. \tag{69}$$

Tekintsünk most két Boole vektor változót, $\underline{x}$ és $\underline{y}$-t az x és az y kategorikus változók helyettesítéseiként. Az $\underline{x}$ és $\underline{y}$ valószínűség eloszlása:

$$P(x_k = 1,\; y_\ell = 1) = \mu_{k\ell}\,, \tag{70}$$

ahol

$$k = 1, 2, \ldots, q\,,$$

$$\ell = 1, 2, \ldots, p\,.$$

A várható értékek:

$$E(x_k y_\ell) = \mu_{k\ell} \quad \text{és} \quad E(\underline{x}\,\underline{y}') = \underline{\mu}_{xy}\,. \tag{71}$$

A $\underline{\mu}_{xy}$ egy $(q \times p)$ típusú mátrix, a kétváltozós valószínűségek táblázata vagy más néven kontingencia táblázat. Helyezzük az $\underline{x}$ és $\underline{y}$ vektorokat az $\underline{u}$ vektorba:

$$\underline{u}' = [\underline{x}'\,,\, \underline{y}'] = [\ldots x_k \ldots, \ldots y_\ell \ldots]\,.$$

A $\underline{u}$ vektor várható értéke:

$$E(\underline{u}) = \mu_u = [\underline{\mu}'_x\,,\, \underline{\mu}'_y] = [\ldots \mu_k \ldots, \ldots \mu_\ell \ldots]\,.$$

Ezek után definiáljuk a szuper–kontingencia táblázatot:

$$E(\underline{u}\,\underline{u}') = E \begin{pmatrix} \underline{x}\,\underline{x}' & \underline{x}\,\underline{y}' \\ \underline{y}\,\underline{x}' & \underline{y}\,\underline{y}' \end{pmatrix} = \begin{pmatrix} \underline{\mu}_{xx} & \underline{\mu}_{xy} \\ \underline{\mu}_{yx} & \underline{\mu}_{yy} \end{pmatrix}. \tag{72}$$

A szuper–kontingencia táblázat tartalmazza a kontingencia táblázatot, a $\underline{\mu}_{xy}$-t, ennek transzponáltját, $\underline{\mu}_{yx}$-t, és a $\underline{\mu}_{xx}$ és $\underline{\mu}_{yy}$ diagonális mátrixokat. A $\underline{\mu}_{xx}$ diagonális mátrix diagonális elemei az egyváltozós valószínűségek ($\mu_{kk} = \mu_k$ a (68) szerint). Általában q kategorikus változó Boole– változóinak szuper–kontingencia mátrixa q^2 szubmátrixot (blokkot) tartalmaz.

Az x és y kategorikus változók megfigyeléseit x_i és y_i jelölje. A helyettesítő Boole–változók megfigyelt értékeit ekkor x_{ki} és $y_{\ell i}$ jelöli. A Boole– változók összege:

$$n_k = \sum_i x_{ki}\,, \qquad n_\ell = \sum_i y_{\ell i}\,, \tag{73}$$

és átlaga:

$$m_k = \overline{\sum}_i x_{ki}\,, \qquad m_1 = \overline{\sum}_i y_{\ell i}\,, \tag{74}$$

ahol

$\bar{\Sigma}$: az átlag müveletet jelöli.

A relatív szorzat momentum mátrix:

$$\underline{M} = \frac{1}{n}\begin{pmatrix} \underline{x}\,\underline{x}' & \underline{x}\,\underline{y}' \\ \underline{y}\,\underline{x}' & \underline{y}\,\underline{y}' \end{pmatrix}. \tag{75}$$

Definiáljuk még a normált szorzat mátrixot:

$$\underline{Q} = \left[q_{k\ell} = \frac{m_{k\ell}}{\sqrt{m_k m_\ell}}\right], \tag{76}$$

a kovariancia mátrixot:

$$\underline{C} = [c_{k\ell} = m_{k\ell} - m_k m_\ell]\,, \tag{77}$$

a normált eltérés mátrixot:

$$\underline{D} = \left[d_{k\ell} = \frac{m_{k\ell} - m_k m_\ell}{\sqrt{m_k m_\ell}}\right], \tag{78}$$

és a korrelációs mátrixot:

$$\underline{R} = \left[r_{k\ell} = \frac{m_{k\ell} - m_k m_\ell}{\sqrt{(m_k - m_k^2)(m_\ell - m_\ell^2)}}\right]. \tag{79}$$

A normált eltérés mátrix felhasználásával számíthatjuk a Yule–féle (phi) együtthatót:

$$\Phi^2 = \frac{1}{n}\chi^2 = \sum_k \sum_\ell d_{k\ell}^2\,, \tag{8}$$

és a kontingencia együtthatókat:

$$C_{\mathrm{Cramer}} = \sqrt{\Phi^2/(\min(pq) - 1)} \tag{81}$$

$$C_{\mathrm{Pearson}} = \sqrt{\Phi^2/(1 + \Phi^2)} = \sqrt{\chi^2/(n + \chi^2)}\,. \tag{82}$$

6.6.1 Kanonikus korrelációelemzés

Legyen adva két változóhalmaz, x_k $(k = 1, 2, \ldots, q)$ és y_ℓ $(\ell = 1, 2, \ldots, p)$. Adottak továbbá a szorzatmátrixok $\underline{M}_{xx}$ és M_{yy}, valamint a várható értékek vektorai: $\underline{m}_k = \underline{0}$ és $\underline{m}_y = \underline{0}$.

A kanonikus korreláció modellje a következő egyenletekből áll:

$$\xi = \sum_k w_k x_k\,, \qquad \mathrm{var}(\xi) = \underline{w}'\underline{M}_{xx}\underline{w} = 1\,, \tag{83}$$

$$\eta = \sum_\ell v_\ell y_\ell\,, \qquad \mathrm{var}(\eta) = \underline{v}'\underline{M}_{yy}\underline{v} = 1\,, \tag{84}$$

$$\eta = \rho\cdot\xi + \varsigma\,, \qquad E(\eta|\xi) = \rho\cdot\xi\,, \tag{85}$$

ahol ξ és η a kanonikus látens változók egységnyi varianciával, $\underline{w} = [w_k]$ és $\underline{v} = [v_\ell]$ a kanonikus súlyegyütthatók, és ρ az első kanonikus korreláció.

A kanonikus korreláció modell skálafüggetlen, ami azt jelenti, hogy például x_k-t megszorozzuk $1/c$ konstanssal és a w_k értékét c-vel, akkor a (85) változatlan marad.

A kanonikus korreláció modellben annyi kanonikus változó–pár állítható elő, amennyi az x_k és y_k változó–halmaz elemeinek a minimuma $(\min(p, q))$

Bizonyítható (lásd LANCASTER, 1969), hogy a kanonikus korrelációs együtthatók négyzetösszege megegyezik a Yule–féle együtthatóval:

$$\Phi^2 = \sum_i^{\min(p,q)} \rho_i^2\,, \tag{86}$$

ami azt is jelenti, hogy a kanonikus korrelációs együtthatók négyzetes átlagának négyzetgyöke egyenlő a Cramer–féle kontingencia együtthatóval.

6.6.2 Főkomponenselemzés

Legyen adva a dichotom változók normált vektora:

$$\underline{u}' = [u_k] = [\ldots x_{k(Q)} \cdots, \ldots y_{\ell(Q)} \cdots]\,,$$

ekkor a dichotom változók lehetséges értékei: $(0, \frac{1}{\sqrt{m_k}})$.

Adottak továbbá a szorzat modell egyenletei:

$$u_k = p_k\eta + e_k \qquad E(u_k\eta) = p_k\eta\,, \tag{87}$$

$$\eta = \sum_k w_k u_k \qquad \mathrm{var}(\eta) = \underline{w}'\underline{Q}\underline{w}\,. \tag{88}$$

Be lehet bizonyítani, hogy a $\underline{Q}$ mátrix sajátértéke:

$$d_{(Q)} = \sum_k p_k^2 ,$$

és az $\underline{x}_Q$ és $\underline{y}_Q$ közötti kanonikus korreláció között az alábbi összefüggés létezik:

$$d_{(Q)} = 1 + \rho_{(Q)} , \tag{89}$$

valamint, hogy a főkomponens (FK) súlyok $w_{k(FK)}$ és a kanonikus súlyok (KK) $w_{k(KK)}$ és $v_{\ell(KK)}$ úgy aránylanak egymáshoz, hogy a kanonikus változók értékeit közvetlenül a főkomponens súlyokból számíthatjuk:

$$\begin{aligned} \xi_{(KK)} &= f \sum_k w_{k(FK)} x_{x(Q)} , \\ \eta &= g \sum_\ell w_{\ell(FK)} y_{\ell(Q)} , \end{aligned} \tag{90}$$

ahol f és g a látens változók standardizálásához szükségesek.

6.6.3 A kategóriák skálázása

Egy kategorikus változót — láttuk az előzőekben — helyettesíthetünk Boole–változókkal úgy, hogy minden kategóriához hozzárendelünk egy $(0, 1)$ értéket felvevő változót. A k-adik Boole–változó akkor veszi fel az 1–et $(y_k = 1)$, ha a kérdéses megfigyelés (i) az y változó k-adik kategóriába esik $(y_i = k)$. Ebből következően felhasználva az előző két alfejezet látens változó–képzését:

$$\sum_k w_k y_{ki} = w_k , \tag{91}$$

vagyis a látens változó értéke az i-edik megfigyelésnél (ha a k-adik kategóriába tartozik) egyenlő lesz a w_k súlyegyütthatóval, ami pedig a k-adik kategóriához rendelt Boole–változó súlya. Így a látens változó a megfigyelt változó intervallum mérési szintű változójának tekinthető.

Több, mint két kategorikus változó esetén a kategóriák skálázásához a szuper–kontingencia mátrixot $(\underline{M})$ normalizáljuk $(\underline{Q})$, meghatározzuk a főkomponens súlyokat $(x_{k(Q)})$ és ezután a kapott súlyokat visszatranszformáljuk:

$$w_{k(M)} = w_{k(Q)} / \sqrt{m_k} . \tag{92}$$

A $\underline{Q}$ mátrix első komponense triviális, a főkomponens együtthatói (loadings) egyenlőek m_k-val, a súlyok pedig 1-gyel. A második (és a további főkomponensek pedig egyenlőek lesznek a $\underline{D}$ (normált eltérés) mátrix első (és többi) főkomponenseivel, így a gyakorlati megoldásoknál a $\underline{D}$ mátrixot használjuk.

6.7 Háromdimenziós útelemzés

Az eddigi modellek kétdimenziós mátrixok elemzését végezték. A háromdimenziós általánosításukhoz nézzük először az ún. Kronecker szorzatot, amit mátrixok között értelmezünk.

Az $\underline{A}$ mátrix a $\underline{B}$ és $\underline{C}$ mátrix Kronecker szorzata:

$$\underset{(JK \cdot PQ)}{\underline{A}} = \underset{(J \cdot P)}{\underline{B}} \otimes \underset{(K \cdot Q)}{\underline{C}} = a_{(jk)(pq)} = b_{jp} c_{kq} \,. \tag{93}$$

Bevezetünk még két jelölést:

– a vektorfüggvényt, jele: vec $\underline{A}$,

$$\text{vec}\underline{A} = (a_{11}.a_{12}, \ldots, a_{1K}, a_{21}, \ldots, a_{jk}, \ldots, a_{JK})' = a_{(jk)} \,.$$

A vec$\underline{A}$ függvény az $\underline{A}$ mátrix sorait sorba belehelyezi egy oszlopvektorba. A zárójelben lévő dupla index kombinált indexet jelöl: $(jk) = (j-1) \cdot J + k$ ahol a második index (k) gyorsabban fut mint az első index (j).

– egy paramétermátrix ($\underline{A}$) modell (pattern) mátrixát, $\underline{M}$-et vagy más jelöléssel $\underline{M}_{\underline{A}}$, $M(\underline{A})$ vagy pattern ($\underline{A}$). Az $\underline{M}$ mátrix azt jelöli, hogy a paramétermátrix mely eleme a szabad paraméter, és melyet rögzítettünk 0-hoz:

$$\underline{A} = \underline{M}_{\underline{A}}\, \underline{A} \,.$$

Ha $m_{jk} = 1$, akkor bármely valós a_{jk}-ra $(m_{jk} \cdot a_{jk} = a_{jk})$. Ha $m_{jk} = 0$, akkor az $a_{jk} = 0$ kötött elem. Ha a modell mátrix egységmátrix, akkor a paraméter mátrix diagonális kell hogy legyen:

$$\underline{A} = \underline{I}\, \underline{A} \,.$$

Az indexek az eddigiektől eltérnek, így összefoglaljuk őket:

$$i = 1, 2, \ldots, I \,,$$
$$j = 1, 2, \ldots, J \,,$$
$$k = 1, 2, \ldots, K \,,$$
$$r = 1, 2, \ldots, R \,,$$
$$q = 1, 2, \ldots, Q \,.$$

A Kronecker főkomponensek

Ha egy $\underline{A}$ mátrixnak Kronecker struktúrája van, és elég nagy mátrix, akkor jól jellemezhetjük a viszonylag kicsi $\underline{B}$ és $\underline{C}$ paraméter mátrixokkal. A modellben

$$\underline{A} = \underline{B} \otimes \underline{C} + \underline{A}_r \,, \tag{94}$$

ahol $\underline{B}$ és $\underline{C}$ mátrixokat az $\underline{A}$ mátrix Kronecker főkomponenseinek nevezzük. A modell becslését a legkisebb négyzetek módszerének kritériuma szerint a következő függvény minimalizálásával kapjuk meg:

$$\Phi = \mathrm{trace}(\underline{A}_r' \, \underline{A}_r) \Longrightarrow \min \,. \tag{95}$$

Ha az $\underline{A}$ mátrix szimmetrikus (multitrait–multimethod) korrelációs mátrix, akkor a reziduális $\underline{A}_r$ mátrix diagonális elemeit elhagyhatjuk a minimalizáláskor. Ekkor a reziduális mátrix:

$$\underline{A}_s = \underline{A}_r \, (\underline{1} - \underline{I}) \,, \tag{96}$$

és a $\underline{B}$ és $\underline{C}$ mátrixokat Kronecker Minres faktornak nevezzük, amelyek a

$$\Phi = \mathrm{trace}(\underline{A}_s' \underline{A}_s) \rightarrow \min \tag{97}$$

függvényt minimalizálják.

A Kronecker főkomponensek illeszkedésének jóságát a következő index mutatja:

$$fitkpc(\underline{A}) = 1 - \mathrm{trace}(\underline{A}_r' \underline{A}_r) / \mathrm{trace}(\underline{A}' \underline{A}) \,. \tag{98}$$

Háromdimenziós modellek

A következő táblázatban Lohmöller és Wold nyomán összefoglalunk különböző kétdimenziós modelleket és a háromdimenziós megfelelőiket.

