Hierarchical Dynamic Factor Model

Yi Chen

1 Model

$$u_{icjt} = \sum_{p=1}^{P_0} \beta_{icpt} \cdot x_{pcjt} + \epsilon_{icjt}$$

where i indexes customer, c indexes category, j index product, and t indexes time. P_c is the number of predictors of brand choice in category c, i.e., the covariate X's dimension.

$$\beta_{icp}(t) = \alpha_{icp} + \sum_{l_G=1}^{L_G} w_{icpl_G}^{\text{Glob}} \cdot \theta_{l_G}^{\text{Glob}}(t) + \sum_{l_c=1}^{L_c} w_{icpl_c}^{\text{Cat}} \cdot \theta_{cl_c}^{\text{Cat}}(t) + \sum_{l_i=1}^{L_i} w_{icpl_i}^{\text{Ind}} \cdot \theta_{il_i}^{\text{Ind}}(t)$$

where

- α_{icp}: the time-invariant level of parameter cp for customer i
- $\theta_{l_G}^{\text{Glob}}$: L_G global factors of all parameters for all customers
- θ_c^{Cat}: L_c factors for category c across all customers
- θ_i^{Ind}: L_i factors for individual i across all categories

Let $s = (c, p)$ and $w_{isl} = \log(1 + \exp(\tilde{w}_{isl}))$ then model

$$\tilde{w}_{il} = (\tilde{w}_{i1l}, ..., \tilde{w}_{iSl}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w)$$

In this way, we can model the correlations between all parameters.
For Global factors:

$$w_{isl}^{\text{Glob}} = \log(1 + \exp(\tilde{w}_{isl}^{\text{Glob}}))$$
$$\tilde{w}_{il}^{\text{Glob}} = (\tilde{w}_{i1l}^{\text{Glob}}, ..., \tilde{w}_{iSl}^{\text{Glob}}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w^{\text{Glob}})$$

For Category factors (K categories in total):

$$w_{icl}^{\text{Cat}} = \log(1 + \exp(\tilde{w}_{icl}^{\text{Cat}}))$$
$$\tilde{w}_{il}^{\text{Cat}} = (\tilde{w}_{i1l}^{\text{Cat}}, ..., \tilde{w}_{iKl}^{\text{Cat}}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w^{\text{Cat}})$$

For Individual factors (N categories in total):
How do we model this? Should individuals be modelled correlated???

Priors:

1. For every latent factor l:
   - $\theta_l(t) \sim \mathcal{GP}(m_l, k(t, t'; 1, \rho_l))$, where $k(t, t'; 1, \rho_l)$ is a Matern-3/2 kernel.
   - $\rho_l \sim \text{Weibull}(0.1, 1)$

2. For w: $\Sigma_w = \text{diag}(\tau)\Lambda_w\text{diag}(\tau)$, where $\tau \sim \text{HalfNormal}(5), \Lambda_w \sim LKJ(2)$
   (Then for every individual i and latent factor l, $\tilde{w}_{il} = (\tilde{w}_{i1l}, ..., \tilde{w}_{iSl}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w)$)

3. For α:
   $\mu_{\alpha,s} \sim \mathcal{N}(0, 5)$ for each s, and $\mu_\alpha = (\mu_{\alpha,1}, ..., \mu_{\alpha,S})$
   $\Sigma_\alpha = \text{diag}(\tau_\alpha)\Lambda_\alpha\text{diag}(\tau_\alpha)$, where $\tau_\alpha \sim \text{HalfNormal}(5), \Lambda_\alpha \sim LKJ(2)$
   For every individual i:
   $\alpha_i = (\alpha_{1l}, ..., \alpha_{1S}) \sim \mathcal{N}(\mu_\alpha, \Sigma_\alpha)$

2 Simplified Model

Suppose we have N customers, K categories, M products in each category, and T time periods. And suppose there are L latent factors.

We start with the simplest version of this model. Suppose $P_c = 1$ for all categories. That means, X has only one dimension. Then

$$u_{icjt} = \beta_{ict} \cdot x_{cjt} + \epsilon_{icjt}$$

And suppose every latent factor is global.

$$\beta_{ic}(t) = \alpha_{ic} + \sum_{l=1}^{L} w_{icl} \cdot \theta_l(t)$$

Then we let $w_{icl} = \log(1 + \exp(\tilde{w}_{icl}))$, and $\tilde{\mathbf{w}}_{il} = (\tilde{w}_{i1l}, ..., \tilde{w}_{iKl})$
Priors:

1. For every latent factor l:
   - $\rho_l \sim \text{Weibull}(0.1, 1)$
   - $\theta_l(t) \sim \mathcal{GP}(m_l, k(t, t'; 1, \rho_l))$, where $k(t, t'; 1, \rho_l)$ is a Matern-3/2 kernel.

2. For w: $\Sigma_w = \text{diag}(\tau)\Lambda_w\text{diag}(\tau)$, where $\tau \sim \text{HalfNormal}(5), \Lambda_w \sim LKJ(2)$
   Then for every individual i and latent factor l, $\tilde{w}_{il} = (\tilde{w}_{i1l}, ..., \tilde{w}_{iKl}) \sim \mathcal{N}(\mathbf{0}, \Sigma_w)$)

3. For α:
   $\mu_{\alpha,c} \sim \mathcal{N}(0, 5)$ for each c, and $\mu_\alpha = (\mu_{\alpha,1}, ..., \mu_{\alpha,K})$
   $\Sigma_\alpha = \text{diag}(\tau_\alpha)\Lambda_\alpha\text{diag}(\tau_\alpha)$, where $\tau_\alpha \sim \text{HalfNormal}(5), \Lambda_\alpha \sim LKJ(2)$ (Σ_α is a $K \times K$ matrix)
   For every individual i:
   $\alpha_i = (\alpha_{i1}, ..., \alpha_{1K}) \sim \mathcal{N}(\mu_\alpha, \Sigma_\alpha)$