THE CONDITIONAL AND UNCONDITIONAL MODELS OF FACTOR ANALYSIS AND THE NUMERICAL SOLUTION OF THEIR ESTIMATING EQUATIONS

The purpose of this note is to compare and to contrast the estimating equations of the conditional and the unconditional models of factor analysis.

The Models

The basic equation underlying both of the models of factor analysis can be written as

$$(1) \qquad y_{t.} = \mu_{t.} + \eta_{t.}; \quad t = 1, \ldots, T.$$

Here, $y_{t.}$ is a row vector of observations on G variables, $\mu_{t.}$ is vector of unobserved systematic variables and $\eta_{t.}$ is a vector of disturbances. It is assumed that successive disturbance vectors are uncorrelated and that

$$(2) \qquad E(\eta_{t.}) = 0 \quad \text{and} \quad D(\eta_{t.}) = \Omega,$$

where $\Omega = \text{diag}\{\omega_1, \ldots, \omega_M\}$ is a diagonal matrix. It is assumed that the systematic vector $\mu_{t.}$ is generated by $K = G - J$ latent factors such that

$$(3) \qquad \mu_{t.} = \xi_{t.}B'; \quad t = 1, \ldots, T,$$

where $\xi_{t.}$ is a row vector of K factors and B is a matrix of order $G \times K$ of $\text{Rank}(B) = K$ comprising the factor loadings.

In the conditional model, the unobserved factors $\xi_{t.}$ are regarded s fixed quantities. In the unconditional model, they have a statistical distribution such that

$$(4) \qquad E(\xi_{t.}) = 0 \quad \text{and} \quad D(\xi_{t.}) = \Phi.$$

It is common practice, when dealing with the unconditional model, to set $\Phi = I$. However, we shall assume that both Ω and Φ are diagonal matrices with G elements which have positive values.

For ease of notation, the T realisations of the relationships of (1) and (3) are compiled to give the following matrix equations:

$$(5) \qquad Y = M + H, \quad M = \Xi B'.$$

Here Y, M and H are matrices of order $T \times G$ which comprise, respectively, the realisations of $y_{t.}$, $\mu_{t.}$ and $\eta_{t.}$, whilst Ξ is a matrix of order $T \times K$ comprising the values of the factors $\xi_{t.}$

The Estimating Equations of the Conditional Model

On the assumption that the disturbance vector is normally distributed, we may write the likelihood function of the conditional model as

$$L(B,\Omega) = (2\pi)^{-GT/2}|\Omega|^{T/2}\exp\{\frac{1}{2}\text{Trace}(Y-M)'(Y-M)\Omega^{-1}\}. \tag{6}$$

The criterion for estimating B is to find the value which minimises

$$\text{Trace}\{(Y-\Xi B')'(Y-\Xi B')\Omega^{-1}\} = (Y-\Xi B')^{c'}(\Omega^{-1}\otimes I)(Y-\Xi B')^{c}, \tag{7}$$

The first step towards minimising the function is to find an expression for the minimising value of $M = \Xi B'$ when B is given. The value is

$$\Xi B' = Y\Omega^{-1}B(B'\Omega^{-1}B)^{-1})B'. \tag{8}$$

Putting this into (7) indicates that the criterion function takes the form of

$$\begin{aligned}&\text{Trace}\{[I-B(B'\Omega^{-1}B)^{-1}B'\Omega^{-1}]Y'Y[I-\Omega^{-1}B(B'\Omega^{-1}B)^{-1}B']\Omega^{-1}\}\\ &=\text{Trace}\{Y'Y\Omega^{-1}-Y'Y\Omega^{-1}B(B'\Omega^{-1}B)^{-1})B'\Omega^{-1}\}\end{aligned} \tag{9}$$

In order to identify a unique estimate of B, it is necessary to specify some further aspects of this matrix. Therefore, the following normalisation is imposed:

$$B'\Omega^{-1}B = I. \tag{10}$$

In that case, the criterion function for estimating B, which is derivable from (9), takes the form of

$$L(B) = \text{Trace}\{B'\Omega^{-1}Y'Y'\Omega^{-1}B\} - \text{Trace}\{\Lambda B'\Omega^{-1}B\}, \tag{11}$$

where Λ is a diagonal matrix of Lagrangean multipliers. Differentiating the function with respect to B^c—the long vector formed from B—gives

$$\begin{aligned}\frac{\partial L}{\partial B^c} &= (B'\Omega^{-1}Y'Y'\Omega^{-1})^r + (\Omega^{-1}Y'Y'\Omega^{-1}B)^rK\\ &\quad + (\Lambda B'\Omega^{-1})^r + (\Omega^{-1}B\Lambda)^rK,\end{aligned} \tag{12}$$

where the symbol r denotes the operation of forming a row vector from the rows of a matrix and K stands for the tensor commutator defined by $A^rK = A'^r$.