Modell	Kétdimenziós	Háromdimenziós
Komponens elemzés	$\underset{(J\times I)}{\underline{Y}} = \underset{(J\times P)}{\underline{A}}\,\underset{(P\times P)}{\underline{D}}\,\underset{(P\times I)}{\underline{X}} + \underset{(J\times I)}{\underline{E}}$	$\underset{(JK\times I)}{\underline{Y}} = \underset{(JK\times PQ)}{(\underline{A}_1\otimes\underline{A}_2)}\,\underset{(PQ\times M)}{\underline{D}}\,\underset{(M\times I)}{\underline{X}} + \underset{(JK\times I)}{\underline{E}}$ Háromdimenziós komponens modell Tucker, 1966
Faktor-elemzés	$\underset{(J\times 1)}{\underline{Y}} = \underset{(J\times P)}{\underline{A}}\,\underset{P\times 1}{\underline{\eta}} + \underset{(J\times 1)}{\underline{\epsilon}}$	$\underset{(K\times 1)}{\underline{y}} = \underset{(JK\times PQ)}{(\underline{A}_1\otimes\underline{A}_2}\,\underset{(PQ\times 1)}{\underline{\eta}} + \underset{(JK\times 1)}{\underline{\epsilon}}$ Háromdimenziós faktormodell, Bloxom (1968), Bentler and Lee (1978)
Út–modell manifeszt változókkal (MVPM)	$\underset{(J\times 1)}{\underline{y}} = \underset{(J\times J)}{\underline{B}}\,\underset{(J\times 1)}{\underline{y}} + \underset{(J\times 1)}{\underline{\epsilon}}$	$\underset{(JK\times 1)}{\underline{y}} = \underset{(JK\times JK)}{(\underline{B}\otimes\underline{B}_2)}\,\underset{(JK\times 1)}{\underline{y}} + \underset{(JK\times 1)}{\underline{\epsilon}}$ Háromdimenziós út–modell (MVP3M) Lohmöller (1983)
Út–modell látens változókkal (LVPM)	$\underset{(J\times 1)}{\underline{y}} = \underset{(J\times P)}{\underline{A}}\,\underset{(P\times 1)}{\underline{\eta}} + \underset{(J\times 1)}{\underline{\epsilon}}$ $\underset{(P\times 1)}{\underline{\eta}} = \underset{(P\times P)}{\underline{B}}\,\underset{(P\times 1)}{\underline{\eta}} + \underset{(p\times 1)}{\underline{\delta}}$	$\underset{(JK\times 1)}{\underline{y}} = \underset{(JK\times PQ)}{(\underline{A}_1\otimes\underline{A}_2}\,\underset{(PQ\times 1)}{\underline{\eta}} + \underset{(JK\times 1)}{\underline{\epsilon}}$ $\underset{(PQ\times 1)}{\underline{\eta}} = \underset{(PQ\times PQ)}{(\underline{B}_1\otimes\underline{B}_2)}\,\underset{(PQ\times 1)}{\underline{\eta}} + \underset{(PQ\times 1)}{\underline{\delta}}$ LVP3M, Lohmöller (1983)
Komponens elemzés kevert mérési skálával	$\underset{(K\times I)}{\underline{Z}} = \text{scaled}\,\underset{(J\times I)}{(\underline{Y})} + \underset{(J\times I)}{\underline{F}}$	$\underset{(JK\times I)}{\underline{Z}} = \text{scaled}\,\underset{(JK\times I)}{(\underline{Y})} + \underset{(JK\times I)}{\underline{F}}$
	$\underset{(J\times I)}{\underline{Y}} = \underset{(J\times P)}{\underline{A}}\,\underset{(P\times P)}{\underline{D}}\,\underset{(P\times I)}{\underline{X}} + \underset{(J\times I)}{\underline{E}}$ PRINCIPALS (Young, Takane and de Leeuw 1978)	$\underset{(JK\times I)}{\underline{Y}} = \underset{(JK\times RR)}{(\underline{A}_1\otimes\underline{A}_2)}\,\underset{(R\times R)}{\underline{D}}\,\underset{(P\times I)}{\underline{X}} + \underset{(JK\times I)}{\underline{E}}$ ALSCOMP3 (Sands and Young 1980)

A táblázatban szereplő jelölések:

$\underline{A}$: faktorsúly mátrix, $(J \times P)$ jelöli a mátrix méretét, vagyis az $\underline{A}$ mátrixnak J sora és P oszlopa van, $(J \times P)$ típusú mátrix,

$\underline{B}$: az útegyütthatók mátrixa,

$\underline{E}, \underline{F}, \underline{\varepsilon}, \underline{\delta}$: a reziduálisokat tartalmazzák,

$\underline{Z}, \underline{Y}$: adatmátrixok,

$\underline{X}$: főkomponensek értékei,

$\underline{y}$: manifeszt változók vektora,

$\underline{\eta}$: látens változók vektora,

$\underline{D}$: sajátértékek mátrixa.

6.8. *Példa látens változók útelemzésére*

(Társadalmi státusz – értékrend – magatartás útmodellek)

A látens változókat a megfigyelt (manifeszt) változók mérési modelljeiként írjuk le. A manifeszt változók három szintjét különítjük el. Az első szinten az egyén backgroundját, hátterét írjuk le: a Család társadalmi státusát a Nagyapa, Apa és anya iskolai végzettségével mérve, a Gyerekkort a különböző településkategóriákban az első 14 életévben eltöltött idővel, Gyerekkori boldogsággal, Jóléttel és Zaklatottsággal kifejezve. A második szint a jelen jellemzését adja: a Társadalmi státus látens változót a Nem, Lakóhely, Életkor, Iskolai végzettség, Jövedelem megfigyelt változók fejezik ki, a Vallásosságot pedig a "Vallásos szellemben nevelték-e Önt a szülei?" és a "Vallásos embernek tartja–e magát?" kérdésekkel mértük. A Család társadalmi státusa, a Gyerekkor, a Társadalmi státus és a Vallásosság látens változók és az őket meghatározó megfigyelt változók kapcsolatát a modellben belsőnek (inwards) tekintjük. Ez alatt azt értjük, hogy a manifeszt változók, mint az ismeretlen dimenzióról összegyűjtött megfigyelt változók, a hatásukat kifejező súlyegyütthatókkal előállítják a látens változót anélkül, hogy azt gondolnánk, hogy a látens változó a felelős a megfigyelt változók varianciáiért. A modellben ezt úgy jelöljük, hogy megfigyelt változókból húzzuk a nyilat a látens változó felé. A manifeszt változók harmadik szintje az endogén változók halmaza. Egy vagy két blokkot különítünk el. Az egyik — a modellekben állandó — az értékrendszert kifejező blokk. Megfigyelt változói valójában már önmagukban is képzett változók, a MINISSA tér három dimenziója (az első kettő 45 fokkal elforgatva). A modellek abban különböznek, hogy az értékrendszer mellé milyen output változókat jelölünk második blokként. Jellegében ezek az emberek magatartásait, vélekedéseit és értékeléseit fejezik ki. A két endogén blokkban a látens és megfigyelt változók kapcsolata külső (outwards) jellegű, feltételünk szerint a látens változó, mint egy közös faktor, dimenzió felelős a manifeszt változók variancia–kovarianciáiért. Jelölésben a látens változóból a megfigyelt változóba mutató nyillal jelezzük ezt.

6.8.1 *A modell látens változói*

A Család társadalmi státusza

A Nagyapa, Nagyanya, Apa és Anya iskolai végzettsége közül a legnagyobb súlya (és ez átlagosan kétszerese az őt követő legfontosabb súlynak) az Apa iskolai végzettségének van. (Ezek a súlyok a különböző modellekben .53-tól .68-ig terjednek.) Másodsorban határozza meg a Család társadalmi státuszát az Anya iskolai végzettsége. A súlyok a modellben. .33-tól .38-ig terjednek. A nagyszülők közül a Nagymama iskolai végzettsége játszik nagyobb szerepet, a súlyok .19 és .26 között mozognak.

(A Bizalom endogén változónál ez a súly megugrik és a .36 értéket veszi fel.) A Nagyapa iskolai végzettségének kicsi az együtthatója, de ez az együttható negatív, ami a társadalom nagymértékű átrétegeződésére utal. Ha generációnként is megvizsgál;juk a Család társadalmi státusz látens változót azt találjuk, hogy a Nagyapa iskolai végzettsége a 40–59 éves generációnál jelenik meg negatív együtthatóval. És ez a súly nagyobb, mint az Anya iskolai végzettségének a súlya. Ugyanennél a generációnál a Család társadalmi státusz látens változót legnagyobb mértékben a Nagyanya iskolai végzettsége határozza meg.

A 60 éves és idősebbeknél a Nagyanya iskolai végzettségének van negatív súlya, és ennél a generációnál az Apa iskolai végzettségének a szerepe a meghatározó. (.81-től .96-ig terjedő együtthatókkal.)

A 20–39 éves generációnál átlagosan kisebbek a faktorsúlyok, kiegyenlítettebb az Apa, Anya és Nagyanya együtthatójának értéke, és nincs köztük negatív súly.

Gyerekkor

A Gyerekkor látens változó részben a gyerekkorban (első 14 év) különböző településeken eltöltött idővel, részben kilenc foku létrán mért boldogsággal és jóléttel, és ötfoku skálán mért zaklatottsággal függ össze. A Gyerekkor látens változó pozitívan kapcsolódik Budapesttel és a Nagyvárosban eltöltött gyerekkori idővel, negatívan a Falun és Tanyán töltött gyerekkorral, pozitívan a Boldogsággal, negatívan (kicsi súllyal) a Zaklatottsággal, és átlagosan kicsi a súlya a Jólétnek.

A modellek Gyerekkori látens változója az őt meghatározó manifeszt változók alapján a boldogabb, urbanizált településen eltöltött gyerekkori évek kontra a falun, tanyán töltött boldogtalan évek dichotómiát fejezi ki.

A különböző generációk közötti alapvető különbség, hogy a fiatal 20–39 éves generációnál lecsökken a Tanya és a Falu települések súlya és ezzel párhuzamosan megnő a városi kategóriák együtthatója (Kisváros, Nagyváros, Budapest), az idősebb generációknál ez éppen ellenkezőleg változik, a falu–város dichotómiában a falusi kategóriáknak van nagyobb szerepe. Még egy; nem jelentős súllyal, de a 20–39 évesek Zaklatottság emlékezete kap a negatív irányba nagyobb együtthatót (-.10).

A gyerekkori Jólét csak a Bizalom, Egyenlőtlenség, Pozició és a Tudat endogén blokkok esetén játszik nagyobb szerepet a Gyerekkor látens változó meghatározásában (.21 és .28 faktorsúlyokkal).

Társadalmi státusz

A Társadalmi státusz látens változót öt manifeszt változóval fejeztük ki: Nem, Lakóhely, Életkor, Iskolában töltött évek, Jövedelem (egy főre jutó jövedelem). Ezek közül alapvetően az Iskola (.56) és a Lakóhely (.46) határozza meg a társadalmi státusz értékét. A modellekben (átlagosan) egységnyi emelkedés a társadalmi hierarchiában a látens változó értékelése szerint elérhető egységnyivel urbanizáltabb településekre költözéssel és egységnyivel több iskolával.

Az Életkor, mint ezt láttuk már korábban is, a társadalmi hierarchiában eltöltött hellyel párhuzamosan mozog a különböző érték és tudati modellekben. A Társadalmi státusz látens változót -.26-tól -.38-ig terjedő súllyal határozza meg.

A Jövedelemnek kicsi a szerepe a társadalmi státusz emelkedésében (.07) az is inkább az idősebb generációknál jellemzőbb (.10-től .27-ig).

A Társadalmi státusz emelkedése — gyenge összefüggésben (.05) — de inkább a férfiakhoz kapcsolódik.

Vallásosság

A Vallás látens változót legnagyobb részt (.80 feletti súlyokkal) az ateista vagyok-tól a rendszeresen járok templomba kategóriákat tartalmazó kérdésekre adott válaszok határozzák meg. Kisebb súlya van a Vallásos szellemű nevelésnek (.25-től .44-ig terjedő együtthatókkal).

6.8.2 A modellek endogén változói

Az alapmodell: Mi határozza meg az emberek értékrendjét?

Mit is nevezünk értékrendnek? A Rokeach értékek axiológiai tere (MINISSA-tér) három dimenziója (az első kettő 45 fokkal elforgatva) a Jólét–Eszme, Közösség–Autonómia és az Indusztriális–Posztindusztriális dichotómiával jellemezhető. Ezek közös, a társadalmi helyzettel magyarázható dimenziója a Jólét, Közösség, Indusztriális értékektől az Eszme, Autonómia, Posztindusztriális értékekig terjed. Ezen belül is legnagyobb súlya a Közösség–Autonómia tengelynek van, amíg az Indusztriális–Posztindusztriális tengely szerepe csekély. Azt mondhatjuk tehát, hogy az Értékrend látens változó negatív oldalán a hagyományos értékek választását, pozitív oldalán pedig a modernizáció értékeinek választását méri. A modellben meghatározott Értékrend látens változót a Család társadalmi státusza, Gyerekkor, Társadalmi státusz, Vallás látens változók 26%-ban (R^2) tudják meghatározni. Ezen belül a Társadalmi státusznak a direkt hatása .40 (útegyüttható), teljes hatása a Vallásosságon keresztül .50, vagyis azt mondhatjuk, hogy egységnyi emelkedés a társadalmi státuszon (amit elérhetünk például úgy, hogy a településrendszerben egységnyivel urbánusabb településre költözünk és eggyel növeljük iskolai végzettségünket) értékrendünk modernizációjában .50 egységnyi fejlődést idéz elő. A Család társadalmi státuszának direkt hatása csekély, de teljes hatásban az iskolázottabb családi háttér–környezet .27 egységgel modernizáltabbá teszi értékrendünket. (Megjegyezzük, hogy a Család társadalmi státuszában legnagyobb súlya az apa iskolai végzettségének van, kisebb, de jelentős szerepe van az anya iskolai végzettségének is, a nagyanya iskoláinak nagyobb a jelentősége, mint a nagyapáé). Általában, a modell szerint a családi és gyerekkori háttérnek a közvetlen hatása az emberek értékválasztásaira csekély és bizonytalan, azonban azzal, hogy az emberek társadalmi státuszát jelentős súllyal meghatározzák, ezen keresztül azt mondhatjuk, hogy iskolázottabb szülők és nagyszülők, kevés gyerekkori év, amit falun vagy tanyán töltöttünk. Több gyerekkorban jólétben eltöltött boldog (kevésbé zaklatott) városiu év az embereket a hagyományos értékek falől az autonóm, modern értékek választása felé segíti.

Modernizáció

Az előző modell értékdimenzióit kicseréltük három, az értékekből választott dichotómiával: Érzelem – Pragmatikus (Megbocsátó, Szeretettel teljes – Alkotó szellemű, Hatékony), Érzelem – Racionalitás (Megbocsátó, Szeretettel teljes – Értelmes, Logikus gondolkodású), Hagyományos – Önálló (Engedelmes, Segítőkész, Tiszta, Udvarias – Bátor, Önálló).

A modell látens változóját nevezhetjük a Modernizációnak, mivel az Érzelem, Hagyományos értékektől a Pragmatikus, Racionális és Önálló értékekig méri az értékválasztást. Ez a modernizációs látens változó hasonlóan működik, mint az előbbi modellben az Értékrend látens változó, melyik szintén egy modernizációs tengelyként értelmezhető.

A társadalmi státuszon felfelé lépkedve csökken a vallásosság (-.50), amely pedig a Modernizáció ellen hat (-.40), így a társadalmi státusz direkt hatása (.32) a Vallás látens változón keresztül még felerősödik, a teljes hatás így .43. A Család és Gyerekkori háttér ebben a modellben is a Társadalmi státuszon keresztül hat a Modernizációra (a totális együtthatók: .18 és .16).

Anómia

A modellben bevezettük az Értékrend mellé endogén változónak az Anómiát. Bár a modell tanusága szerint az anómikusság nem függ túlságosan a társadalmi státusztól és a háttértől, még kevésbé az értékrendtől, a látens változók teljes hatása nem elhanyagolható:

Család társadalmi státusza:	–.19
Gyerekkor:	–.08
Társadalmi státusz:	–.27

Vagyis a társadalmi hierarchián felfelé haladva kevésbé anomikusak az emberek Magyarországon. Ahhoz mondjuk, hogy valaki egységnyivel csökkentse anómiáját, nem kell mást tennie, mint hogy például elköltözik faluról nagyvárosba, és képezi magát egy kicsit, mondjuk a szakma után elvégzi a közepiskolát és az egyetemet. De ha a társadalmi hierarchiában az előbb említett mozgást megteszi és úgy dönt, anomikus marad, célszerű, ha nem enged a szekularizációnak és megmarad a hagyományos értékrendnél.

Három generációra külön–külön is megbecsültük a modell paramétereit. A 20–39 éves generációnál nagyon lecsökken a modell státusz direkt hatása (–.09) és a teljes hatás is csak akkora, mint a teljes mintában a közvetlen hatás (–.18). Megnőtt viszont az értékek fontossága. Eszerint egységnyi lépés a hagyományos értékek felé, és .16 egységgel nő az anómia. Nő az anómia szintje az alacsony családi háttérrel. Nő az anómia a gyerekkorban eltöltött városi élettel is, bár ez a gyerekkor indirekt hatása közvetítőkön keresztül nézve lecsökken, de nem fordul át (míg a többi generációnál és a teljes mintában is ez a gyengébb közvetlen hatás miatt átfordul).

A 40–59 éves generációnál az Anómia varianciájának mindössze 6%-át tudjuk az öt látens változóval megmagyarázni. Egyedül a társadalmi státusz látens változó közvetlen hatása jelentős (–.21), vagyis a társadalmi hierarchián egységgel lejjebb menve az anómia .21 egységgel nő. Az Értékrend választása ennél a rétegnél nem kapcsolódik szisztematikusan az anómiához.

Társadalmi célok (Inglehart)

A 12 Inglehart célkitűzés közül a modell látens változója azokat állítja elő pozitív súllyal, amelyeket Inglehart postmateriálisnak nevezett. A Társadalmi célok látens változó negatív együtthatóval fejezi ki a materiális célokat. Ezt a szabályt két cél megszegi, "Az ország gazdasági egyensúlyát biztosítani" és "Városainkat, falvainkat és tájainkat szebbé tenni" célok, amelyek megcserélődnek, a "gazdasági egyensúly..." a postmateriális célokhoz kerül, a "városaink szebbé tevése..." pedig materiális céllá válik. A társadalmi célok látens változót azért elnevezhetjük posztmateriális – materiális dimenziónak (azért nem fordítva, mert a manifeszt változók mérésekor rangsorolást alkalmaztunk).

