Enkele schatters van factor scores Bayesiaans bezien

Pieter Vijn

Subfaculteit Psychologie, Universiteit van Amsterdam.

Samenvatting

Bekende schatters van factorscores zijn die van Thurstone en Bartlett. Beide kunnen eenvoudig afgeleid worden met de Bayesiaanse methode. Ze verschillen in de aprioriverdeling van de factoren.

Inleiding

Factor scores kunnen geschat worden via een aantal methoden, Mulder et. al. (1977). Van de schatter van Thurstone en Bartlett zullen we een Bayesiaanse afleiding geven. In het geval waarin de variantie-covariantie matrix van de factoren volledig onbekend is, kan er een gebied aangegeven worden, waarin de schatter van Thurstone moet liggen.

Het model

Uitgangspunt is het algemeen gangbare factoranalyse model, zoals dat beschreven is in Mulder et. al. (1977). We volgen hun notatie:

x	: de vector van gemeten variabelen	$(n\times 1)$
Λ	: de matrix met factorladingen	$(n\times p)$
e	: de vector van errortermen	$(n\times 1)$
f	: de vector van factoren	$(p\times 1)$
$\hat{f}$	: de vector van factorscore-schatters	$(p\times 1)$

Het model: $x = \Lambda f + e$ (1)

De veronderstellingen zijn:

- $E(e) = 0, \quad E(f.e') = 0$
- de factoren en variabelen zijn gestandaardiseerd
- $E(xx') = \Sigma,\ E(ff') = \Phi,\ E(ee') = \Psi,\ \Psi$ diagonaal

en:

$$\Sigma = \Lambda\,\Phi\,\Lambda' + \Psi \qquad (2)$$

De schatter van Thurstone

Dit is de schatter met de kleinste kwadratische fout:

$$\hat{f}_r = \Phi \Lambda' \Sigma^{-1} x \tag{3}$$

Deze niet-zuivere schatter kan geschreven worden als:

$$\hat{f}_r = (\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Lambda' \Psi^{-1} x \tag{4}$$

We herkennen Φ^{-1} als de inverse van de variantie-covariantie matrix der factoren; ofwel de precisiematrix. De variantie-covariantie matrix van $\hat{f}_r$ is $(\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1}$.

De afleiding van (4) kan via directe inversie van $(\Lambda \Phi \Lambda' + \Psi)$ of door een resultaat van Lawley en Maxwell (1971, vergl. (8.7), pag. 109) in de vorm (4) te schrijven.

<u>De schatter van Bartlett</u>

Deze schatter is in feite een gegeneraliseerde kleinste kwadratenschatter. We kunnen het model (1) ook schrijven als:

$$\Psi^{-\frac{1}{2}} x = \Psi^{-\frac{1}{2}} \Lambda f + e \tag{5}$$

met $E\ e = 0$ en $Eee' = I$.

De kleinste kwadratenschatter van model (5) is direct:

$$\hat{f}_b = (\Lambda' \Psi^{-1} \Lambda)^{-1} \Lambda' \Psi^{-1} x \tag{6}$$

Dit is een BLUE schatter (Best- Linear- Unbiased Estimater). De variantie-covariantie matrix van de Bartlettschatter is $(\Lambda' \Psi^{-1} \Lambda)^{-1}$.

<u>Bayesiaanse Interpretatie</u>

In de Bayesiaanse aanpak wordt over de, in de klassieke statistiek vaste doch onbekende, populatieparameters een kansverdeling verondersteld. We maken allereerst de, in maximum likelihood factor-analyse gebruikelijke, aanname dat de fouten uit een normale verdeling komen:

$$e \approx N(0, \Psi) \tag{7}$$

De observatievector x volgt dan, gegeven de factor f, een normale verdeling:

$$x | f \approx N(\Lambda f, \Psi) \tag{8}$$

De regel van Bayes kan dan gebruikt worden indien er een aprioriverdeling over f, p(f), verondersteld wordt.

$$\text{"Bayes"}: \quad p(f|x) \sim p(x|f)\, p(f) \tag{9}$$

waarin $p(f|x)$ de aposteriori verdeling van f is, gegeven de observatievector x.

De prior p(f)

Voor de prior p(f) bestaan ruwweg twee mogelijkheden:

a) informatieve prior

We veronderstellen hier dat f apriori uit een normale verdeling komt met het gemiddelde f_a en variantie-covariantie matrix P:

$$f \approx N(f_a, P) \tag{10}$$

b) niet-informatieve prior

Soms is het handig een prior te gebruiken zonder voorkennis, f_a en P. Daarvoor wordt een niet-integreerbare verdeling gebruikt, Box en Tiao (1973, pag. 115)

$$P(f) \sim c \tag{11}$$

waarin c een willekeurige constante is. Deze niet-informatieve prior ontstaat ook indien we van de informatieve prior de precisiematrix P^{-1} naar een nulmatrix laten gaan.

De aposterioriverdeling p(f/x)

We hebben nu: $f \approx N(f_a, P)$ en

$x|f \approx N(\Lambda f, \Psi)$

Een standaardresultaat is dan, Chamberlain en Leamer (1976):

$$f|x \approx N(\tilde{f}, H^{-1}) \tag{12}$$

met:

- het aposteriorigemiddelde

$$\tilde{f} = (\Lambda' \Psi^{-1} \Lambda + P^{-1})^{-1} (\Lambda' \Psi^{-1} \Lambda \hat{f}_b + P^{-1} f_a)$$

- de variantie-covariantiematrix

$$H^{-1} = (\Lambda' \Psi^{-1} \Lambda + P^{-1})^{-1}$$

Het aposteriorigemiddelde $\tilde{f}$ is een matrix-gewogen gemiddelde van de Bartlett schatter $\hat{f}_b$ en het apriorigemiddelde f_a. De precisiematrix H is de som van de precisiematrix der Bartlett schatter en die van een aprioriverdeling.

