Posterior Distributions for the CDAFS models

For the CDAFS models,

$$y_{it} = k \iff \tau_{i,k-1} < z_{it} \le \tau_{i,k},\ \ i = 1, \ldots, m_i, t = 1, \ldots, \mathrm{T},$$

$$\mathbf{z}_t = \mathbf{\Lambda}\mathbf{f}_t,$$

$$\mathbf{f}_t = \sum_{l=1}^{L} \mathbf{B}_l \mathbf{f}_{t-l} + \mathbf{v}_t.$$

To simply the display of formulas, we only give the full conditional posterior distribution of the lag 1 CDAFS model. For more complex models, the forms of the posteriors should be the same.

Conditional distribution for $\mathbf{z}_t$

For the underlying variable $\mathbf{z}_t$, the conditional distribution is a truncated multivariate normal distribution,

$$\mathbf{z}_t | \mathbf{y}_t, \mathbf{\Lambda}, \mathbf{f}_t, \mathbf{Q}, \boldsymbol{\tau} \sim MN(\mathbf{\Lambda}\mathbf{f}_t, \mathbf{Q})\mathbf{I}(\mathbf{z}_t \in \mathbf{A}),$$

where MN represents the multivariate normal distribution, and $\mathbf{I}(\mathbf{z}_t \in \mathbf{A})$ is an indicator function which has the value 1 if $\mathbf{z}_t \in \mathbf{A}$ and 0 otherwise. Furthermore, $\mathbf{A}$ is a p-dimensional cube formed by the thresholds in a p-dimensional space,

$$\mathbf{A} = (\tau_{1,y_1-1}, \tau_{1,y_1}] \times, \ldots, \times (\tau_{p,y_p-1}, \tau_{p,y_p}]$$

To sample from the posterior distribution, the method described by Geweke (1991) can be used.

Conditional distribution for $\mathbf{f}_t$

Let the prior distribution of $\mathbf{f}_0$ be

$$\mathbf{f}_0 \sim MN(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0).$$

The posterior distribution of $\mathbf{f}_t$ is

$$\mathbf{f}_t|\mathbf{f}_{j\neq t}, \mathbf{f}_0, \mathbf{z}_t\boldsymbol{\Lambda}, \mathbf{B}, \mathbf{Q}, \mathbf{D} \sim MN(\boldsymbol{\Sigma}_t \cdot \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t),$$

with

$$\boldsymbol{\Sigma}_t^{-1} = \mathbf{D}^{-1} + \boldsymbol{\Lambda}^t\mathbf{Q}^{-1}\boldsymbol{\Lambda} + \mathbf{B}^t\mathbf{D}^{-1}\mathbf{B}$$

and

$$\boldsymbol{\mu}_t^t = \mathbf{f}_{t-1}^t\mathbf{B}\mathbf{D}^{-1} + \mathbf{z}_t^t\mathbf{Q}^{-1}\boldsymbol{\Lambda} + \mathbf{f}_{t+1}^t\mathbf{D}^{-1}\mathbf{B}.$$

Conditional distributions for $\boldsymbol{\Lambda}$ *and* $\mathbf{Q}$

Let $\boldsymbol{\lambda}_i$ be the ith row in the factor loading matrix $\boldsymbol{\Lambda}$ and the variance of e_{it} be q_i , the ith diagonal element of $\mathbf{Q}$.

Let the prior distribution of $\boldsymbol{\lambda}_i^t$ be

$$\boldsymbol{\lambda}_i^t \sim MN(\boldsymbol{\lambda}_{i0}^t, \boldsymbol{\Sigma}_{i0}).$$

Conditional on the other parameters and the observed data, the posterior distribution of $\boldsymbol{\lambda}_i^t$ is

$$\boldsymbol{\lambda}_i^t|\mathbf{f}_t, \mathbf{z}_t, q_i \sim MN(\boldsymbol{\lambda}_{i1}^t, \boldsymbol{\Sigma}_{i1}), \quad (1)$$

with

$$\boldsymbol{\lambda}_{i1}^t = (\boldsymbol{\Sigma}_{i0}^{-1} + q_i^{-1}\mathbf{X}^t\mathbf{X})^{-1}(\boldsymbol{\Sigma}_{i0}^{-1}\boldsymbol{\lambda}_{i0}^t + q_i^{-1}\mathbf{X}^t\mathbf{z}_i)$$

and

$$\mathbf{\Sigma}_{i1} = (\mathbf{\Sigma}_{i0}^{-1} + q_i^{-1}\mathbf{X}^t\mathbf{X})^{-1},$$

where $\mathbf{X}$ is a $q \times \mathrm{T}$ matrix of factor scores and $\mathbf{z}_i$ is a $\mathrm{T} \times 1$ vector with the observations for the ith underlying variable.