Megjegyezzük, hogy ez a látens változó nagyon hasonlít a MINISSA első tengelyére, vagyis jól értelmezhető a modell többi blokkjától függetlenül is mint a célok immanens dimenziója.

A Posztmateriális – Materiális látens változót a családi és gyerekkori háttér nem befolyásolja. A Társadalmi státusz közvetlen hatása –.28, teljes hatása (az értékrenden és a valláson keresztül) –.40, vagyis a társadalmi hierarchián egységgel lejjebb a célok választása a materiális irányban mozog .4o egységgel, a társadalmi hierarchián feljebb jutva a cél–váltás a postmateriális irányba mozog ugyanilyen mértékben. Ha külön blokkban szerepeltetjük a Materiális és Posztmateriális célokat, az derül ki, hogy a Materiális célok választása a társadalmi hierarchiával ellentétes irányba változik, ennek a kapcsolatnak az együtthatója .22, azonban a Posztmateriális célok választását csak gyengén befolyásolja a társadalmi státusz.

Generációnként ez a kép különbözik. Míg a 20–39 éveseknél a Materiális dimenzió háttérbe kerül (.05), és fontosabbá válik a Posztmateriális cél (–.13).

A 40–59 éveseknél ez megfordul, a társadalmi státusz nagyobb súllyal befolyásolja a materiális célokat (.28).

A 60 éves és ennél idősebbeknél szintén a Posztmateriális célok választása válik fontosabbá a társadalmi státusz emelkedésével (.23).

Bizalom

A modellben szereplő társadalmi pozicíót kifejező blokkok egyáltalán nem, vagy csak gyengén befolyásolják a Bizalom látens változót, de a modernebb értékrend is csak kis súllyal esik latba (.04) abban, hogy valaki megbízik-e az emberekben vagy sem. Mégis, a magasabb Családi társadalmi státusz háttérként inkább az önmagukban való bízást és nem általában az emberekben való bizalmat erősíti (-.06). Ez az elért társadalmi státusz hatásával ellensúlyozódik, és átfordul egy nagyon gyenge, vagy semmilyen másokban való bízásba. A Bizalmat, ha gyengén is, de erősíti a társadalmi hierarchiában elfoglalt magasabb pozíció (.10).

Mit is tartalmaz a Bizalom látens változó?

Az egyik oldalán az "emberekben általában meg lehet bízni" kijelentésnek van nagy súlya és kisebb az "emberek többsége szívesen segít másokon" kijelentés fontossága, a dimenzió másik oldalát az "emberek többsége csak önmagával törődik" és "ha törésre kerül a sor, nem sok jót lehet várni az emberektől" kijelentésekkel való egyetértés határozza meg. Az értékrend modernizációja alig, csak nagyon-nagyon gyengén (.04) növeli a Bizalmat.

Más képet kapunk azonban generációnként., A 20–39 éveseknél a Társadalmi státusszal a *Bizalmatlanság* növekszik (-.10), a 40–59 éveseknél viszont már a magasabb társadalmi státusszal megjön a Bizalom is (.05), és a 60 évesek vagy idősebbeknél ez egy kicsit még fokozódik is (.07).

Megjegyzendő, hogy ezek a súlyok meglehetősen kicsik, így a kapcsolódás csak nagyon gyengének mondható.

Az Értékrend látens változó eltérő tartalmat kap generációnként.

A 20–39 éveseknél megnő az Indusztriális - posztindusztriális tengely fontossága, és ellentétes előjelű lesz a Közösség - Autonóm és Jólét - Eszme tengellyel. Ennél a rétegnél az Eszme, Autonómia értékek együttjárása az Indusztriális értékekkel növeli legjobban a Bizalmat.

A Jólét, Közösség és Posztindusztriális értékek együttes választása növeli legjobban viszont a Bizalmatlanságot, vagyis a csak magunkban bízást. Megjegyezendő, hogy az a furcsa modernizáció ennél a generációnál is csak itt, a Bizalom output változó-blokk esetén jelent meg.

Participáció, beleszólás

A részvételt és beleszólást a szűkebb és tágabb környezet döntéseinek meghozatalába a társadalmi háttérrel és státusszal, vallással és értékrenddel a következőképpen tudjuk magyarázni:

	Participáció Direkt hatás	Teljes hatás	Indirekt hatás
Család társadalmi státusza	–.01	.13	.14
Gyerekkor	–.03	.11	.14
Társadalmi státusz	.17	.32	.15
Vallás	–.15	–.18	–.03
Értékrend	.16	.16	–

Összesen a fenti tényezők a variancia 13%-át tudják reprodukálni. Látható, hogy a családi és gyerekkori háttér közvetlen hatása elhanyagolható, a jelenlegi társadalmi státuszra gyakorolt jelentős hatásukkal azonban a súlyuk felemelkedik .13, illetve .11-re.

A társadalmi státusz egységnyi emelkedésével a Participáció .17 egységnyivel emelkedik. De ha azt is figyelembe vesszük, hogy a Vallásosság .50 egységnyivel csökken, valamint azt, hogy az Értékrend .41 egységnyivel lesz modernebb, és ezekkel együtt azt, hogy a szekularizáció .15 egységgel, a Modernizáció pedig .16 egységgel növeli a Participációt, a Társadalmi státusz egységnyi emelkedése összességében –32 egységnyivel növeli az emberek beleszólásait a döntésekbe.

A Társadalmi státusznak a súlya a Participációban a különböző generációknál:

	Útegyütthatók Direkt hatás	Teljes hatás	Indirekt hatás
20–39 évesek	.19	.29	.10
40–59 évesek	.22	.37	.15
60 évnél idősebb	.48	.48	–

Teljesítménymotiváció

A Teljesítménymotiváció látens változót, amely tartalmilag a dolgos életet jelenti és a sikeresen befejezett munkával függ össze nagy súllyal, kisebb mértékben a kitartással, a modell látens változói csak nagyon kevéssé magyarázzák, mindössze a variancia 2%-át.

Az Értékrendben történő elmozdulás az Eszme, Autonómia és Posztindusztriális értékek felé növeli .13 egységgel a Teljesítménymotiváltságot. A többi útegyüttható közül a Társadalmi státuszé kicsi, de pozitív .05, a Vallásosság is pozitívan hat (.09), a teljesítménymotivációra.

Tehát óvatosan az mondható, hogy a magasabb Társadalmi státusz elősegíti a modernebb Értékrend kialakulását, amely pozitívan hat a Teljesítménymotivációra.

Egyenlőtlenség

Az Egyenlőtlenség látens változó 12 ma még meglévő társadalmi egyenlőtlenség megszűnését, illetve csökkenését fejezi ki.

A társadalmi változókkal tulajdonképpen magyarázni nem tudjuk, egyedül az Értékrend-i modernizáció súlya nagyobb (.14), de ezzel együtt is csak 0.03-ra emelkedik fel a többszörös korreláció négyzete.

Nagyon halovány kapcsolódást találunk még a Társadalmi státusznál és a Vallásnál. Eszerint a státusz emelkedésével az egyenlőtlenség növekedését látják az emberek. A Vallásosság is az egyenlőtlenség növekedése érzetét kelti. És ez az Értékrenden keresztül már összességében .10 útegyütthatót jelent. A magasabb státusszal kapcsolódó modernebb értékrend a társadalmi státusz hatását pozitívra változtatja, és teljes hatásában a társadalmi hierarchia .09 súllyal az egyenlőtlenség csökkenése irányában alakítja a véleményeket.

Pozíció

A Pozíciót, vagyis azt, hogy az emberek hol képzelik el a maguk életét az elképzelhető legjobb élet és az elképzelhető legrosszabb élet között, a modellbe bevont változó–blokkokkal magyarázni nem tudjuk (.01 a többszörös korreláció).

Sérelem

A 11 igazságtalanságot, méltánytalanságot kifejező Sérelem látens változó érzékenyen reagál az Értékrend változására (.33).

Vagyis egységnyivel modernebb, Autonóm, Eszme (kisebb súllyal Posztindusztriális) értékrendszerrel .33 egységnyivel több esetet élünk meg igazságtalanságként, méltánytalanságként. A társadalmi státusz emelkedésével viszont csökken a méltánytalanságok súlya .16 egységgel, amit viszont a státusznak az értékrendre gyakorolt hatása semlegesít.

A boldogabb, városiasabb környezetben eltöltött gyerekkor is elősegíti, hogy bizonyos eseteket súlyos igazságtalanságként éljünk meg (a súlya .09). A Vallásosság viszont, ha nem is nagyon, de segít abban, hogy ne érezzünk nagyon súlyosnak igazságtalanságokat.

Tolerancia

A modellben a Társadalmi státusz .21-es együtthatóval befolyásolja az emberek toleranciáját. Másként fogalmazva alacsonyabb iskolai végzettséggel, falusiasabb településen élve és öregedve az emberek intoleranciája .21 egységgel fokozódik. Kedvezően hat az emberek toleranciájára az iskolázottabb családi környezet (.11) és a modernebb értékrend is (.06). Az Értékrend közvetítésével a Társadalmi státusz direkt hatása még növekszik, és totális hatásként a modellben egységnyi státusz–emelkedés a Toleranciát .30 egységgel növeli.

Tudat

Több, különböző magatartást és percepciót kifejező kérdéscsoport átlagos válaszértékeiből állítottuk össze a Tudat elnevezésű endogén változó–blokkot. A Tudat látens változó értelmezéséhez azt vesszük sorra, hogy milyen súllyal tudtuk vele előállítani a manifeszt változók átlagos értékeit. A Tudat látens változó csökkenő mértékű *pozitív súllyal* kapcsolódik: a Participációhoz a Toleranciához, az Egyenlőtlenségek csökkenéséhez, Posztmateriális célok, Vállalkozási hajlam, Teljesítménymotiváció, Bizalom, Fogyasztói elvek, csökkenő fontosságú *negatív súllyal* határozza meg: az Anómiát, a Sérelmet, az Elégedettséget. A Tudat látens változó a társadalomhoz való pozitív viszonyulását, a társadalmi változásokhoz való adaptációs készséget fejezi ki.

A társadalmi státusznak ehhez a pozitív Tudatra gyakorolt hatása a modell útegyütthatója szerint elég erős (.32). Ha a Társadalmi státuszt meghatározó együtthatókat is figyelembe vesszük, és a státusznak az Értékrenden és Valláson keresztül kifejtett indirekt hatását is, akkor a Tudat egységnyi változtatását a szocializáció irányában elérhetjük úgy, hogy megemeljük az iskolázottságot két egységgel, két egységnyivel urbanizáltabbá tesszük a településeket és egységnyivel növeljük a jövedelmeket, de egységnyi urbanizációs költséget spórolhatunk, ha egységnyivel "fiatalabbá" tesszük az embereket. A legdrágább út a tudati modernizációban, ha a személyes jövedelem emelésével próbálkozunk, mivel egységnyi iskola négy egységnyi jövedelemnek felel meg, egységyni lakóhelyváltozás az urbanizáció irányában megfelel két és félszeres jövedelem emelkedésnek. A Vallás, mint visszahuzó erő szerepel a modellben, közvetlenül –.18 együtthatójával, (teljes hatása a Tudatra –.20). A társadalmi háttér két látens változójának közvetlen súlya elég kicsi, de a totális hatás már jelentősnek mondható. A Családi státusz teljes hatása .28, a Gyerekkoré .14.

A Tudat-ot nagyon különbözően befolyásolja generációnként a modell többi változója. Az Értékrend modernizációja (a Posztindusztriális értékek is nagyobb súlyt kapva) legnagyobb mértékben a 20–39 éves generációnál hat (.22 az útegyüttható). A 40–59 éveseknél ez a hatás .14, a 60 éveseknél öregebbeknél már csak .09. Ezzel szinte párhuzamosan nő a Társadalmi státusz szerepe a Tudatra. Míg a fiatalabb generációnál a státusz közvetlen hatása gyakorlatilag nulla (a teljes hatás is csak .16), a 40–59 éveseknél a társadalmi hierarchiában felfelé haladva .35 egységgel változik a Participáció, Egyenlőtlenség csökkenése, Teljesítménymotiváció, Bizalom, magasabb Pozíció értékelés, és Posztmateriális értékek irányában az emberek tudata. A legidősebbeknél a társadalmi státusz közvetlen hatása .42.

Praktikusan a pozitív Tudatra a 20–39 éves generációnál leginkább az Értékrend modernizációjával lehet hatni, az idősebb generációnál pedig egyre inkább a Társadalmi státusz növelésével.

6.8.3 Ökológiai blokk felvétele a modellbe

Az eddigi modellekben szereplő társadalmi háttér, társadalmi státusz és értékrend, valamint más tudati, viselkedési blokkok mellé most az egyén ökológiai környezetét is felvonultatjuk magyarázó, exogén blokként. A területi adatokat az Akadémia Adatarchivuma Településsoros adatbázisából vettük. Ez az adatállomány az MTA I. Természettudományi Főosztálya és a Városépítési Tudományos és Tervező Intézet között 1980-ban kötött együttműködési megállapodás alapján jött létre és a KSH különböző adatfelvételei — Népszámlálások, rendszeres településstatisztikai adatfelvétele — nyomtatásban nyilvánosságra hozott településsoros és gépi adathordozókra rögzített adatait tartalmazza. Ebből az adatállományból 18 mutatót választottunk ki és rendeztünk saját adatbázisunkhoz, amelyek a település infrastruktúráját, foglalkozási szerkezetét, demográfiai összetételét jellemzik.

Az Ökológiai blokk látens változóját a 18 mutató csökkenő fontosság (súly) szerinti sorrendben a következőképpen állította elő:

Pozitív súllyal	Negatív súllyal
Érettségizettek aránya	
	Mezőgazdasági aktív keresők aránya
Átlagos lakásnagyság	
	Egyetemet végzettek aránya
	Ivóvízzel ellátott lakások aránya
Általános iskolát végzettek aránya	
Al- és ágybérlők száma	
	0–14 évesek aránya
Emeletes épületek aránya	
Fürdőszobával ellátott lakások száma	
	Eljárók aránya

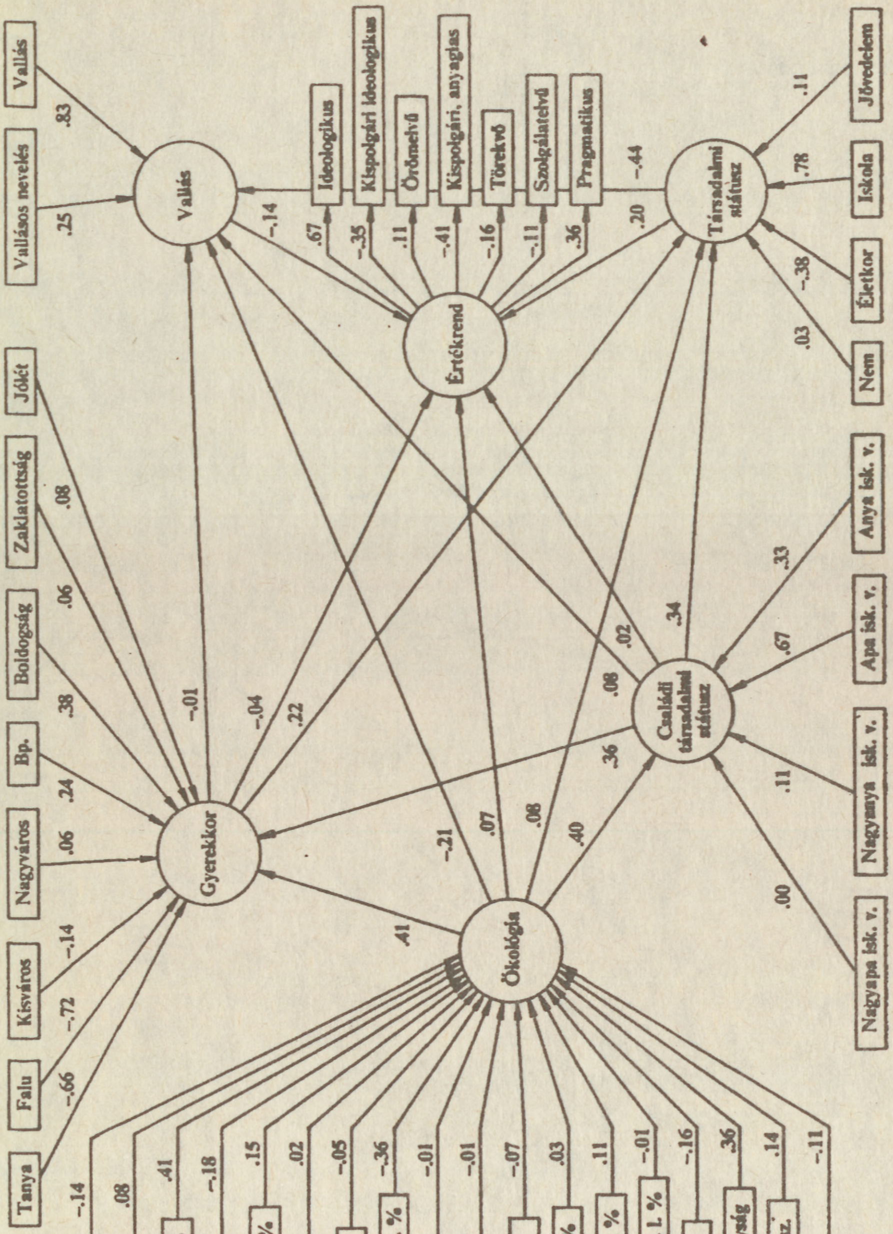
A többi mutató súlya már .1 alatt van.

