ABDUL RAUF

0607050

ANALISIS FAKTOR DAN

KESIMPULAN UNTUK STRUKTUR MATRIKS KOVARIANS

Analisis factor adalah salah satu metode statistik multivariat yang mencoba menerangkan hubungan antar sejumlah peubah-peubah yang saling independen antara satu dengan yang lain sehingga dibuat kumpulan peubah yang lebih sedikit dari jumlah peubah awal.

Analisis faktor dapat dianggap sebagai perluasan dari analisis komponen utama. Kedua teknik ini dapat dipakai untuk menentukan matriks covarians Σ. Tetapi perkiraan yang berdasarkan pada model analisis faktor lebih rumit.

9.2 Model Faktor Orthogonal

Misalkan X adalah vektor random yang terobservasi, dimana terdiri dari p komponen mempunyai mean μ dan kovariansi matriks Σ. Dalil model factor menyatakan bahwa X secara linear tergantung dengan (*common factor*) $F_1, F_2, \ldots, F_m$ dan faktor khusus (*specific factor*) $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_p$. Modelnya adalah :

$$
\begin{aligned}
X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \ldots + \ell_{1m}F_m + \varepsilon_1 \\
X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \ldots + \ell_{2m}F_m + \varepsilon_2 \\
&\vdots \\
X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \ldots + \ell_{pm}F_m + \varepsilon_{1p}
\end{aligned}
\qquad (9\text{-}1)
$$

Atau dalam notasi matriks, $\underset{(px1)}{X-\mu} = \underset{(pxm)}{L}\underset{(mx1)}{F} + \underset{(px1)}{\varepsilon}$ (9-2)

Koeffisien l_{ij} dinamakan loading dari variable ke-*i* pada factor ke-*j* sehingga matriks L dinamakan matriks dari faktor loading. Selisih p$X_1 - \mu_1, \dots, X_p - \mu_p$dinyatakan dalam hal p+m variabel random $F_1, F_2, \dots, F_m, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ yang tidak terobservasi.

Dengan banyaknya jumlah yang tidak terobservasi, verifikasi langsung model faktor dari observasi $X_1, X_2, \dots . X_p$ tidak berguna. Tetapi penambahan asumsi tentang vektor random F dan ε maka model (9-2) menandakan hubungan kovarians. Penambahan asumsi sebagai berikut:

$$E[F] = \underset{(mx1)}{0},\ Cov(F) = E[FF^t] = \underset{(mxm)}{\mathrm{I}}$$

$$E[\varepsilon] = \underset{(px1)}{0}, Cov(\varepsilon) = E[\varepsilon\varepsilon^t] = \underset{(pxp)}{\Psi} = \begin{bmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \end{bmatrix}$$

Ψ matriks diagonal

dimana F dan ε saling bebas(independent), sehingga

$$Cov(\varepsilon, F) = E[\varepsilon F'] = \underset{pxm}{0}$$

Model faktor orthogonal dengan m faktor umum

$$\underset{px1}{X} = \underset{px1}{\mu} + \underset{pxm}{L}\underset{(mx1)}{F} + \underset{px1}{\varepsilon} \qquad (9\text{-}4)$$

$\mu_i = rata - rata\ variabel\ ke\ i$

$\varepsilon_i = faktor\ khusus\ ke\ i$

$F_j = faktor\ umum\ ke\ j$

$l_{ij} = loading\ dari\ variabel\ ke\ i\ dan\ faktor\ ke\ j$

Model factor orthogonal menghasilkan struktur kovariansi untuk X, dari model (9-4)

$$(X-\mu)(X-\mu)' = (LF+\varepsilon)(LF+\varepsilon)'$$

$$= (LF+\varepsilon)((LF)'+\varepsilon')$$

$$= LF(LF)' + \varepsilon(LF)' + LF\varepsilon' + \varepsilon\varepsilon'$$

Sehingga menghasilkan kovariansi

$$\Sigma \quad = Cov(X) = E(X-\mu)(X-\mu)'$$

$$= LE(FF') \quad L' + E(\varepsilon F')L' + LE(F\varepsilon') + E(\varepsilon\varepsilon')$$

