Supplement:

Explore Conditional Dependencies in Item Response Tree Data

1 Four Models

Here I provide the specific formulation of the four models utilized for the verbal aggression data. I assume the three-category item response data have been expanded to two sets of sub-responses based on a linear tree with two sub-trees (two nodes and two branches).

Denote $Y_{p,i} = m$ as the raw response $(m = 0, 1, 2)$ from person p $(p = 1, \ldots, N)$ to item i $(i = 1, \ldots, I)$. $m = 0$ indicates 'No', $m = 1$ indicates 'Perhaps', and $m = 2$ indicates 'Yes' in the verbal aggression data example. Let $Y^*_{p,i,k}$ denote person p's sub-response to item i in Node k following the mapping matrix given below:

$Y_{p,i}$	$Y^*_{p,i,1}$	$Y^*_{p,i,2}$
0	0	-
1	1	0
2	1	1

The conditional probabilities of $Y^*_{p,i,k} = 1$ is formulated in the four models (M1, M2, and M3, and M4) as follows:

- M1: IRTree model with one latent trait (sequential IRT model)

$$\text{logit}\left(\Pr\left(Y^*_{p,i,k} = 1 | \theta_p, \alpha_{i,k}, \beta_{i,k}\right)\right) = \alpha_{i,k}\theta_p + \beta_{i,k}, \quad \theta_p \sim \mathsf{N}(0, 1), \tag{1}$$

  where $\alpha_{i,k} \in \mathbb{R}$ and $\beta_{i,k} \in \mathbb{R}$ are the item slope and intercept for item i in node k, and $\theta_p \in \mathbb{R}$ is the latent trait for person p.

- M2: IRTree model with two node-specific latent traits

$$\text{logit}\left(\Pr\left(Y^*_{p,i,k} = 1 | \theta_{p,k}, \alpha_{i,k}, \beta_{i,k}\right)\right) = \alpha_{i,k}\theta_{p,k} + \beta_{i,k}, \quad (\theta_{p,1}, \theta_{p,2})' \sim \mathsf{N}(\mathbf{0}, \Sigma_\theta), \tag{2}$$

where $\alpha_{i,k} \in \mathbb{R}$ and $\beta_{i,k} \in \mathbb{R}$ are the item slope and intercept for item i in node k, and $\theta_{p,k} \in \mathbb{R}$ is the latent trait for person p in node k. Σ_θ is a two-by-two covariance matrix.

- M3: LSIRM with one latent trait

$$\text{logit}\left(\Pr\left(Y^*_{p,i,k} = 1|\theta_p, \alpha_{i,k}, \beta_{i,k}, \mathbf{a}_p, \mathbf{b}_i\right)\right) = \alpha_{i,k}\theta_p + \beta_{i,k} - \gamma d(\mathbf{a}_p, \mathbf{b}_i), \quad \theta_p \sim \mathsf{N}(0,1), \tag{3}$$

where where $\alpha_{i,k} \in \mathbb{R}$ and $\beta_{i,k} \in \mathbb{R}$ are the item slope and intercept for item i in node k, and $\theta_p \in \mathbb{R}$ is the latent trait for person p. In addition, $\gamma > 0$ is the weight of the distance term and $d(\mathbf{a}_p, \mathbf{b}_i) = ||\mathbf{a}_p - \mathbf{b}_i||_2$ is a two-dimensional Euclidean distance, and $\mathbf{a}_p \in \mathbb{R}^2$ and $\mathbf{b}_i \in \mathbb{R}^2$ are the latent positions of person p and item i in the two-dimensional space.

- M4: LSIRM with two node-specific latent traits

$$\text{logit}\left(\Pr\left(Y^*_{p,i,k} = 1|\theta_{p,k}, \alpha_{i,k}, \beta_{i,k}, \mathbf{a}_p, \mathbf{b}_i\right)\right) = \alpha_{i,k}\theta_{p,k} + \beta_{i,k} - \gamma d(\mathbf{a}_p, \mathbf{b}_i), \quad (\theta_{p,1}, \theta_{p,2})' \sim \mathsf{N}(\mathbf{0}, \Sigma_\theta), \tag{4}$$

where $\alpha_{i,k} \in \mathbb{R}$ and $\beta_{i,k} \in \mathbb{R}$ are the item slope and intercept for item i in node k, and $\theta_p \in \mathbb{R}$ is the latent trait for person p. Σ_θ is a two-by-two covariance matrix. In addition, $\gamma > 0$ is the weight of the distance term and $d(\mathbf{a}_p, \mathbf{b}_i) = ||\mathbf{a}_p - \mathbf{b}_i||_2$ is a two-dimensional Euclidean distance, and $\mathbf{a}_p \in \mathbb{R}^2$ and $\mathbf{b}_i \in \mathbb{R}^2$ are the latent positions of person p and item i in the two-dimensional space.

2 Estimation

A fully Bayesian approach is applied to estimate the four models based on the following priors: $\alpha_{i,k} \sim \mathsf{N}(0, 3^2),\ k = 1, 2, \quad \beta_{i,k} \sim \mathsf{N}(0, 3^2),\ k = 1, 2, \quad \theta_{p,k} \sim \mathsf{N}(0, 1),\ k = 1, 2, \quad c_\theta \sim \mathsf{N}(0, 3^2)$, $\gamma \sim$ Half $\mathsf{N}(0, 1), \quad a_{p,d} \sim \mathsf{N}(0, 1),\ d = 1, 2, \quad b_{i,d} \sim \mathsf{N}(0, 1),\ d = 1, 2$. Here c_θ is the Cholesky element, such that $\Sigma_\theta = L \cdot L'$, where $L = \begin{bmatrix} 1 & 0 \\ c_\theta & 1 \end{bmatrix}$. The diagonal elements of L are set to 1 for the identifiability of the covariance matrix. In this case, the correlation is $\rho_\theta = c_\theta / \sqrt{1 + c_\theta^2}$. The estimation is implemented by using NIMBLE (de Valpine et al., 2017). In all analyses, the number of iterations and burn-in are set to 15,000 and 5,000, respectively. All analyses show reasonable convergence.

References

de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple Lang, D., and Bodik, R. (2017). Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26:403–413.