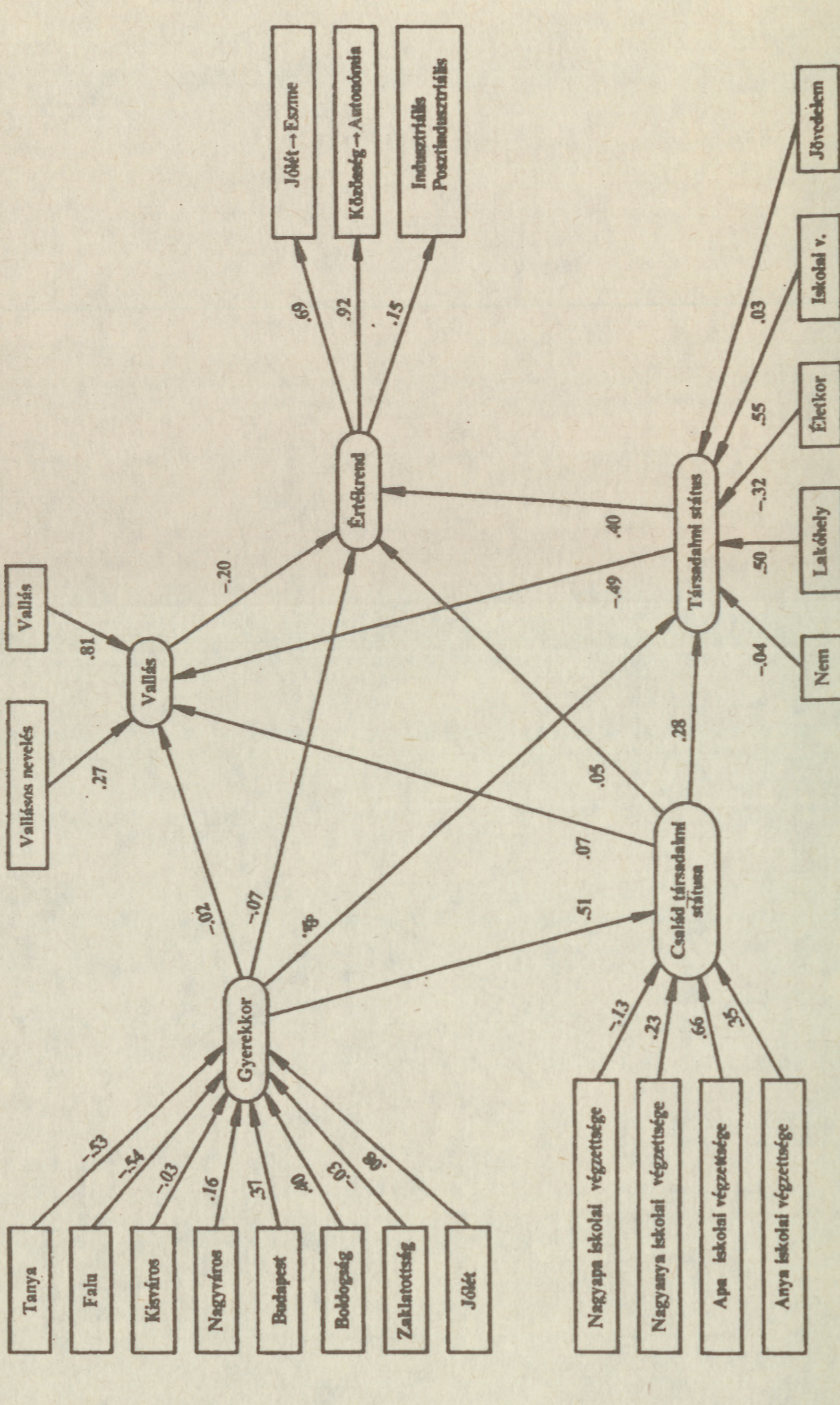
Az Ökológia látens változó pozitív felét nagyobb súllyal a legalább középiskolát végzettek aránya, infrastruktúrában az urbanizáltabb kép (emeletes épületek aránya), fürdőszobával ellátott lakások aránya és a fejlettebb kereskedelmi hálózat határozza meg, míg a negatív oldalon a Mezőgazdasági aktív kereső aránya a meghatározó, kisebb mértékű az eljárók aránya és ugyanerre az oldalra esik az egyetemet végzettek aránya és az ivóvízzel ellátott lakások aránya, vagyis a bal oldal önmagában is heterogén, legalábbis két dimenziót is magában foglal. Ez utalhat arra, hogy a modellben az ökológiát két látens változóval is képviseltethettük volna. Így mindenesetre azzal a dichotómiával jellemezzük az ökológia látens változót, hogy foglalkoztatásában a mezőgazdasági terület, amelyik a teljes foglalkoztatást nem tudja biztosítani (vagy ennek másik véglete, magasan kvalifikált népesség, magas gyerekszám és jó ivóvíz), mint lakóhelyi környezet terjed a ma nálunk már szellemi elitséget jelentő középiskolai végzettségűek magas arányáig, az ezzel együttjáró városi kultúráig, a zsúfoltsággal a városi településekig, de ezzel a magas lakáshiányig is. Rövidebben a mezőgazdasági és agglomerációs alvó–településtől a szocialista jellemzővel illethető városainkig.

Ez az Ökológiai látens változó a modell egyetlen olyan exogén változója, amelyik egyben predeterminált is. Jelentős a direkt hatása az egyén családjának társadalmi státuszára és a Gyerekkor látens változóra. Az Ökológia látens változó, mely az ember társadalmi környezetének viszonyát fejezi ki, a modellünk szerint közvetlenül az ember társadalmi backgroundjára hat leginkább (.40 és .41 útegyütthatókkal), lényegesen kisebb a közvetlen hatása az egyén társadalmi státuszára és értéktudatára (.10 és .14 útegyütthatókkal). De ha nem feledkezünk meg az indirekt hatásokról sem, akkor már a modell szerint azt állíthatjuk, hogy egységnyi emelkedése az ökológiának a társadalmi státuszon több mint egyharmadot emel, és majdnem egyharmaddal segíti az értéktudat modernizálódását. Egy teljes egységnyit az értéktudaton úgy modernizálhatunk, hogy az Ökológia egységnyi emelése mellett a Társadalmi státuszon másfél egységnyit emelünk (pl. két egységgel emeljük az iskolázottságot, vagy csak egy egységnyivel, de akkor még sok pénzzel). Ezt a kapcsolódást akkor kaptuk, amikor a modellben az értékrendet a Rokeach–értékteszttel mértük, méghozzá úgy, hogy az értékek három fő dimenziójával, a Jólét–Eszme Közösség–Autonómia, Indusztriális–Posztindusztriális dimenziókkal fejeztük ki az Értékrend látens változót. Amikor a Rokeach–tesztek három kiragadott (de nem alapvető) dichotómiáját választottuk az értékrend modernizálódásának kifejezésére, (Érzelem–Pragmatikus, Érzelem–Racionális, Hagyományos–Önálló), az Ökológia közvetlen és teljes hatása lényegesen kisebb volt (.20 és .13).

Az *Ökológia* látens változó közvetlen hatása a tudat más területeire ellentétes a *Társadalmi státusz* hatásával. Míg az *Ökológia* direkt hatásában anomikussá tesz, a *Társadalmi státusz* megfordítja ezt a hatást, és teljes hatásábban már gyengén ugyan, de csökken az Anómia (–.70).

Amikor az Anómiához még hozzávettünk más tudati változókat, az Intoleranciát, Participációt, Teljesítménymotivációt, Vállalkozási hajlamot, Bizalmat, Fogyasztói elveket, Pozíciót, Igazságosságot, Elégedettséget és Egyenlőséget, és ezek fejezték ki a *Tudat* látens változót, az Ökológia pozitív (.12-es) súlya az Anómiához és Intoleranciához kapcsolódik, és csak a Társadalmi státusz hatása közvetítésével fordul át a teljes hatása a társadalomhoz való pozitív viszonyt kifejező tudati oldalhoz.


Vallás
Vallásos nevelés
Vallás
Jólét
Zaklatottság
Boldogság
Bp.
Nagyváros
Kisváros
Falu
Tanya
Gyerekkor
Ökológia
Értékrend
Társadalmi státusz
Családi társadalmi státusz
Ideologikus
Kispolgári ideologikus
Örömelvű
Kispolgári, anyagias
Törekvő
Szolgálatelvű
Pragmatikus
Jövedelem
Iskola
Életkor
Nem
Anya isk. v.
Apa isk. v.
Nagyanya isk. v.
Nagyapa isk. v.
0–14 évesek %
60–X évesek %
Érettségizettek %
Egyetemet v.%
Általános isk. v. %
Keresők %-a
K-ből férfiak %
Mezőgazdasági k. %
Ipari k. %
Építőipari k. %
Közlekedési k. %
Kereskedelmi k. %
Emeletes épületek %
Fürdőszobával ell. l. %
Ivóvízzel ell. l. %
Átlagos lakásnagyság
Al- és ágybérlők sz.
Eljárók %


Tanya
Falu
Kisváros
Nagyváros
Budapest
Boldogság
Zaklatottság
Jólét
Gyerekkor
Vallásos nevelés
Vallás
Vallás
Értékrend
Jólét → Eszme
Közösség → Autonómia
Indusztriális Posztindusztriális
Társadalmi státus
Jövedelem
Iskolai v.
Életkor
Lakóhely
Nem
Család társadalmi státusa
Nagyapa iskolai végzettsége
Nagyanya iskolai végzettsége
Apa iskolai végzettsége
Anya iskolai végzettsége

Irodalomjegyzék

AIGNER, O.J. and GOLDBERGER, A.S.: Latent Variables in Socio-economic Models North-Holland Publishing Company 1977, 383 p.

AKAIKE, H.: Factor Analysis and AIC, Psychometrika, Vol. 52, No. 3,317-332 p.

ALKER, H.R., DEUTSCH, K.W. and SOETZEL, A.H.: Mathematical Approaches to Politics, Elsevier Scientific Publishing Company, 1973, 475 p.

ANDERSEN, E.B.: Discrete Statistical Models with Social Science Applications, North-Holland Publishing Company, Amsterdam, 1980, 383 p.

ANDERSON, J.C. and GERBING, D.W.: The effect of sampling error on convergence, improper solutions, and goodness–of–fit indices for maximum likelihood confirmatory factor analyis, Psychometrika, 1984, 49, 155–73 p.

ANDERSON, T.W.: The teaching of practical statistics, John Wiley & Sons, 1987, 199 p.

ANDERSON, T.W.: An estimation of parameters in latent structure analysis, Psychometrika, 1954, 19, 1–10 p.

ANDERSON, T.W.: Some scaling models and estimation procedures in the latent class model, In Probability and Statistics, The Herald Cramér Volume, ed. U. Grenander, Stockholm: Alwqvist & Wicksell and Wiley, 1959, 9–38 p.

ANDERSON, T.W.: An introduction to multivariate statistical analysis, 2nd edn. New York: Wiley, 1984.

ANDERSON, T.W. and RUBIN, H.: Statistical inference in factor analysis, Third Berkeley Symp. Math. Statist and Prob., 1956, 5, 111–50 p.

ANDORKA, R.: A társadalmi mobilitás változásai Magyarországon, Gondolat, 1982, 327 p.

ARATÓ, M.: Fejezetek a matematikai statisztikából számítógépes alkalmazásokkal, Számítógépalkalmazási Kutató Intézet, Közl. 22. Budapest

BARTHOLOMEW, D.J.: Latent Variable Models and Factor Analysis, London: Oxford University Press, 1987, 193 p.

BARTHOLOMEW, D.J.: Factor analysis for categorical data, J. Roy. Statist. Soc., 1980, B, 42, 293–321 p.

BARTHOLOMEW, D.J.: Mathematical Methods in Social Science, Chichester: Wiley. 1981a.

BARTHOLOMEW, D.J.: Posteriori analysis of the factor modell, Br. J. Math. Statist. Psychol., 1981b, 34, 93–9. p.

BARTHOLOMEW, D.J.: Latent variable models for ordered categorical data, J. Econometrics, 1983, 22, 229–43 p.

BARTHOLOMEW, D.J.: The foundations of factor analysis, Biometrika, 1984a, 71, 221–32 p.

BARTLETT, M.S.: Tests of significance in factor analysis, Br. J. Psychol. (Statistical Sect.), 1950, 3, 77–85 p.

BARTLETT, M.S.: Factor analysis in psychology as a statistician sees it, Uppsala Symp. Psychol. Factor Analysis, Uppsala: Almqvist & Wiksell, 1953, 23–34 p.

BENTLER, P.M. and BONETT, D.G.: Significance test and goodness-of-fit in the analysis of covariance structures, Psychological Bulletin, 1980, 88, 588-606

BENTLER, P.M.: Multivariate analysis with latent variables: causal modelling, Annu. Rev. Psychol., 1980, 31, 419–56 p.

BENTLER, P.M.: Linear systems with multiple levels and types of latent variables, in Systems Under Indirect Observation, I (eds K.G. Jöreskog and H. Wold), North Holland, Amsterdam, 1982.

BENTLER, P.M. and WEEKS, D.G.: Linear structural equations with latent variables, Psychometrika, 1980, 45, 290–308 p.

BENTLER, P.M.: Structural Modeling and Psychometrika: An Historical Perspective on Growth and Achievements, Psychometrika, 1986 Vol. 51, No. 1 March 1986 35-51 p.

BLALOCK, H.M.: Correlation and causality: the multivariate case, Social Forces, 1961, 39, 246–51 p.

BLALOCK, H.M.: Making causal inferences for unmeasured variables from correlations among indicators, Amer. J. Sociol., 1963, 69, 53–62 p.

BOCK, R.D. and AITKIN, M.: Marginal maximum likelihood estimation of item parameters: application of an EM algorithm, Psychometrika, 1981, 46, 443–59 p.

BOCK, R.D. and LIEBERMAN, M.: Fitting a response model for n dichotomously scored items, Psychometrika, 1970, 35, 179–97.

BOOLEM, K.: Sample size and Bentler and Bonett's nonnormed fit index, Psychometrika, 1986, 51, 375-377

BOOMSMA, A.: Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation, Psychometrika, 1985, 50, 229–42 p.

BORG, I. and LINGOES, J.: Multidimensional Similarity Structure Analysis, Springer-Verlag, 1987, 390 p.

BROWNE, M.W.: A comparison of factor analytic techniques, Psychometrika, 1968, 33, 267-333

BROWNE, M.W.: Covariance structures, in Topics in Applied Multivariate Analysis (ed D.M. Hawkins), Cambridge University Press, Cambridge, 1982.

BROWNE, M.W.: Asymptotically distribution–free methods for the analysis of covariance structures, Br. J. Math. Statist. Psychol., 1984, 37, 62–83 p.

BROWN, M.B. and BENEDETTI, J.: On the mean and variance of the tetrachoric correlation coefficient, Psychometrika, 1977, 42, 347-355

BRYANT E.H. and ATCHLEY, W.R.: Multivariate Statistical Methods Within-Groups Covariation, Dowden, Hutchinson and Ross, Inc., 1975, 436 p.

CAROLL, R.J.: Tranformation and weighting in regression, Chapmann and Hall Ltd., 1988, 249 p.

CAROLL, R.J.: An analytical solution for approximating simple structure in factor analysis, Psychometrika, 1953, 18, 23–38.

CATELL, R.B.: Factor Analysis, New York: Harper and Row., 1952

CHIKÁN, A., FÜSTÖS, L., PAPRIKA, Z.: Analysis of Multicriteria Decision on Capital Allocation by Multivariate Statistics, in: Papers on applications I. ed. by Gy. Meszéna, Department of Mathematics, Karl Marx University of Economics, Budapest, May 1987, 95–133 p.