Given the prior distribution of q_i,

$$q_i \sim IG(\frac{v_0}{2}, \frac{\delta_0}{2}),$$

the posterior of q_i is

$$q_i|\boldsymbol{\lambda}_i, \mathbf{f}_t, \mathbf{z}_t \sim IG(\frac{v_1}{2}, \frac{\delta_1}{2}), \tag{2}$$

with

$$v_1 = v_0 + \mathrm{T},$$

$$\delta_1 = \delta_0 + \sum_{t=1}^{\mathrm{T}} (z_{it} - \boldsymbol{\lambda}_i \mathbf{f}_t)^2,$$

where IG represents the inverse Γ distribution.

Conditional distributions for **B** *and* **D**

If we give **B** and **D** multivariate normal and Wishart prior distributions, we can obtain their conjugate posterior distribution as in Zhang, Hamaker, and Nesselroade (Under Review). In the current study, we assume that **D** is a correlation matrix for identification purpose. We then give a uniform distribution U(-1,1) for each off-diagonal element in **D** as suggested by Chib and Greenberg (1998). We derive the posterior distribution accordingly.

Let $\mathbf{b}_i$ represent the ith row of $\mathbf{B}$ with $i = 1, \ldots, p$. Given the prior distribution of $\mathbf{b}_i$ be

$$\mathbf{b}_i \sim MN(\mathbf{b}_{i.0}, \Sigma_{bi.0}),$$

The posterior distribution is

$$\mathbf{b}_i | \mathbf{f}_t, \mathbf{D}, \mathbf{y}_t \ \sim MN(\mathbf{b}_{i.1}, \Sigma_{bi.1}),$$

where

$$\mathbf{b}_{i.1} = \Sigma_{bi.1}(\Sigma_{bi.0}^{-1}\mathbf{b}_{i.0} + \sum_{t=1}^{\mathrm{T}}[\mathbf{f}_{t-1}\mathbf{d}_i\mathbf{f}_t - \sum_{j \neq i}^{q} \mathbf{f}_{t-1} d_{ji} \mathbf{b}_j \mathbf{f}_{t-1}]),$$

and

$$\Sigma_{bi.1} = (\Sigma_{bi.0}^{-1} + \sum_{t=1}^{\mathrm{T}} \mathbf{f}_{t-1} d_{ii} \mathbf{f}'_{t-1})^{-1}.$$

In the above formula, d_{ij} is the (i,j)th element in the matrix $\mathbf{D}$ and $\mathbf{d}_i = (d_{i1}, \ldots, d_{iq})$.

For the correlation matrix D, we give all the off-diagonal elements a uniform prior distribution,

$$d_{ij} \sim U(-1, 1), i = 1, \ldots, q-1, j = 2, \ldots, q.$$

There is no closed form for the posterior distribution of d_{ij}. However, the Metropolis-Hastings algorithm used in Chib and Greenberg (1998) can be employed to sample the posterior distribution.

In the current study, we first calculated the thresholds by using the method in Olsson (1979) and then fixed all the thresholds in the Bayesian step. However, we can also obtain the posterior distribution of the thresholds and use the Metropolis-Hastings method to sample the thresholds from the posterior distributions (e.g., Song & Lee, 2002).

References

Chib, S., & Greenberg, E. (1998). Analysis of multivariate probit models. *Biometrika*, *85*(2), 347-361.

Geweke, J. (1991). Efficient simulation from the multivariate normal and student t distribution subject to linear constraints. In E. M. Keramidas (Ed.), *Computing science and statistics: Proceedings of the twenty third symposium on the interface* (p. 571-578). Alexandria, VA: American Statistical Association.

Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, *44*, 443-460.

Song, X.-Y., & Lee, S.-Y. (2002). Bayesian estimation and model selection of multivariate linear model with polytomous variables. *Multivariate Behavioral Research*, *37*, 453-477.

Zhang, Z., Hamaker, E. L., & Nesselroade, J. R. (Under Review). Comparisons of four methods for estimating dynamic factor models. *Structural Equation Modeling.*