On setting the derivative to zero, we obtain, via some minor manipulations, the equation

$$Y'Y\Omega^{-1}B = B\Lambda, \tag{13}$$

which is the estimating equation for B.

In order to find an estimate of Ω, we must consider the criterion function

$$L(\Omega) = -T\log|\Omega| - \text{Trace}\{(Y-M)'(Y-M)\Omega^{-1}\}, \tag{14}$$

which is derivable from (6) by taking logarithms. Let $\omega = [\omega_{11}, \omega_{22}, \ldots, \omega_{MM}]$ be the vector of the diagonal elements of $\Omega = \Omega(\omega)$. Then, by using a chain rule to differentiate the function in respect of ω, we get

$$\frac{\partial L}{\partial \omega} - -T\Omega^r \frac{\partial(\Omega^{-1})^c}{\partial \omega} + \{(Y-M)'(Y-M)\}^r \frac{\partial(\Omega^{-1})^c}{\partial \omega}. \tag{15}$$

By setting this derivative to zero, we can obtain an equation which can be written as

$$T\Omega(\omega) = \text{diag}\{(Y-M)'(Y-M)\}. \tag{16}$$

By using the condition $B'\Omega^{-1}B = I$, we can obtain from (8) the expression

$$M = \Xi B' = Y\Omega^{-1}BB'. \tag{17}$$

Substituting the latter into (16) gives

$$\begin{aligned} T\Omega(\omega) &= \text{diag}\{Y'Y - Y'Y\Omega^{-1}BB' - BB'\Omega^{-1}Y'Y \\ &\qquad + BB'\Omega^{-1}Y'Y\Omega^{-1}BB'\} \\ &= \text{diag}\{Y'Y - B\Lambda B'\}, \end{aligned} \tag{17}$$

where the second equality follows from using the condition under (13) which also implies that $\Lambda = B'\Omega^{-1}Y'Y\Omega^{-1}B$.

Thus, by gathering equations (13) and (17), it can be seen that the estimating equations of the conditional model of factor analysis are given by

$$\begin{aligned} Y'Y\Omega^{-1}B = B\Lambda, \qquad \Lambda = B'\Omega^{-1}Y'Y\Omega^{-1}B \\ T\Omega = \text{diag}\{Y'Y - B\Lambda B'\}. \end{aligned} \tag{19}$$

The Estimating Equations of the Unconditional Model

The basic equation of the model can be written as

$$y_{t.} = \xi_{t.}B' + \eta_{t.}; \quad t = 1, \ldots, T. \tag{20}$$

Under the assumptions of the unconditional model, the dispersion matrix of $y_{t.}$ becomes

$$\begin{aligned} D(y_{t.}) &= B(\xi_{t.})B' + D(\eta_{t.}) \\ &= B\Phi B' + \Omega. \end{aligned} \tag{21}$$

Under the assumption that $\xi_{t.}$ and $\eta_{t.}$ are normally distributed, the likelihood function of the unconditional model can be written as

$$L(B, \Phi, \Omega) = (2\pi)^{MT/2}|B\Phi B'|^{T/2} \exp\{\frac{1}{2}\text{Trace}[Y'Y(B\Phi B' + \Omega)^{-1}]\}. \tag{22}$$

The parameters can be estimated by finding the values which minimise the function

$$L = -T \log |B\Phi B'| - \text{Trace}[Y'Y(B\Phi B' + \Omega)^{-1}. \tag{23}$$

Differentiating with respect to B^c gives

$$\frac{\partial L}{\partial B^c} = -2T\{\Phi B'(N\Phi B')^{-1}\}^r + 2\{\Phi B'(B\Phi B')^{-1}Y'Y\Phi B'(N\Phi B')^{-1}\}^r. \tag{24}$$

Setting this to zero and rearranging gives the first-order condition

$$T^{-1}Y'Y(B\Phi B' + \Omega)^{-1}B = B. \tag{25}$$

Now consider the identity

$$(B\Phi B' + \Omega)^{-1} = \Omega^{-1} - \Omega^{-1}B(B'\Omega^{-1}B + \Phi^{-1})^{-1}B'\Omega^{-1}. \tag{26}$$