CLIFF, N.: Analyzing multivariate data, San Diego: Harcourt Brace Jovanovich Corp., 1987, 494 p.

CLOGG, C.C.: Some latent structure models for the analysis of Likert–type data, Social Sci. Res., 1979, 8, 287–301 p.

CLOGG, C.C.: New developments in latent structure analysis 215-246 p., in D.J. JACKSON and E.F. Borgatta (eds.): Factor Analysis and Measurement in Sociological Research, Beverly Hill. CA: Sage, 1980

CLOGG, C.C. and GOODMAN, L.A.: Simultaneous Latent Structure Analysis in Several Groups in Sociological Methodology, 1985, 81-110 p.

COHRAN, W.G.: Sampling Techniques, John Wiley and Sons, Inc., 1963, 413 p.

COLEMAN, J.S.: Introduction to Mathematical Sociology, New York: Free Press, 1964.

COX, C.R. and HINKLEY, D.V.: Theoretical Statistics, Chapman and Hall Ltd., 1974. 511 p.

COX, D.R. and Dakes, D.: Analysis of Survival Data Monographs on Statistics and Applied Probability, Chapman and Hall Ltd., 1984. 201 p.

CRAMÉR, H.: Mathematical Methods of Statistics, Princeton University Press, 1966. 574 p.

CRISTOFFERSSON, A.: Factor analysis of dichotomized variable, Psychometrika, 1975, 40, 5-32.

CRONBACH, L. J.: Coefficiant alpha and the internal structure of test, Psychometrika, 1951, 16, 297-334.

CRONBACH, L. J.: Internal Consistency of Tests: Analysis Old and New, Psychometrika, Vol. 53, No. 1, 63-70 p.

CSEH-SZOMBATHY L.: Családszociológiai problémák és módszerek, Gondolat, Budapest, 1979, 402 p.

CSEH-SZOMBATHY L. – FERGE ZS.: A szociológiai felvétel módszerei, Közgazdasági és Jogi Könyvkiadó, Budapest, 1975.

CSEH-SZOMBATHY L. – LÉDERER P. (szerk.): Az empirikus szociológiai kutatás statisztikai alapjai, Tankönyvkiadó, Budapest, 1973, 280 p.

D'AGOSTON, R.B. and STEPHENS, M.A.: Goodness-of-fit Techniques, Statistics: textbooks and monograps Vol. 68., Marcel Dekker, Inc., 1986, 560 p.

DAVIDSON, M.L.: Multidimensional Scaling, John Wiley and Sons, Inc., 1983. 242 p.

DAVIS, P.M. and COXON, A.P.M.: (ed.) Key Tests in Multidimensional Scaling, Heinemann Educational Books, London, 1982, 347 p.

DERRY, W.D. and LEWIS-BECK, M.S.: New Tools for Social Scientists, Sage Pub., 1986

DESARBO, W.S. and CAROLL, J.D.: Three–way Metric Unfolding Via Alternating Weighted Least Squares, Psychometrika, Vol. 50, No. 3, 275-300 p. September 1985

DILLON W.R. and GOLDSTEIN, M.: Multivariate Analysis Methods and Applications, John Wiley and Sons, Inc., 1984. 587 p.

DRAPER, N.R. and SMITH, H.: Applied Regression Analysis, John Wiley and Sons, Inc., 1966, 407 p.

EATON, M.L.: Multivariate Statistics, A Vector Space Approach, John Wiley and Sons., Inc., 1983, 512 p.

ÉLTETŐ ÖDÖN – ZIERMANN, M.: Matematikai statisztika, Tankönyvkiadó, Budapest, 1961.

EVERITT, B.S.: The Analysis of Contingency Tables, Monographs on Statistics and Applied Probability, Chapmann and Hall Ltd., 1977, 128 p.

EVERITT, B.S.: An Introduction to Latent Variable Models, Monographs on Statistics and Applied Probability, Chapmann and Hall Ltd., 1984, 107 p.

FEINBERG, S.E.: The Analysis of Cross-Classified Categorical Data, The Massachusetts Institute of Technology, 1980, 198 p.

FIELDING, A.: Latent Structure Models, 1977, 125-57 p., in C. PAYNE and C.A. O'MUIRCHEARTAIGH (eds.), The Analysis of Survey Data Volume I., Exploring data Structures, New York: John Wiley & Sons, Inc.

FINNEY, D.J.: Probit Analysis, Cambridge University Press, 1977, 333 p.

FISCHER, G.H.: Logistic latent trait models with linear constraints, Psychometrika, 1983, 48, 3–26 p.

FLESS, J.L.: Statistical Methods For Rates and Proportions, New York: John Wiley & Sons, Inc., 1973

FLETCHER, R. and POWELL, M.J.D.: A rapidly convergent descent method for minimization, Comput. J., 1963, 2, 163–68.

FORGÁCSNÉ KOVÁCS ERZSÉBET – TÖRÖKNÉ MATITS Á.: Gazdasági adatrendszerek struktúrájának elemzése, Tankönyvkiadó, Budapest, 1986, 132 p.

FORMANN, A.K.: The Latent class analysis of polytomous data, Biometrical Journal, 1978, 20, 755-771, A note on parameter estimation for Lazarsfeld's latent class analysis, Psychometrika, 1978, 43, 123-126, Linear logistic latent class analysis, Biometrical Journal, 1982, 24, 171-190

FORMANN, A.K.: Constrained latent class models: Theory and applications, British Journal of Mathematical and Statistical Psychology, 1985, 38, 87-111.

FORMANN, A.K.: Latent Class Models for Nonmonotone Dichotomous Items, Psychometrika, Volume 53, Number 1, 45-63 p.

FULLER, E.L. and HEMMERLE, W.J.: Robustness of the maximum–likehood estimation procedure in factor analysis, Psychometrika, 1966, 31, 255-266.

FULLER, W.A.: Measurment Error Models, John Wiley and Sons, 1987, 440 p.

FÜSTÖS, L.: Szociológiai kutatások sokváltozós matematikai statisztikai módszerei, MTA Szociológiai Kutató Intézet Kiadványai, Budapest, 1979, 220 old.

FÜSTÖS, L.: Methods of Measuring Characteristics of Distributions, Budapest, Center for Value Sociology, Institute for Sociology, The Hungarian Academy of Sciences, Institute for Culture, 1980, 60 p.

FÜSTÖS, L., MANCHIN, R., TÓTH K.: SZOCPROG 1.3 verzió, Társadalomstatisztikai programrendszer, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1981, 45 old.

FÜSTÖS, L.: Klaszterelemzés, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1982, 54 old.

FÜSTÖS, L., PAPRIKA, Z.: Innováció nemzetközi mércével, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1982, 25 old.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: Cluster Analysis, Szigma, 1977, X. évf. 3. sz.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: Cluster Analysis, Acta Oeconomics, Vol. 26, (3.4) 291–334 p.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: A sokdimenziós skálázás egyes újabb módszerei I., Szigma, 1982, XV. évf. 3. sz.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: A sokdimenziós skálázás egyes újabb módszerei II., Szigma, 1983, XVI. évf. 3. sz.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: A sokdimenziós skálázás egyes újabb módszerei III., Szigma, 1986, XIX. évf. 1–2. sz.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: Bevezetés az adatelemzés sokváltozós módszereibe, Tankönyvkiadó, Budapest, 1983, 265 old.

FÜSTÖS, L.: Sokdimenziós skálázás módszerei: MINISSA, INDSCAL, PREFMAP, PROFIT, PARAMAP, MRSCAL, HICLUS, MINIRSA, MDPREF, UNICON, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1980–1983, 341 old. (10 füzet)

FÜSTÖS, L.: Lineáris egyenleterendszerek általános modelljei, LISREL, LVPLS, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1983, 161 old.

FÜSTÖS, L.: Három kívánság I–II, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1984, 98 old.

FÜSTÖS, L., MÓNUS, Z.: Bevezetés az adatelemzésbe, Szocprog–PC 1.1' verzió, Társadalomstatisztikai programrendszer személyi számítógépre, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1984, 283 old.

FÜSTÖS, L.: Loglineáris modell kontingencia táblák elemzésére, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1985, 54 old.

FÜSTÖS, L., MESZÉNA, GY., SIMONNÉ MOSOLYGÓ N.: Bevezetés az adatelemzés sokváltozós módszereibe, Akadémiai Kiadó, Budapest, 1985, 265 old.

FÜSTÖS, L., MESZÉNA, GY., RESS, S., SIMONNÉ MOSOLYGÓ N.: LISREL — Das allgemeine Lineare Modell von Strukturgleichungen, in Faktoranalyse, Beitrage zum 2. Anwenderseminar in Jena, Friedrich–Schiller–Universitat, Jena, 1987, 33–69 p.

FÜSTÖS, L., MESZÉNA, GY., RESS, S., SIMONNÉ MOSOLYGÓ N.: Strukturális kapcsolatok általános lineáris modellje (LISREL), Szigma, 1987–88, XX. évf. 1. szám

FÜSTÖS, L.: Az adatelemzés statisztikai módszerei, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1988, 587 old.

FÜSTÖS, L.: Az exploratív faktorelemzés módszerei, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1988, 55 old.

FÜSTÖS, L.: 'Ertéktérkép' (16 ország értéktérképe a gyermeknevelési elvek tükrében), Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1988, 44 old.

FÜSTÖS, L., KÖNYVES TÓTH, I.: Gyermeknevelési elvek, (A magyar társadalom és Kővágóörs, egy helyi társadalom–értéktérképének összehasonlító vizsgálata.), Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1988, 88 old.

FÜSTÖS, L.: The methods of exploratory factor analysis, Módszertani füzetek, MTA Szociológiai Kutató Intézet, Budapest, 1988, 51 old.

FÜSTÖS, L., KOVÁCS, E. Számítógépes adatelemzés statisztikai módszerei, Tankönyvkiadó, Budapest, 1989, 587 old.

GIBSON, W.A.: Three multivariate models: factor analysis, latent structure analysis and latent profile analysis, Psychometrika, 1959, 24, 229–52 p.

GOODMAN, L.A.: Exploratory latent structure analysis using both identifiable and unidentifiable models, Biometrika, 1974, 61, 215–31 p.

GOODMAN, L.A.: The analysis of systems of qualitive variable when some of the variables are unobservable. Part I., A modofied latent structure approach, American Journal of Sociology, 1974, 79, 1179-1259 p.

GOODMAN. L.A.: Analyzing Qualitative/Categorical Data, Long-linear models and Latent-Structure Analysis, Addison-Wesley Publishing Company, London, 1978, 467 p.

GOODMAN L.A.: The Analysis of Cross–Classified Data, Having Ordered Categories, Harvard University Press, Cambridge 1984, 414 p.

GORSUCH, R.L.: Factor analysis, Lawrence Erbaum Associates Publishers, London, 1983.

GREEN, B.F.: A general solution of the latent class model of latent structure analysis and latent profil analysis, Psychometrika, 1951, 6, 151-166

GREENACRE, M.J.: Theory and applications of correspondence analysis, London: Academic Press, 1984

GUTTMAN, L.: A new approach to factor analysis: The Radex, Chapter in Paul F. Lazarsfeld (ed.), Mathematical Thinking in the Social Sciences, New York: Columbia Univ. Press, 258–348 p.

GUTTMAN, L.: The determinancy of factor score matrices with implications for five other basic problems of common factor theory, Br. J. Statist. Psychol., 1955, 8, 65–82 p.

HABERMAN, S.J.: Analysis of Qualitative Data, Volume 2: New Developments, Academic Press, 1979, 612 p.

HAGENAARS, J.A.: Latent Structure models with Direct Effects Between Indicators: Local Dependence Models in Sociological Methods and Research, 1988, Vol. 16, No. 3, 379-405, Sage Publications, Inc.

HANKISS, E., MANCHIN, R., FÜSTÖS, L., SZAKOLCZAI, Á.: Kényszerpályán? MTA Szociológiai Kutató Intézet, Értékszociológiai Műhely kiadványai, Budapest, 1982, 382 old.

HANKISS, E., MANCHIN, R., FÜSTÖS, L., SZAKOLCZAI, Á.: Folytonosság és Szakadás, MTA Szociológiai Kutató Intézet, Értékszociológiai Műhely kiadványai, Budapest, 1982, 604 old.

HANKISS, E., MANCHIN, R., FÜSTÖS, L., SZAKOLCZAI, Á.: The Role of Values in Various Cultural Contexts, UNESCO kiadványai, Budapest–Paris, 1982, 84 p.

HANKISS, E., MANCHIN, R., FÜSTÖS, L., SZAKOLCZAI, Á.: Modernization of Value Systems, Budapest–Vienna, 1982, (Symp. on Cultural Indicators).

HANKISS, E., MANCHIN, R., FÜSTÖS, L.: The Role of Goals in People's Lives, UNESCO kiadványai, Budapest–Paris, 1980, 103 p..

HANKISS, E., MANCHIN, R., FÜSTÖS, L.: The Role of Values and Value Deficiencies in Primary Health Care Recording Systems, in: M. Lipkin–Karel Kupka: Psychosocial Factors, Affecting Health, New York, 1982.

HANKISS, E., MANCHIN, R., FÜSTÖS, L.: Cross-National QOL Research, An Outline for a Conceptual Framework, UNESCO kiadványai, Budapest–Paris, 1981, 102 p.

HARMAN, H.H.: Modern factor analysis, Chicago: University of Chicago Press, 1976.

HARRIS, C.W.: Relationships between two systems of factor analysis, Psychometrika, 1956, 21, 267-333.

HARRIS, C.W.: Canonical factor models for the description of change, in C.W. HARRIS (Ed.), Problems in measuring change, Madison: University of Visconsin Press, 1963

HARRIS, C.W.: Some Rao-Guttmann relationships, Psychometrika, 1962, 27, 247-263

HARRIS, C.W. and KAISER, H.F.: Oblique factor analytic solutions by orthogonal transformations, Psychometrika, 1964, 29, 347–62 p.

HARRIS, C.W.: On factors and factor scores, Psychometrika, 1967, 32, 363-379

HAWKINS, D.M.: Identification of Outhiers, Monographs on Applied Probability and Statistics, London, New York 1980, 188 p.

HEARN, D. and BAKER, M.P.: Computer Graphics, Prentice-Hall International, 1986, 352 p.

VAN DER HEIJDEN, P.G.M. and DE LEEUW, J.: Correspondence analysis and Complementary to Loglinear Analysis, Psychometrika, 1985, 50, 429-447.

HEYWOOD, H.B.: On finite sequences of real numbers, Proc. Roy. Soc., 1931, Ser. A, 134, 486–510 p.

HIDY PÉTER – KOVÁCS ERZSÉBET: A lokális döntések természetéről, Művelődéskutató Intézet, Budapest, 1986, 163 p.

HOPPÁL, M. és SZECSKŐ T.: Értékek és változások, I. II., Tömegkommunikációs Kutatóközpont, Budapest, 1987, 277 p.

HORST, P.: Factor Analysis of Data Matrix, New York, Molt, Rinehart and Vinston, 1965

HOTELLING, H.: Analysis of a complex of statistical variables into principal components, J. Educ. Psychol., 1933, 24, 417–41 and 498–520 p.

HOWE, W.G.: Some Contributions to Factor Analysis, Oak Ridge National Laboratory, Oak Ridge, 1955.

IHARA, M. and KANO, Y.: A New Estimator of the Uniquenesss in Factor Analysis, Psychometrika, Vol. 51, No. 4, December 1986, 563-566 p.

JÁNOSSY, L.: A valószínűségelmélet alapjai és néhány alkalmazása — különös tekintettel mérési eredmények kiértékelésére, Tankönyvkiadó, Budapest, 1965, 206 p.

JÁNOSSY, L.: Mérési eredmények kiértékelésének elmélete és gyakorlata, Akadémiai Kiadó, Budapest, 1968, 527 p.

JENNRICH, R.I. and ROBINSON, S.M.: A Newton–Raphson algorithm for maximum likelihood factor analysis, Psychometrika, 1969, 34, 111–23 p.

JOHNSON, R.A. and WICHERN, D.W.: Applied Multivariate Analysis, Englewood Cliffs, N.J.: Prentice–Hall, 1982.

JÖRESKOG, K.G.: A general approach to confirmatory maximum likelihood factor analysis, Psychometrika, 1969, 34, 183–220 p.

JÖRESKOG, K.G.: A general method for analysis of covariance structures, Biometrika, 1970, 57, 239–51 p.

JÖRESKOG, K.G.: Simultan Factor Analysis in Several Populations, Psychometrika, Vol., 1971, 36, 409-26 p.

JÖRESKOG, K.G.: Statistical Analysis of Congeneric Test, Psychometrika, Vol. 1971, 36, 109-33 p.