$$\Sigma = \mathrm{LL}' + \Psi$$

Kovarians untuk variable random X dan factor umum F diperoleh

$$(X-\mu)F' = (\mathrm{LF} + \varepsilon)\mathrm{F}'$$

$$= \mathrm{LFF}' + \varepsilon\ \mathrm{F}'$$

$$\mathrm{Cov(X,F)} = \mathrm{E}[(\mathrm{X} - \mu)\mathrm{F}']$$

$$= \mathrm{LE(FF')} + \mathrm{E}(\varepsilon \mathrm{F}') = \mathrm{L}$$

Struktur kovarians untuk model factor orthogonal

1. $Cov(X) = LL' + \Psi$

$Cov(X_i, X_j) = l_{i1}l_{j1} + \cdots + l_{im}l_{jm}$

$$var(X_i) = l_{i1}^2 + \cdots + l_{im}^2 + \psi_i \qquad (9\text{-}5)$$

2. Cov(X,F) = L

atau $Cov(X_i, F_j) = l_{ij}$

Bagian dari variansi variable ke I yang diberikan oleh factor umum dinamakan komunality. Sedangkan variabel X_i yang berasal dari faktor khusus disebut ”keunikan” atau ” variansi khusus ”. komunality ditandai dengan h_i^2, ditulis sebagai berikut.

$$var(X_i) = \sigma_{ii} = l_{i1}^2 + \cdots + l_{im}^2 + \psi_i$$

Jadi $\sigma_{ii} = h_i^2 + \psi_i$

dengan

σ_{ii} = variansi dari variabel random ke i

ψ_i = variansi khusus ke i

h_i^2 = komulatis ke i

$h_i^2 = l_{i1}^2 + \cdots + l_{im}^2$

Contoh 9.1

Misal matriks covarian

$$\Sigma = \begin{bmatrix} 19 & 30 & 2 & 12 \\ 30 & 57 & 5 & 23 \\ 2 & 5 & 38 & 47 \\ 12 & 23 & 47 & 68 \end{bmatrix}$$

Dengan persamaan

$$\begin{bmatrix} 19 & 30 & 2 & 12 \\ 30 & 57 & 5 & 23 \\ 2 & 5 & 38 & 47 \\ 12 & 23 & 47 & 68 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 7 & 2 \\ -1 & 6 \\ 1 & 8 \end{bmatrix} \begin{bmatrix} 4 & 7 & -1 & 1 \\ 1 & 2 & 6 & 8 \end{bmatrix} + \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

atau

$$\Sigma = LL' + \Psi$$

Karena ∑ punya struktur dengan m = 2. maka

$$L = \begin{bmatrix} l_{11} & l_{12} \\ l_{21} & l_{22} \\ l_{31} & l_{32} \\ l_{41} & l_{42} \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 7 & 2 \\ -1 & 6 \\ 1 & 8 \end{bmatrix}, \qquad \Psi = \begin{bmatrix} \psi_1 & 0 & 0 & 0 \\ 0 & \psi_2 & 0 & 0 \\ 0 & 0 & \psi_3 & 0 \\ 0 & 0 & 0 & \psi_4 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Komunaliti dari X_1 ?

$$h_1^2 = l_{11}^2 + l_{12}^2 = 4^2 + 1^2 = 17$$

Dan variansi dari X_1 ?

$$\sigma_{ii} = h_i^2 + \psi_i$$

$$= 17 + 2 = 19$$

Cara yang sama dapat digunakan untuk variabel X_2.