<u>De Bayes schatter $\tilde{f}$</u>

In de beschrijving van model (1) was het uitgangspunt: $E\, f = 0$ en $E\, f\, f' = \Phi$. Met de keuze van $f_a = 0$ en $P = \Phi$ wordt dat apriori gerealiseerd. Substitutie in (12) en gebruikmakend van (6) geeft voor de Bayes schatter $\tilde{f}$ (het gemiddelde van de aposterioriverdeling):

$$\tilde{f} = (\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1} \Lambda' \Psi^{-1} x \qquad (13)$$

en we zien dat $\tilde{f}$ identiek is aan de schatter van Thurstone in (4). Bovendien blijkt nu dat onder een niet-informatieve prior, $P^{-1} \to 0$ (of ook: Φ^{-1} wordt als een nulmatrix verondersteld), de Bayes schatter $\tilde{f}$ identiek is aan de Bartlett schatter in (6).

Samenvattend:

Als $\tilde{f} = N(0, \Phi)$ dan $\tilde{f} = \hat{f}_r$, als $p(f) \sim c$ dan $\tilde{f} = \hat{f}_b$.

De variantie-covariantie matrices zijn identiek aan de klassieke resultaten. Uit de struktuur van $(\Lambda' \Psi^{-1} \Lambda + \Phi^{-1})^{-1}$ volgt direkt dat de variantie van de Thurstone schatter nooit kleiner is dan die van de Bartlett schatter. Indien de matrix $\Lambda' \Psi^{-1} \Lambda$ ongeveer een nulmatrix zou zijn (bijvoorbeeld in een zeer slecht passend factormodel) dan convergeert de Thurstone schatter naar een Gar schatter. De variantie-covariantie matrix zou dan gelijk worden aan Φ en daarmee struktuur bewarend zijn, Mulder et. al (1977). Een simpele Bayesiaanse oplossing voor de Gar schatter is niet voorradig.

<u>Φ volledig onbekend</u>

De schatters zijn operationeel indien de matrices Λ, Ψ en Φ bekend zijn.

Daartoe worden schattingen $\hat{\Lambda}$, $\hat{\Psi}$ en $\hat{\Phi}$gebruikt, deze zijn natuurlijk niet perfect. We gaan er nu in het volgende van uit dat Λ en Ψ bekend zijn maar dat Φ volledig onbekend is.

Grenzen voor het aposteriorigemiddelde $\tilde{f}$

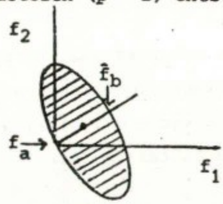
Indien Φ onbekend is, blijkt het toch mogelijk een gebied aan te geven, waarin $\tilde{f}$ (lees: de Thurstone schatter) moet liggen. Daartoe wordt gebruik gemaakt van een stelling uit Chamberlain en Leamer (1976), toegepast op onze situatie:

Het aposteriorigemiddelde $\tilde{f}$ ligt in de ellipsoide met als centrum het middelpunt van het lijnsegment dat $\hat{f}_b$ (de Bartlett schatter) en f_a (het apriorigemiddelde) verbindt, de grens van de ellipsoide is inclusief $\hat{f}_b$ en f_a.

In formulevorm:

$$(\tilde{f} - \tfrac{1}{2}\hat{f}_b)' \, (\Lambda' \, \Psi^{-1} \, \Lambda) \, (\tilde{f} - \tfrac{1}{2} \hat{f}_b) < \frac{\hat{f}_b' \, . \, \Lambda' \, \Psi^{-1} \, \Lambda \, . \, \hat{f}_b}{4} \qquad (14)$$

In het geval van twee factoren (p = 2) ontstaat de volgende figuur:



Ongeacht de grootte van Φ, altijd zal de schatter $\tilde{f}$ in het gearceerde gebied liggen. Indien Φ volledig bekend is bestaat dit gebied uit één punt: de Thurstone schatter. In het geval van orthogonale gestandaardiseerde factoren, $\Phi = I$ geldt, gebruik makend van een stelling uit Chamberlain en Leamer (1976):

$$0 \leq |\hat{f}_{ri}| \leq |\hat{f}_{bi}| \quad , \; i = 1,\ldots\ldots,p \qquad (15)$$

De coördinaten van de Thurstone schatter $\hat{f}_r$ liggen dan elementsgewijs tussen 0 en die van de Bartlett schatter.

Literatuur

Box, E.P. and Tiao, G.C., Bayesian Inference in Statistical Analysis, Addison Wesley, 1973.

Chamberlain, G. and Leamer, E.E. Matrix weighted Averages and Posterior Bounds, J.R. Statist. Soc. B, 38, 85-94, 1976

Lawley, D.N. and Maxwell, A.E., Factor Analysis as a statistical method, London, 1971 (2e ed.).

Mulder, J. De Pijper, W.M., Saris, W.E., Het schatten van factorscores, MDN, 8, 56-66, 1977.