JÖRESKOG, K.G.: Some contributions to maximum likelihood factor analysis, Psychometrika, 1967, 32, 443-482 p.

JÖRESKOG, K.G. and GOLDBERGER, A.S.: Factor analysis by generalized least squares, Psychometrika, 1972, 37, 243-260 p.

JÖRESKOG, K.G.: General Method for Estimations a Linear Structural Equation Systems, 1973, 85-122 p., in A.S. GOLDBERGER and O.D. DUNCAN (eds), Structural Equation Models in the Social Sciences, New York: Seminar Press

JÖRESKOG, K.G.: Structural equation models in the social sciences: Specification, estimation and testing, In P.R. Krishnaiah (ed.), Applications of Statistics, Amsterdam: North–Holland, 1977.

JÖRESKOG, K.G.: Basic ideas of factor and component analysis, In K.G. Jöreskog and D. Sorbom (ed), Advances in Factor Analysis and Structural Equation Models, Cambridge, Mass: Abt Books, 1979.

JÖRESKOG, K.G. and GOLDBERGER, A.S.: Factor analysis by generalized least squares, Psychometrika, 1972, 37, 243–59 p.

JÖRESKOG, K.G. and SÖRBOM, D.: Statistical models and methods for analysis of longitudinal data, In Latent Variables in Socioeconomic Models, ed. D.J. Aigner and A.S. Goldberger, Amsterdam: North–Holland, 285–325 p.

KAISER, H.F.: Image analysis, in C.W. HARRIS (Ed.), Problems in measuring change, Madison: University of Wisconsin Press, 1963

KAISER, H.F.: The varimax criterion for analytic rotation in factor analysis, Psychomterika, 1958, 23, 187–200 p.

KENDALL, M.G.: A Course in Multivariate Analysis, London: Griffin, 1957, 1975.

KENDALL, M.G. and BABINGTON SMITH, B.: Factor analysis, J. Roy. Statist. Soc., 1950, B, 12, 60–94 p.

KENDALL, M.G. and LAWLEY, D.N.: The principles of factor analyis, J. Roy. Statist. Soc., 1956, A, 119, 83–4 p.

KINDLER, J. – PAPP, O.: Komplex rendszerek vizsgálata, Műszaki Könyvkiadó, Budapest, 1977, 262 p.

KISH, L.: Statistical desing for research, John Wiley and Sons, Inc., 1987, 267 p.

KLEIJEN, J.P.C.: Statistical Tools for Simulation Practitioners, Marcel Dekker, Inc., New York and Basel 1987, 429 p.

VAN DE KLOOT, W.A. and KROONENBERG, P.M.: External Analysis with Three-Mode Principal Component Models, Psychometrika, Vol. 50, No. December 1985 4, 479-494 p.

KOEMLER, K.: Goodness-of-fit test for log-linear models in sparse contingency tables, J. of the Amer. Stat. Asso., 1986, 81. 483-493 p.

KOLOSI, T.: Státusz és réteg, in Rétegeződés-modell vizsgálat III, MSZMP KB Társadalomtudományi Intézet, Budapest, 1984, 280 p.

KOLOSI, T. és RUDAS T.: Empirikus problámamegoldás a szociológiában, OMIKK-TÁRKI, Budapest, 1988, 213 p.

KOOPMAN, R.F.: On Bayesian estimation in unrestricted factor analysis, Psychometrika, 1978, 43, 109–10 p.

KRANE, W.R. and MCDONALD, R.P.: Scale invariance and the factor analysis of correlation matrices, Br. J. Math. Statist. Pasychol., 1978, 31, 218–28 p.

LAZARSFELD, P.F. and N.W. HENRY: Latent Structure Analysis, Boston, Houghton–Mifflin, 1968

LAWLEY, D.N. and MAXWELL, A.E.: Factor Analysis as a Statistical Method, London Butterworths, 1971, 153 p.

LEE, H.B. and COMREY, A.L.: An empirical comparison of two minimum residual factor extraction methods, Multivariate Behavioral Res., 1978, 13, 497–507 p.

LEE, S.–Y.: A Bayesian approach to confirmatory factor analyis, Psychometrika, 1981, 46, 153–60 p.

LEIK, R.K. and MEEKER, B.T.: Mathematical Sociology, Prentice-Hall, Inc., 1975, 242 p.

LOEHLIN, I.C.: Latent Variable Models: An Introduction to Factor, Path and Structural Analysis, Hillsdale, N.J.: Erlbaum, 1987, 273 p.

MACCULLEN, R.: A comparison of factor analysis programs in SPSS, BMDP, and SAS, Psychometrika, 1983, 48, 223–31 p.

MARIDA, K.V. and KENT, J.T. and BIBBY, J.M.: Multivariate Analysis, Academic Press, Inc., 1982, 518 p.

MARTON Á.: Robusztusság a statisztikában, Statisztikai Szemle 60, 1982, 8–9, 905–909 p.

MARTON, Á. – VINCZE, I.: A matematikai statisztika a gazdasági és társadalmi jelenségek vizsgálatában, Statisztikai Szemle 61, 1983, 1, 43–58 p.

MASTERS, G.N.: A Comprasion of Latent Trait and Latent Class Analysis of Likert–type Data, Psychometrika, Vol. 50, No. 1, March 1985, 69-82 p.

MASTERS, G.N. and WRIGHT, B.D.: The essential process in a family of measurement models, Psychometrika, 1984, 49, 529-544 p.

MAXWELL, A.E.: Multivariate Analysis in Behavioural Research, Chapman and Hall, London, 1977.

MCCULLAGH, P. and NELDER, J.A.: Generalized Linear Models, Monographs on Statistics and Applied Probability, Chapman and Hall Ltd., 1983. 261 p.

MCCUTCHEON, A.: Latent Class Analysis, Beverly Hills, CA: Sage, 1967

MCDONALD, R.P. and BURR, E.J.: A Comparsion of four Methods of Constructing Factor Score, Psychometrika, 1967, 32, 381-401 p.

MCDONALD, R.P.: The measurement of factor indeterminancy, Psychometrika, 1974, 39, 203-222 p.

MCDONALD, R.P.: Factor Analysis and Related Methods, Lawrenc Erlbaum Associates Publisher, 1985, 259 p.

MESZÉNA GY. és ZIERMANN M.: Valószínűségelmélet és matematikai statisztika, Közgazdasági és Jogi Könyvkiadó, Budapest, 1981, 554 p.

MESZÉNA, GY. (szerk.): Sztochasztikus módszerek a döntéselőkészítésben, Tankönyvkiadó, Budapest, 1984, 252 p.

MILL, M.O.: Correspondence Analysis, A neglected multivariate method, Applied Statistics, 1974, 23, 340-354 p.

MISLEVY, R.J.: Recent developments in the factor analysis of categorical variables, J. of Educational Statistics, 1986, 11, 3-31 p.

MISLEVY, R.J.: Estimating latent distributions, Psychometrika, 1984, 49, 359–81 p.

MISLEVY, R.J.: Estimation of latent group effects, J. Am. Statist. Assoc., 1985, 8O, 993–7 p.

MOOIJAART, A.: Two kinds of factor analysis for ordered categorical variables, Multivariate Behavioural Res., 1983, 18, 423–41 p.

MOOIJAART, A.: Factor analysis for non–normal variables, Psychometrika, 1985, 50, 323–42 p.

MORRISON, D.F.: Multivariate Statistical Methods, New York: McGraw–Hill, 1967.

MÓRI, F.T. és SZÉKELY J. GÁBOR: (Szerk.) Többváltozós statisztikai analízis, Műszaki Könyvkiadó, Budapest, 1986, 393 p.

MULAIK, S.A.: The Foundations of Factor Analysis, New York: McGraw–Hill, 1972.

MULAIK, S.A.: Factor analysis and Psychometrika: major developments, Psychometrika, 1986, 51, 23–33 p.

MULAIK, S.A. and MCDONALD, R.P.: The effect of additional variables on factor indeterminacy in models with a single common factor, Psychometrika, 1978, 43, 177–92 p.

MUNCK, I.M.E.: Model Building in Comparative Eduction, Applications of tha LISREL Method to Cross-National Survey Data, International Association for the Evaluation of Educational Achievement IEA Monograph Studies No. 10, Almqvist and Wiksell International, Stockholm, 1979, 199 p.

MUNDRUCZÓ GY.: Sztochasztikus modellek közgazdasági alkalmazásának kérdései, különös tekintettel a mérési hibákra, Kandidátusi értekezés, Budapest, 1979, 155 p.

MUNDRUCZÓ GY. – KERÉKGYÁRTÓ GYNÉ: Alkalmazott regresszió számítás, Tankönyvkiadó, Budapest, 1975, 193 p.

MUTHÉN, B.: Contributions to factor analysis of dichotomous variable, Psychometrika, 1978, 43, 551-560 p.

MUTHÉN, B.: A structural probit model with latent variables, J. Amer. Statist. Ass., 1979, 74, 807–11 p.

MUTHÉN, B. and A. CHRISTOFFERSSON: Simultaneous factor analysis of dichotomous variables in several groups, Psychometrika, 1981, 46, 485-500 p.

MUTHÉN, B.: A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators, Psychometrika, 1984, 46, 115-132 p.

MUTHÉN, B. and KAPLAN, D.: A comporison of some methodogics for the factor analysis of non–normal likert variables, British Journal of Mathematical and Statistical Psychology, 1985, 38, 171-189 p.

NAMBOODIRI, N.K. and CARTER, L.F. and BLALOCK, H.M.: Applied Multivariate Analysis and Experimental Design, McGraw-Hill, Inc., 1975, 688 p.

O'MUIRCHEARTAIGH, C.A. and PAYNE, C. (Ed.): Analysis of Survey Data, Volume 1: Exploring Data Structures, (A. FILDING: Latent Structure Models, Chapter 5, 125-157 p.), Volume 2: Model Fitting, John Wiley & Sons, Inc., 1977, 273 p., 255 p.

ORLÓCI, L. and KENEKEL, N.C.: Introduction to Data Analysis with Applications in Population and Community Biology, UWO Biology 224 A, 352 B. The University of Western Ontario, London, Ontario, 1983.

ORLÓCI, L.: Multivariate Analysis in Vegetation Research, The University of Western Ontario, Dr. W. Junk B.V. – Publishers – The Hague, 1975, 276 p.

PLEWIS, J.: Analysing Change, Measurement and Explanation Using Longitudinak Data, John Wiley & Sons, 1985, 182 p.

PRÉKOPA, A.: Valószínűségelmélet műszaki alkalmazásokkal, Műszaki Könyvkiadó, Budapest, 1972.

PRÉKOPA, A. és ÉLTETŐ ÖDÖN: Matematikai Jegyzetek, IV. rész. Matematikai Statisztika, Statisztikai Kiadó, Budapest, 1961.

RAO, C.R.: Estimation and tests of significance in factor analysis, Psychometrika,, 1955, 20, 93–111 p.

RASH, G.: Probabilistic models for some intelligence and attainments test (2nd ed.), Chicago: University of Chicago Press, 1980

RÉNYI, A.: Valószínűségszámítás, Tankönyvkiadó, Budapest, 1954.

ROST, J.: A Latent Class Model for Rating Data, Psychometrika, Vol. 50, No. 1, March 1985, 37-49 p.

SARIS, W.E. and STRONKHORST, L.M.: Causal modelling in nonexperimental research: An introduction to the LISREL approach, Amsterdam: Sociometric Research Foundation, 1984

SCHÖNEMANN, P.H. and WAND, M.: Some new results on factor indeterminancy, Psychometrika, 1972, 37, 61-91 p.

SEARLE, S.R.: Linear Models, John Wiley & Sons, Inc., 1971, 532 p.

SEBER, G.A.F.: Multivariate Observations, John Wiley & Sons, Inc., 1984, 686 p.

SIKOS, T.T. (szerk.): Matematikai és statisztikai módszerek alkalmazási lehetőségei a területi kutatásokban, Akadémiai Kiadó, Budapest, 1984, 301 p.

SINGER, B.: Grade of membership representations concepts and problems, in: ANDERSON, T.W. and K.B. ATHREYA (eds.), Festschrift für Samuel Karlin, New York: Academic Press, 1989

SÖRBOM, D.: A General Method for Studying Differences in Factor Means and Factor Structure Between Groups, British Journal of Mathematical and Statistical Psychology, 1974, Vol. 27, 229-239 p.

SPATH, H.: Cluster Analysis Algorithms, Computers and Their Applications, Ellis Horwood Limited, 1980, 221 p.

SVÁB J.: Többváltozós módszerek a biometriában, Mezőgazdasági Kiadó, Budapest, 1979, 221 p.

TAKANE, Y. and LEEUW, J.: On the Relationship between Item Response Theory and Factor Analysis of Discretized Variables, Psychometrika, Vol. 52, No. 3, 393-408 p.

TAKÁCS L. – ZIERMANN, M.: Valószínűségszámítás, Tankönyvkiadó, Budapest, 1954.

TELEGDI L.: Bináris változók struktúrájának vizsgálata, Alkalmazott Matematikai Lapok 13(1987–88) 17–42 p.

TIKU, M.L. TAN and BALAKRISHNAN, N.: Robust Inference Statistics: textbooks and monographs, Vol. 71. Marcel Dekker, Inc., 1986, 321 p.

TUKEY, J.W.: Exploratory Data Analysis, Addison–Wesley Publishing Company, 1977, 499 p.

UPTON, G.J.G.: Analysis of Cross–tabulated Data, John Wiley and Sons, Inc., 1978. 148 p.

VELICER, W.F.: The empirical comparison of the similarity of principal component, image and factor patterns, Multivariate Behavioral Research, 1977, 12, 3–22 p.

VINCE, I.: Matematikai statisztika ipari alkalmazásokkal, Műszaki Könyvkiadó, Budapest, 1968.

WILEY, D.E.: The identification problem for structural equation models with unmeasured variables, in Structural Equation Models in the Social Sciences (eds A.S. Goldberger and O.D. Ducan), Seminar Press, New York, 1973.

WOODBURY, H.A. and MANTON, K.G.: Grade of Membership Analysis of Depression-related Psychiatric Disorders, Sociological methods and research, Vol. 18 No 1. August, 1989, Sage Publications, Inc.

WRIGHT, B.D. and MASTERS, G.N.: Rating scale analysis, Chicago: MESA Press, 1982

WRIGLEY, N.: Categorical Data Analysis for Geographers and Environmental Scientists, Longman, 1985, 392 p.

YANA, H. and MUKHERJEE, D.N.: A Generealized Method of Image Analysis from an Intercorrelation Matrix which may be Singular, Psychometrika, Vol. 52, No. 4, 555-564 p.

Melléklet

I. Matematikai összefoglaló

A következőkben összefoglaljuk azokat a matematikai alapfogalmakat, amelyek a látens változós módszerek megértéséhez elengedhetetlenül szükségesek.

I.1 Mátrixaritmetikai műveletek

A mátrix a számoknak olyan halmaza, amely sorokba és oszlopokba van rendezve. Az n sorból és m oszlopból álló mátrixot $(n \times m)$ típusúnak nevezzük.

Pl. egy (3×4) típusú mátrix:

$$\underline{X} = \begin{pmatrix} 6 & 8 & 3 & 4 \\ 2 & 5 & 1 & 0 \\ 3 & 1 & 4 & 9 \end{pmatrix} .$$

A mátrix jelölésére aláhúzott negybetűket használunk. Az $\underline{X}$ mátrix i-edik sorának j-edik elemét x_{ij}-vel jelöljük.

Az elemek különböző elrendezései szerint a következő speciális mátrixokról beszélünk:

a) A nullmátrix: minden eleme zérus. Jele $\underline{0}$.

b) A kvadratikus mátrix: n sorból és n oszlopból áll.

c) A diagonális mátrix: olyan kvadratikus mátrix, amelynek a főátlón kívül minden eleme zérus.

d) Az egységmátrix olyan diagonális mátrix, amelynek minden főátlóbeli eleme egyes (1). Jele: $\underline{I}$.

Egy $\underline{X}$ mátrix transzponáltján azt az $\underline{X}'$-vel jelölt mátrixot értjük, amelyet $\underline{X}$-ből úgy kapunk, hogy sorait rendre felcseréljük oszlopaival.

e) Szimmetrikus a mátrix: ha egyenlő a saját transzponáltjával. ($\underline{X} = \underline{X}'$, vagyis $x_{ij} = x_{ji}$).

Az olyan mátrixot, amely egy sorból vagy egy oszlopból áll, vektornak nevezzük. Az $(1 \times m)$ típusú mátrixot sorvektornak ($\underline{x}'$), az $(n \times 1)$ típusút pedig oszlopvektornak $\underline{x}$ nevezzük (és aláhúzott kisbetűvel jelöljük).

g) Az egységvektor: egy eleme egyes, a többi zérus. Jele: $\underline{e}_i$, ahol i mutatja az egyes elem sorszámát (helyét).

h) Az összegzővektor: minden eleme egyes. Jele $\underline{1}$.