Model factor mengasumsikan bahwa p+p(p-1)/2=p(p+1)/2 varian dan kovarian untuk X dapat diperoleh dari *pm* factor loading l_{ij} dan p variansi khusus ψ_i. Saat m = p, setiap matriks covarians dapat diperoleh persisnya sebagai LL', sehingga Ψ dapat menjadi matriks nol. Tetapi saat m cukup kecil ke p maka analisis faktor paling berguna. Pada kasus ini model factor menyediakan penjelasan sederhana dari kocarian dalam X dengan parameter yang lebih sedikit dari p(p+1)/2 parameter dalam $\sum$.

Yang disayangkan untuk analisis factor, kebanyakan matriks covarian tidak dapat difaktorkan sebagai LL'+Ψ, dengan jumlah factor m lebih kecil dari p.

Contoh 9.2 (tidak ada solusi yang sesuai)

Misal p =3 dan m=1 dan misalkan variabel random $X_1, \dots, X_3$ mempunyai matriks kovarian

$$\sum = \begin{bmatrix} 1 & 0.9 & 0.7 \\ 0.9 & 1 & 0.4 \\ 0.7 & 0.4 & 1 \end{bmatrix}$$

Gunakan model factor

$$X_1 - \mu_1 = l_{11}F_1 + \varepsilon_1$$

$$X_2 - \mu_2 = l_{21}F_2 + \varepsilon_2$$

$$X_3 - \mu_3 = l_{31}F_3 + \varepsilon_3$$

Dimana $\sum = LL' + \Psi$

Atau $1 = l_{11}^2 + \psi_1$ 0.9= $l_{11}l_{21}$ 0.7= $l_{11}l_{31}$

$1 = l_{21}^2 + \psi_2$ 0.4= $l_{21}l_{31}$

$1 = l_{31}^2 + \psi_3$

Pasangan dari persamaan

$0.7 = l_{11} l_{31}$

$0.4 = l_{21} l_{31}$

Maka

$l_{21} = \frac{0.4}{0.7} l_{11}$

Subtitusikan l_{21} ke persamaan

$0.9 = l_{11} l_{21}$

Hasilnya l_{11}^2 = 1.575 atau l_{11} = $\mp$ 1.255. karena var(F_1) = 1 (dengan asumsi) dan var (X_1) = 1, l_{11} = Cov(X_1,F_1) = Corr(X_1,F_1). Koefisien korelasi tidak dapat lebih besar dari yang lainnya sehingga nilai|l_{11}| =1.255 terlalu besar. Dalam persamaannya juga

$1 = l_{11}^2 + \psi_1$ atau $\psi_1 = 1 - l_{11}^2$

Menghasilkan ψ_1=1-1.575 = -0.575

Yang hasilnya tidak memenuhi. Karena nilainya negative untuk Var(ε_1) = ψ_1.

Untuk m>1, ada hubungan yang tidak dapat dipisahkan dengan model factor. Untuk mengetahuinya, misal T matriks orthogonal m x m sehingga TT' = T'T = I. persamaan (9-2) dapat ditulis

$$X - \mu = LF + \varepsilon = LTT'F + \varepsilon = L^* F^* + \varepsilon$$

Dengan $L^* = LT$ dan $F^* = T'F$

Karena $E(F^*) = T'E(F) = 0$ dan $Cov(F^*) = T'Cov(F)T = T'T = I$

Factor F dan faktor $F^* = T'F$ punya bentuk yang sama .Walaupun loading L dan loading L^* berbeda tetapi menghasilkan matriks covarians Σ yang sama. Sehingga matriks covarians Σ yang dihasilkan adalah sebagai berikut.

$\Sigma = LL' + \Psi$= LTT'L'+ Ψ = (L^*)(L^*)'+ Ψ

Factor loading L dapat ditentukan pada matriks orthogonal T. sehingga loading

L dan $L^* = LT$

memberikan penyajian yang sama. Communality yang diberikan oleh LL' = (L^*)(L^*)', tidak dipengaruhi oleh pemilihan T.