By using the condition $B'\Omega^{-1}B = I$ and the fact that

$$\{\Phi(I - \Phi^{-1})\}\{I - (I - \Phi^{-1})^{-1}\} = I, \tag{27}$$

it can be see that

$$\begin{aligned}(B\Phi B' + \Omega)^{-1}B &= \Omega^{-1}B\{I - (B'\Omega^{-1}B + \Phi^{-1})B'\Omega^{-1}B\}\\ &= \Omega^{-1}B\{I - (I + \Phi^{-1})^{-1}\}\\ &= \Omega^{-1}B(I - \Phi)^{-1}.\end{aligned} \tag{28}$$

On substituting this result in (25), we get

$$T^{-1}Y'Y\Omega^{-1}B(I - \Phi)^{-1} = B, \tag{29}$$

or simply

$$T^{-1}Y'Y\Omega^{-1}B = B(I - \Phi), \tag{30}$$

which represents the estimating equation for B.

Now consider differentiating the criterion of (23) in respect of Ω^c. This gives

$$\begin{aligned}\frac{\partial L}{\partial \Omega^c} &= -T(B\Phi B' + \Omega)^{-1r} - Y'Y\frac{\partial (B\Phi B' + \Omega)^{-1c}}{\partial \Omega^c}\\ &= -T(B\Phi B' + \Omega)^{-1r} - (Y'Y)^r\{(B\Phi B' + \Omega)^{-1c} \otimes (B\Phi B' + \Omega)^{-1c}\}\\ &= -T(B\Phi B' + \Omega)^{-1r} - \{(B\Phi B' + \Omega)^{-1}Y'Y(B\Phi B' + \Omega)^{-1}\}^r.\end{aligned} \tag{31}$$

Using this derivative in the condition

$$\frac{\partial L}{\omega} = \frac{\partial L}{\Omega^c}\frac{\partial \Omega^c}{\omega} = 0 \tag{32}$$

results in the following estimating equation:

$$T\Omega = \text{diag}\{Y'Y - TB\Phi B'\}. \tag{33}$$

Now consider differentiating the criterion function in respect of Φ. This gives

$$\begin{aligned}\frac{\partial L}{\partial \Phi^c} &= -T(B\Phi B' + \Omega)^{-1r}(B \otimes B)\\ &\qquad + (Y'Y)^r\{(B\Phi B' + \Omega)^{-1} \otimes (B\Phi B' + \Omega)^{-1}\}(B \otimes B)\\ &= -T\{B'(B\Phi B' + \Omega)^{-1}B\}^r\\ &\qquad + T\{B'(B\Phi B' + \Omega)^{-1}Y'Y(B\Phi B' + \Omega)^{-1}B\}^r.\end{aligned} \tag{34}$$

The estimating equation for Φ is derived from the condition that

$$\text{diag}\{T\ B'(B\Phi B' + \Omega)^{-1}B - B'(B\Phi B' + \Omega)^{-1}Y'Y(B\Phi B' + \Omega)^{-1}B\} = 0. \tag{35}$$

Using the identity of (28) and the condition that $B'\Omega^{-1}B = I$, this can be written as

$$\text{diag}\{T(I + \Phi)^{-1} + \Omega)^{-1}B - (I + \Phi)^{-1}B'\Omega)^{-1}Y'Y\Omega)^{-1}B(I + \Phi)^{-1}\} = 0. \tag{36}$$

Since $I + \Phi$ is diagonal if Φ is diagonal, this is simply

$$T(I + \Phi) = B'\Omega^{-1}Y'Y\Omega^{-1}B = \Lambda. \tag{37}$$

By gathering the equations (30), (33) and (37), it can be seen that the estimating equations of the unconditional model of factor analysis are given by

$$\begin{aligned} Y'Y\Omega^{-1}B = TB(I + \Phi) = B'\Lambda, \qquad \Lambda = B'\Omega^{-1}Y'Y\Omega^{-1}B,\\ T\Omega = \text{diag}\{Y'Y - TB\Phi B'\}.\end{aligned} \tag{19}$$

Comparison with the equations under (19) shows that the estimating equations for the conditional model and the unconditional model differ only in respect of the estimator of the dispersion matrix Ω of the disturbances.