Összeadás — kivonás

Ha $\underline{A}$ és $\underline{B}$ azonos típusú mátrixok, akkor a két mátrix összegén azt a mátrixot értjük, amelyet úgy kapunk, hogy $\underline{A}$ és $\underline{B}$ megfelelő elemeit összeadjuk.

$$\underline{C} = \underline{A} + \underline{B}, \qquad \text{ahol} \quad c_{ij} = a_{ij} + b_{ij}.$$

A kivonás hasonló módon értelmezhető:

$$\underline{D} = \underline{A} - \underline{B}, \qquad \text{ahol} \quad d_{ij} = a_{ij} - b_{ij}.$$

Skalárral való szorzás

Egy adott $\underline{A}$ mátrixnak egy λ skalárral alkotott szorzatán azt a mátrixot értjük, amelynek minden eleme az $\underline{A}$ mátrix megfelelő elemének λ-szorosa.

Pl.

$$3\begin{pmatrix} 1 & 2 \\ 4 & -3 \\ 0 & 7 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 12 & -9 \\ 0 & 21 \end{pmatrix}.$$

Szorzás

Két vektor skaláris szorzata

Ha $\underline{a}$ és $\underline{b}$ azonos elemszámú vektor, akkor skaláris szorzatukon azt a skalárt értjük, amelyet úgy kapunk, hogy a két vektor megfelelő elemeinek szorzatát összeadjuk.

Vagyis:

$$\underline{a}'\underline{b} = [a_1, a_2, \dots a_n] \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + \dots +$$

$$+ \dots + a_n b_n = \sum_{i=1}^{n} a_i b_i .$$

Az $\underline{a}$ és $\underline{b}$ vektor (nem kell azonos elemszámúnak lenni) diadikus szorzatán olyan mátrixot értünk, melynek i-edik sora:

$$a_i[b_1, b_2, \dots, b_m] = [a_i b_1, a_i b_2, \dots a_i b_m],$$

és a teljes mátrix:

$$\underset{(n\times 1)(1\times m)}{\underline{a}\underline{b}'} = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_m \\ a_2b_1 & a_2b_2 & \dots & a_2b_m \\ a_nb_1 & a_nb_2 & \dots & a_nb_m \end{pmatrix} = \underset{(n\times m)}{\underline{C}} \ .$$

A két mátrix szorzata

Két mátrix szorzata csak abban az esetben van értelmezve, ha az első tényező mátrix oszlopainak a száma megegyezik a második tényező mátrix sorainak a számával. Legyen $\underline{A}$ mátrix $(n\times m)$ típusú és $\underline{B}$ mátrix $(m\times k)$ típusú. A két mátrix szorzataként kapott $\underline{C} = \underline{A}\,\underline{B}$ mátrix $(n\times k)$ típusú mátrix, melynek általános eleme:

$$c_{ij} = \underline{a}_i'\underline{b}_j = \sum_{g=1}^{m} a_{ig}b_{gj}\,, \qquad \text{vagy szavakban:}$$

a $\underline{C}$ mátrix i-edik sorának j-edik elemét úgy kapjuk meg, hogy az $\underline{A}$ mátrix i-edik sorának és a $\underline{B}$ mátrix j-edik oszlopának skaláris szorzatát vesszük.

A mátrixok szorzására nem érvényes a kommutativitás, vagyis általában nem áll fenn, hogy

$$\underline{A}\cdot\underline{B} = \underline{B}\cdot\underline{A}.$$

A mátrixok szorzását definiálhatjuk a következőképpen is:

$$\underline{A}\,\underline{B} = \underline{A}[\underline{b}_1, \underline{b}_2, \dots, \underline{b}_k] = [\underline{A}\,\underline{b}_1, \underline{A}\,\underline{b}_2, \dots, \underline{A}\,\underline{b}_k]\,,$$

ahol

$$\underline{A}\,\underline{b}_i = b_{1i}\underline{a}_1 + b_{2i}\underline{a}_2 + \dots + b_{mi}\underline{a}_m\,,$$

ami az $\underline{A}$ mátrix oszlopvektorainak lineáris kombinációja.

Pl.

$$\underline{A} = \begin{pmatrix} 1 & 5 \\ 0 & 2 \\ -3 & 4 \end{pmatrix}, \qquad \underline{B} = \begin{pmatrix} 2 & 3 & 0 \\ 4 & 1 & 5 \end{pmatrix},$$

$$\underline{A}\cdot\underline{B} = \begin{pmatrix} 22 & 8 & 25 \\ 8 & 2 & 10 \\ 10 & -5 & 20 \end{pmatrix} \quad \text{és} \quad \underline{B}\cdot\underline{A} = \begin{pmatrix} 2 & 16 \\ -11 & 42 \end{pmatrix}.$$

Bizonyítás nélkül megjegyezzük a következő azonosságot:

$$(\underline{A}\cdot\underline{B})' = \underline{B}'\underline{A}'$$

$$(\underline{A}\cdot\underline{B}\cdot\underline{C})' = \underline{C}'\underline{B}'\underline{A}'.$$

Mátrixok blokkokra bontása

A mátrixokat néha blokkokra bontjuk, particionáljuk.

Pl.

$$\begin{pmatrix} 1 & 2 & \vdots & 3 & 6 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 4 & \vdots & 2 & 3 \\ 5 & 0 & \vdots & 1 & 2 \\ 2 & 9 & \vdots & 7 & 4 \end{pmatrix}.$$

A particionált mátrixok hasonló módon viselkednek az egyes műveleteknél, csak a blokkok között teljesülni kell a műveletek feltételeinek.

Pl.

$$\underline{A} \cdot \underline{B} = [\underline{A}_1, \underline{A}_2] \begin{pmatrix} \underline{B}_1 \\ \underline{B}_2 \end{pmatrix} = \underline{A}_1 \underline{B}_1 + \underline{A}_2 \underline{B}_2$$

Ha $\underline{A}$ $(n \times m)$ típusú és $\underline{B}$ $(m \times k)$ típusú, hogy a műveleteket el tudjuk végezni, $\underline{A}_1$ mátrix oszlopainak a száma meg kell hogy egyezzen $\underline{B}_1$ sorainak a számával, és hasonlóan $\underline{A}_2$ és $\underline{B}_2$ szorzóknál is.

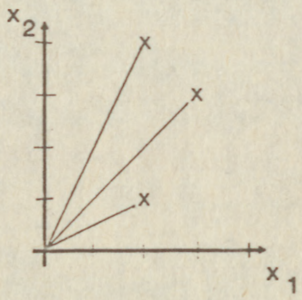
A mátrix geometriai értelmezése

Egy adott $(n \times m)$ típusú $\underline{X}$ mátrix minden oszlopa felfogható úgy, mint az n–dimenziós tér egy–egy pontjának derékszögű koordinátái.

Pl. az

$$\underline{X} = \begin{pmatrix} 2 & 3 & 2 \\ 1 & 3 & 4 \end{pmatrix}$$

mátrix oszlopvektorai meghatározzák az alábbi ábrán látható pontokat, és a hozzájuk tartozó irányvektorokat.



Egy adott $\underline{a}$ vektor hosszán (abszolút értékén) értjük azt a nem negatív valós számot, amelyet a

$$|\underline{a}| = \sqrt{\underline{a}' \underline{a}} \qquad \text{formula határoz meg.}$$

Két adott $\underline{a}$ és $\underline{b}$ vektor φ hajlásszögét a következő képlet határozza meg:

$$\cos\varphi = \frac{\underline{a}'\underline{b}}{|\underline{a}||\underline{b}|}.$$

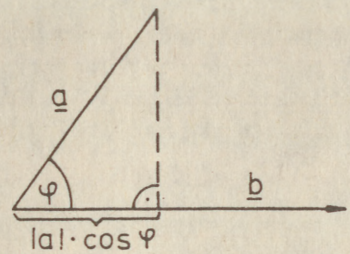
Ebből következik, hogy két vektor akkor alkot derékszöget, ha skaláris szorzatuk egyenlő zérussal ($\underline{a}'\underline{b} = 0, \quad \cos\varphi = 0, \quad \varphi = \frac{\pi}{2}$).

Ekkor azt mondjuk, hogy a két vektor ortogonális.

Egy $\underline{a}$ vektornak $\underline{b}$ irányára merőleges vetületét a következőképpen definiáljuk:

$$|\underline{a}| \cdot \cos\varphi$$

Ábrája:



A vetület előjeles szám: pozitív, ha az $\underline{a}$ és $\underline{b}$ által bezárt szög hegyesszög, és negatív, ha φ tompaszög.

A mátrixok rangja

Az $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_k$ vektorokat lineárisan függetleneknek nevezzük, ha a $\underline{0}$ vektort csak triviálisan állítja elő, azaz

$x_1\underline{a}_1 + x_2\underline{a}_2 + \dots x_k\underline{a}_k = \underline{0}$ egyenletnek csak egy megoldása van, és az $x_1 = x_2 = \dots = x_k = 0$.

Egy mátrix rangján a mátrix oszlopvektoraiból (sorvektoraiból) maximálisan kiválasztható lineárisan független vektorok számát értjük. Egy kvadratikus mátrixot szingulárisnak nevezünk, ha a mátrix rangja kisebb mint oszlopainak (ill. sorainak) száma, ellenkező esetben nem szinguláris.

A determináns

Minden kvadratikus $\underline{A}$ mátrixot jellemezhetünk egy skalárral, amelynek nagysága függ a mátrix elemeitől. Ezt a számot nevezzük az $\underline{A}$ mátrix determinánsának. Jelölése: $|\underline{A}|$. A determinánst a következőképpen állíthatjuk elő.

Ha $\underline{A}$ rendje (oszlopainak ill. sorainak száma) n, az $\underline{A}$ mátrix determinánsa ($|\underline{A}|$) az a_{ij} elemek ($i = 1, 2, \ldots, n$), indexeinek permutációjával előállított alternatív előjelű szorzatainak összege. Pl. ha $n = 2$, akkor

$$|\underline{A}| = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21} ,$$

ha $n = 3$, akkor

$$|\underline{A}| = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} +$$

$$+ a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} .$$

Megfigyelhettük, hogy az első index (i) növekvő sorrendbe írva azonos marad minden tagban, míg a második index-szel az összes lehetséges permutációt leírtuk. Az előjel azzal függ össze, hogy a j indexben páros vagy páratlan számú megfordítás van, azaz páros vagy páratlan számú kisebb index előz meg egy nagyobbat. Ha páros, pozitív, ha páratlan, akkor negatív az előjel.

Nagy n estén már nehézkessé válik ez a kifejtés. Ilyenkor másik módszert használhatunk.

Definiáljuk ehhez az aldetermináns fogalmát. Egy $\underline{A}$ mátrix aldeterminánsán az i-edik sor és a j-edik oszlop elhagyása után megmaradó $\underline{A}_{ij}$ mátrix determinánsát értjük ($|\underline{A}_{ij}|$), amihez $(-1)^{i+j}$ előjel tartozik. Így $c_{ij} = (-1)^{i+j}|\underline{A}_{ij}|$-t előjeles aldeterminánsnak nevezzük.

Az $\underline{A}$ mátrix i-edik sora szerint kifejtett determinánsa:

$$|\underline{A}| = (-1)^{i+1}a_{i1}|\underline{A}_{ij}| + (-1)^{i+2}a_{i2}|\underline{A}_{i2}| + \ldots +$$

$$+ (-1)^{i+n}a_{in}|\underline{A}_{in}| = a_{i1}c_{i1} + a_{i2}c_{i2} + \ldots + a_{in}c_{in} .$$

Néhány szabályt közlünk, bizonyítás nélkül:

a) Egy mátrix determinánsa egyenlő transzponáltjának determinánsával, azaz $|\underline{A}| = |\underline{A}'|$.

b) Ha egy mátrixban két szomszédos oszlopot (vagy két szomszédos sort) felcserélünk, akkor a determináns előjele ellenkezőre változik (abszolút értéke nem változik).

c) Ha egy mátrixban az egyik oszlop (vagy sor) egy másik oszlop (vagy sor) skalárszorosa (azaz a mátrix szinguláris), akkor a mátrix determinánsa zérus.

d) Ha egy mátrixban az egyik oszlop (vagy sor) elemeihez hozzáadjuk egy másik oszlop (vagy sor) elemeinek skalárszorosát, akkor a determináns értéke változatlan marad.

d) Ha egy mátrixban valamely oszlopot (vagy sort) megszorzunk egy λ skalárral, akkor az így kapott mátrix determinánsa az eredeti mátrix determinánsának a λ skalárszorosa lesz.

f) Két n-ed rendű mátrix szorzatának determinánsa egyenlő a két mátrix determinánsának szorzatával, azaz

$$|\underline{A} \cdot \underline{B}| = |\underline{A}| \cdot |\underline{B}|$$

Mátrixok inverze

Legyen $\underline{A}$ kvadratikus mátrix. Ha találunk egy olyan $\underline{B}$ mátrixot, hogy $\underline{A} \cdot \underline{B} = \underline{B} \cdot \underline{A} = \underline{I}$ legyen, akkor $\underline{B}$ mátrixot $\underline{A}$ mátrix inverzének nevezzük, és az $\underline{A}^{-1}$ szimbólummal jelöljük ($\underline{B} = \underline{A}^{-1}$). Legyen $\underline{C}$ az előjeles aldeterminánsok mátrixa ($c_{ij} = (-1)^{i+j}|\underline{A}_{ij}|$). Az $\underline{A}$ mátrix inverzét a következőképpen határozhatjuk meg:

$$\underline{A}^{-1} = \underline{C}'/|\underline{A}| .$$

Láthatjuk, hogy az inverz nem létezik, ha $|\underline{A}|$ egyenlő nullával, és ha $\underline{A}$ nem kvadratikus.

Az $\underline{A}$ mátrix determinánsa akkor zérus, ha az $\underline{A}$ szinguláris, és $|\underline{A}| \neq 0$, ha $\underline{A}$ nem szinguláris.

Gyakran találkozunk a következő azonossággal:

$(\underline{A} \cdot \underline{B})^{-1} = \underline{B}^{-1}\underline{A}^{-1}$, ahol feltesszük, hogy $\underline{A}$ és $\underline{B}$ kvadratikus, és $\underline{B}^{-1}$ és $\underline{A}^{-1}$ létezik.

Gyakran van szükségünk diagonális mátrixok inverzére. Egy diagonális mátrix inverze egy olyan diagonális mátrix, melynek diagonális elemei az eredeti mátrix diagonális elemeinek reciprokai.

Példa inverz meghatározására:

$$\underline{A} = \begin{pmatrix} -2 & 5 & 1 \\ 3 & -1 & 4 \\ 2 & 0 & 7 \end{pmatrix}$$

A determináns értékének meghatározásához használjuk az első sor szerinti kifejtést. Így

$$|\underline{A}| = -2 \begin{pmatrix} -1 & 4 \\ 0 & 7 \end{pmatrix} - 5 \begin{pmatrix} 3 & 4 \\ 2 & 7 \end{pmatrix} + \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} = -49 .$$

Az előjeles aldeterminánsokat tartalmazó mátrix:

$$\underline{C} = \begin{pmatrix} \begin{pmatrix} -1 & 4 \\ 0 & 7 \end{pmatrix} & -\begin{pmatrix} 3 & 4 \\ 2 & 7 \end{pmatrix} & \begin{pmatrix} 3 & -1 \\ 2 & 0 \end{pmatrix} \\ \begin{pmatrix} 5 & 1 \\ 0 & 7 \end{pmatrix} & \begin{pmatrix} -2 & 1 \\ 2 & 7 \end{pmatrix} & -\begin{pmatrix} -2 & 5 \\ 2 & 0 \end{pmatrix} \\ \begin{pmatrix} 5 & 1 \\ -1 & 4 \end{pmatrix} & -\begin{pmatrix} -2 & 1 \\ 3 & 4 \end{pmatrix} & \begin{pmatrix} -2 & 5 \\ 3 & 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} -7 & -13 & +2 \\ -35 & -16 & 10 \\ 21 & 11 & -17 \end{pmatrix}$$

Az $\underline{A}$ mátrix inverze:

$$\underline{A}^{-1} = \underline{C}'|\underline{A}| = -\frac{1}{49} \begin{pmatrix} -7 & -35 & 21 \\ -13 & -16 & 11 \\ 2 & 10 & -17 \end{pmatrix}$$

Ellenőrizhetjük, hogy $\underline{A} \cdot \underline{A}^{-1} = \underline{E}$

Mátrixok sajátértéke és sajátvektora

Legyen $\underline{A}$ egy $(n \times n)$-es kvadratikus mátrix. Bebizonyítható, hogy létezik olyan k vektor, hogy $\underline{A}\,\underline{k} = \lambda \underline{k}$, ahol λ skalár (a triviális megoldást $\underline{k} = \underline{0}$ kizárva). Azt a $\underline{k}$ vektort, mely kielégíti az egyenletet, sajátvektornak nevezzük, és az ezzel összefüggő λ skalárértéket $\underline{A}$ mátrix sajátértéknek.

Az $\underline{A}\,\underline{k} = \lambda\,\underline{k}$-val azonos az

$$\underline{A}\,\underline{k} - \lambda \underline{k} = \underline{0}, \quad \text{ill.} \quad (\underline{A} - \lambda \underline{E})\underline{k} = \underline{0} \quad \text{forma.} \tag{1}$$

A fenti homogén egyenletrendszer n ismeretlent tartalmaz ($\underline{k}$ elemei), ezen kívül még λ skalár ismeretlen. Az egyenletrendszernek akkor van a triviálistól különböző megoldása, ha $\underline{A}$ mátrix rangja kisebb mint n, vagy ami ugyanazt jelenti

$$|\underline{A} - \lambda \underline{E}| = 0 .$$

A determináns jó szolgálatot tesz, mivel kifejtésével megkaphatjuk az $(n + 1)$-edik ismeretlen λ-t. Mivel a determináns kifejtésével az n-edfokú egyenlethez jutunk, általában n különböző gyököt kapunk λ-ra.

Ez azt is jelenti, hogy általában $\underline{A}$ $(n \times n)$ mátrixnak n különböző sajátértéke van, és így n különböző sajátvektora is.

Ha $\underline{K}$ mátrix tartalmazza a sajátvektorokat, és $\underline{L}$ diagonális mátrix a sajátértékeket, az $\underline{A}\,\underline{k} = \lambda \underline{k}$ általános formája:

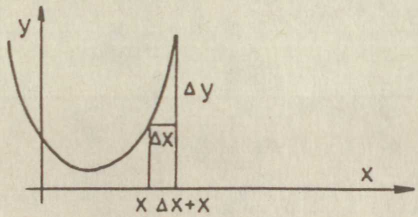
$$\underline{A}\,\underline{K} = \underline{K}\,\underline{L} .$$

Differenciálhányados

Egy $y = f(x)$ függvény legyen pl:

$$y = x^2 - 2x + 3,$$

amelynek képe az alábbi ábrán látható:



Ha x értékét Δx mennyiséggel növeljük, akkor y értéke is megváltozik Δy értékkel. Általában

$$y + \Delta y = f(x + \Delta x).$$

Az ábrából láthatjuk, hogy $\frac{\Delta y}{\Delta x}$ hányados az iránytangensét jelöli az (x, y) és $(x + \Delta x, y + \Delta y)$ pontokat összekötő egyenesek.

Ha Δx kicsi, közelítően az (x, y) ponthoz húzott érintő iránytangensét kapjuk. A függvény x helyhez tartozó differenciálhányadosának a

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

határértéket nevezzük. Ha a differenciálhányadost x függvényeként fogjuk fel, akkor az ún. derivált függvényhez jutunk.

Jelölése:

$$y' = f'(x) = \frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$$

amit y első deriváltjának is nevezünk.

Pl. $y = bx^2$ deriváltját a következőképpen kaphatjuk:

$$y + \Delta y = b(x + \Delta x)^2 = bx^2 + 2bx\Delta x + b(\Delta x)^2,$$

amiből

$$\Delta y = 2bx\Delta x + b(\Delta x)^2,$$

és

$$\frac{\Delta y}{\Delta x} = 2bx + b(\Delta x).$$

Ha x eléggé kicsi, akkor a jobboldal második tagját elhanyagolhatjuk, így

$$\frac{dy}{dx} = 2bx \quad \text{vagy} \quad y' = 2bx\,.$$

Általánosságban, ha

$$y = ax^n\,,$$

akkor

$$y' = nax^{n-1}\,.$$

Deriválási szabályok

(1). Ha $y = u + v$ és u, v függyvénye x-nek, akkor

$$y' = u' + v'\,.$$

pl.

$$u = x^2,\ b = 2x^3$$
$$y = u + v = x^2 + 2x^3$$
$$y' = u' + v' = 2x + 6x^2\,.$$

(2). Ha u és v függvénye x-nek, valamint $y = uv$, akkor

$$y' = u'v + uv'$$

$$u = x^2,\ v = 2x^3 \quad \text{és} \quad y = uv, \quad \text{akkor}$$
$$y' = (2x)2x^3 + (x^2)(6x^2) = 10x^4\,.$$

Természetesen ugyanezt kaptuk volna közvetlenül $y = 2x^5$-ből.

(3). Ha u és v függvénye x-nek és $y = \frac{u}{v}$, akkor

$$y' = \frac{u'v - uv'}{v^2}\,.$$

Pl.

$$u = x^2,\ v = 2x^3$$
$$y' = \frac{4x^4 - 6x^4}{4x^6} = \left(-\frac{1}{2}\right) x^{-2}\,.$$

(4). Ha y függvénye u-nak és u függvénye x-nek, vagyis ha $y = f(u)$, és $u = g(x)$, akkor

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

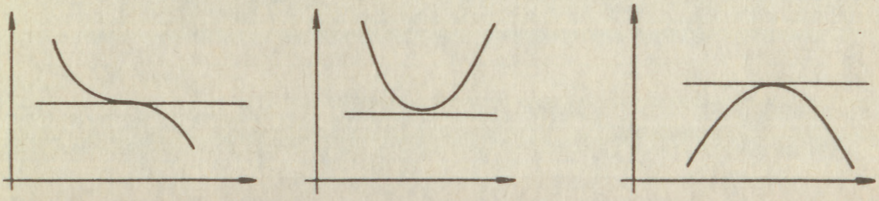
pl.

$$y = u^2 + 4, \; u = 2x - 3$$

$$\frac{dy}{dx} = (2u) \cdot (2) = 4u = 8x - 12$$

Szélső értékek

Legyen $y = f(x)$ és $y' = g(x)$. Ha találunk olyan x_1 értéket, melyre $y' = 0$, akkor az $y = f(x)$ függvény érintője az x tengellyel párhuzamos lesz. Ezt mutatja a következő ábra:



a) inflexiós pont   b) minimum   c) maximum

Ha y' deriváltja, azaz y második deriváltja y'' pozitív ebben a pontban (x_1), akkor a függvénynek minimuma van, és ha y'' negatív, az x_1 pontban az $y = f(x)$ függvénynek maximuma van.

Parciális derivált

Legyen y változó több független változó függvénye. Pl. $y = x_1^2 + 2x_2^2$. Tekintsük az egyik független változót konstansnak és vegyük a másik változó szerinti deriváltját a függvénynek. Az így kapott deriváltat nevezzük parciális deriváltnak. ($\partial y / \partial x_i$).

A szélső értéke a függvénynek ott lehet, ahol a parciális deriváltak egyenlőek nullával.

A példában

$$\frac{\partial y}{\partial x_1} = 2x_1 \qquad \frac{\partial y}{\partial x_2} = 4x_2 \, .$$

A szélső értékek lehetséges helye $2x_1 = 0$ és $4x_2 = 0$, azaz $x_1 = x_2 = 0$.

A szélsőértékről úgy is eldönthetjük, hogy maximum–e vagy minimum, hogy a szélsőérték környezetében megvizsgáljuk a függvényértékeket.

A Lagrange–féle multiplikátor–módszer

Nagyon gyakran $y = f(x)$ függvény szélsőértékét valamilyen egyenlettel megadott feltétel mellett kell keresnünk. A szélsőértéket ilyenkor a Lagrange–féle multiplikátor módszerrel határozhatjuk meg. Adott az $y = f(\underline{x})$ többváltozós függvény (az $\underline{x}$ vektorváltozó, $\underline{x} = [x_1, x_2, \ldots, x_n]$) és a $g(\underline{x}) = 0$ feltétel.

Képezzük a Lagrange–féle függvényt:

$$L = f(\underline{x}) - \lambda' g(\underline{x}) .$$

Az $f(\underline{x})$ függvénynek ott van feltételes szélsőértéke (értékei), ahol az L függvény feltétel nélküli szélső értékkel rendelkezik. $\partial L / \partial x_i = 0$, ahol $i = 1, 2, \ldots, n$.

Vektor derivált

Legyen n változó függvénye $y = \underline{a}'\underline{x}$, kifejtve $y = a_1x_1 + a_2x_2 + \ldots + a_nx_n$.

A parciális deriváltak $\partial y/\partial x_1, \partial y/\partial x_2, \ldots, \partial y/\partial x_n$ az $a_1, \ldots, a_n$ együtthatókat adják.

Egyszerűbben is leírhatjuk a deriváltakat $\partial y/\partial \underline{x} = \underline{a}'$. Ezt nevezzük vektor deriváltnak.

1. Az $f = \underline{x}'\underline{A}\,\underline{y}$ függvényt bilineáris formának hívjuk ($\underline{A}$ $n \times m$ típusú mátrix). Részletesen:

$$\begin{aligned} f =& x_1a_{11}y_1 + x_1a_{12}y_2 + \ldots + x_1a_{1m}y_m + \\ &+ x_2a_{21}y_1 + x_2a_{22}y_2 + \ldots + x_2a_{2m}y_m + \\ &\vdots \\ &+ x_na_{n1}y_1 + x_na_{n2}y_2 + \ldots + x_na_{nm}x_m \end{aligned}$$

A parciális deriváltak:

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= a_{11}y_1 + a_{12}a_2 + \ldots + a_{1m} = \underline{a}_1'\underline{y} \\ \frac{\partial f}{\partial x_2} &= \underline{a}_2'\underline{y} \\ &\vdots \\ \frac{\partial f}{\partial x_n} &= \underline{a}_n'\underline{y} \end{aligned}$$

Összesítve $\partial f/\partial \underline{x}' = \underline{A}\underline{y}$ vagy sorvektoronként

$$\frac{\partial f}{\partial \underline{x}} = \underline{y}'\underline{A}' .$$

2. Az $f = \underline{x}'\underline{A}\,\underline{x}$ formát az $\underline{A}$ mátrix kvadratikus formájának nevezik, ahol $\underline{A}$ szimmetrikus mátrix. Az előzőhöz hasonlóan kaphatjuk a parciális deriváltakat:

$$\frac{\partial f}{\partial \underline{x}'} = 2\underline{A}\,\underline{x} \quad \text{vagy}$$

$$\frac{\partial f}{\partial \underline{x}} = 2\underline{x}'\underline{A}\,.$$

A szélső értékek meghatározása ezek után mátrixjelölésekkel sem okozhat gondot.

I.3 Eloszlásfüggvény — Normális eloszlás

Egy ξ valószínűségi változó eloszlásfüggvényén az

$$F(x) = P(\xi < x)$$

függvényt értjük, ahol x bármely valós szám lehet.

Ebből a definícióból következik, hogy az eloszlásfüggvény:

— monoton nem csökkenő függvény,

— értékei a 0 és 1 közé esnek.

Egy ξ változó empirikus eloszlásfüggvényének az

$F_n(x) = \frac{k}{n}$ függvényt nevezzük, ahol k az x-nél kisebb mintaelemek számát jelenti.

Sűrűségfüggvény

Egy ξ változó sűrűségfüggvényén az

$f(x) = F'(x)$ függvényt értjük (az eloszlásfüggvény deriváltját), ha az $F(x)$ folytonos és véges sok hely kivételével mindenütt differenciálható.

A sűrűségfüggvény nem–negatív függvény, és a függvény alatti terület értéke egyenlő 1-gyel.

A ξ változó empirikus sűrűségfüggvényét a minta hisztogrammjának nevezzük.

Normális eloszlás

A matematikai statisztika többváltozós módszerei jelentős részben azon feltevésen alapulnak, hogy a változók együttes eloszlása normális eloszlást követ.

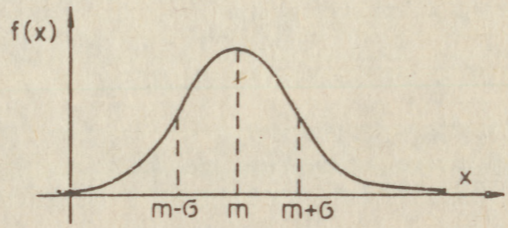
Ezért a változók eloszlása vizsgálatának mindig fontos szerepet kell tulajdonítanunk.

Egy x változó normális eloszlású, ha sűrűségfüggvénye

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}},$$

ahol az m és σ az eloszlás paraméterei. Az m paraméter a változó várható értéke (átlaga) σ paraméter pedig a szórása.

A normális eloszlás ábrája:



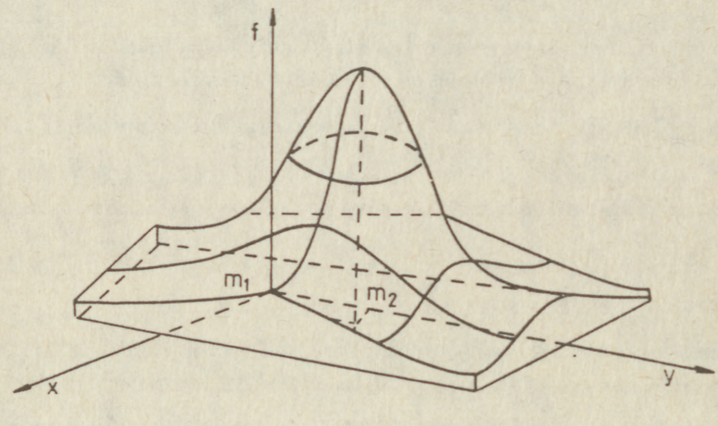
Az x és y, két folytonos változó együttes eloszlását kétdimenziós normális eloszlásúnak nevezzük, ha sűrűségfüggvénye a következő alakú:

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2} - 2r\frac{(x-m_1)(y-m_2)}{\sigma_1\sigma_2} + \frac{(y-m_2)^2}{\sigma_2^2}\right]},$$

ahol

m_1 x változó várható értéke,
m_2 y változó várható értéke,
σ_1 x változó szórása,
σ_2 y változó szórása,
r a két változó korrelációs együtthatója.

Az eloszlás ábrája:



Általánosságban n változó együttes normális eloszlását a következőképpen definiáljuk.

Az n változót az

$\underline{x} = [x_1, x_2, \ldots, x_n]$ vektorba foglaljuk össze.

A vektorváltozó komponenseinek várható értékeit tartalmazza az

$\underline{m} = [m_1, m_2, m_3, \ldots, m_n]$ n–elemű átlagvektor.

$\underline{C}$ mátrix elemei pedig a változók kovarianciáit tartalmazák.

Ezek felhasználásával az $\underline{x}$ n–dimenziós vektorváltozót normális eloszlásúnak nevezzük, ha sűrűségfüggvénye:

$$f(\underline{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\underline{C}|}} e^{-\frac{1}{2}(\underline{x}-\underline{m})'\underline{C}^{-1}(\underline{x}-\underline{m})}.$$

A normális eloszlás jelentőségét az adja, hogy a központi határeloszlás tétele értelmében nagy számú azonos eloszlást követő független változók összege megközelítően normális eloszlású.

Végezetül megemlítünk három fontosabb tételt:

– Ha az $\underline{x}$ vektorváltozó normális eloszlású és komponensei páronként korrelálatlanok, akkor a komponensek egymástól független változók. (Általában korrelálatlanságból még nem következtethetünk függetlenségre.)

– Ha az $\underline{x}$ vektorváltozó normális eloszlású, akkor komponensei is (a peremeloszlások) egydimenziós normális eloszlásúak. Ez a tétel nem megfordítható, tudniillik abból, hogy $\underline{x}$ vektorváltozó komponensei $(x_i - k)$ normális eloszlásúak, nem következik, hogy az együttes eloszlásuk is normális.

– Ha az $\underline{x}$ vektorváltozó normális eloszlású, akkor az x_i változónak a többi változóra vonatkozó regressziós függvénye lineáris.

Kiadja a Társadalomkutatási Informatikai Egyesülés
és az Országos Műszaki Információs Központ és Könyvtár
Felelős kiadó: dr. Horváth Péter
Készült az OMIKK nyomdaüzemében
1991.
Budapest, I., Gyorskocsi u. 5 – 7.
Felelős vezető: Tóth Károly
Nyomdai megrendelés száma: 383
Felelős szerkesztő: dr. Hetényi Pálné

1739-3

600 Ft

TÁRSADALOMKUTATÁSI MÓDSZERTANI TANULMÁNYOK III.

Füstös László: Látens változós modellek