

---

## Description of Ordinal Cluster Model

**Response model model:** Let  $X_{ij} \in \{0, \dots, k_{max}\}$ , where  $1 \leq j \leq n_{respondents}$  is the respondent index and  $1 \leq i \leq n_{items}$  is the item index. Our parameters are respondent parameters  $\theta_j$ , item parameters for each ordinal response category  $\tau_{i,k}$ , respondent latent space positions  $z_j$ , and item-response latent space parameters  $w_{i,k}$ . Then for  $0 \leq k \leq k_{max}$ ,

$$P(X_{ij} = k) = c \exp(\theta_j + \tau_{i,k} - \|z_j - w_{i,k}\|_2^2)$$

Where  $c$  is the normalization constant  $c = \sum_{k=0}^{k_{max}} \exp(\theta_j + \tau_{i,k} - \|z_j - w_{i,k}\|_2^2)$ ; this is just one way of writing the softmax function.

**Priors** On the latent space parameters  $z_j$  and  $w_{i,k}$  we assign a Gaussian mixture prior with fixed number of clusters  $G$ :

$$z_j, w_{i,k} \sim \sum_{g=1}^G \lambda_g MVN_d(\mu_g, \sigma_g I_d)$$

The prior probability for cluster assignment is Dirichlet with parameter  $\nu$  for conjugacy:

$$[\lambda_1, \dots, \lambda_G] \sim \text{Dirichlet}(\nu)$$

and means are given normal hyperiors:

$$\forall g : \mu_g \sim MVN_d(0, \sigma_\mu^2 I_d)$$

All prior variances are given uninformative Inverse-Gamma hyperiors:

$$\sigma_\theta^2, \sigma_\tau^2, \sigma_\mu^2, \forall g : \sigma_g \sim \text{InvGamma}(\epsilon, \epsilon)$$

where  $\epsilon$  is set to be sufficiently small as to be uninformative. If the sampler behaves poorly by estimating prior variances that are unsuitably large or small, these priors may be changed to be informative in order to force estimated parameters to fall within a reasonable range.

**Sampling.** Sampling is Gibbs on each set of parameters. For parameters without conjugate Gibbs distributions, Metropolis-Hastings is used with normal proposals and variances adjusted adaptively throughout the sampling (see <http://probability.ca/jeff/ftpd/adaptex.pdf>).

---

The parameters  $\mu_g, \sigma_g^2$ , and  $K_{z_j}, K_{w_i}$  have conjugate Gibbs distributions and can be sampled directly. This method is analogous to that used in Handcock, Raftery, and Tantrum (2007).

$$\begin{aligned}\mu_g | \text{others} &\sim MVN \left( \frac{m_g \bar{z}_g}{m_g + \sigma_g^2 / \omega^2}, \frac{\sigma_g^2}{m_g + \sigma_g^2 / \omega^2} \right) \\ \sigma_g^2 | \text{others} &\sim \text{InvGamma} \left( \epsilon + \frac{m_g}{2}, \epsilon + \frac{m_g}{2} (\bar{z}_g - \mu_g) \right) \\ P(K_{z_j} = g | \text{others}) &= \frac{\lambda_g \phi_d(z_j; \mu_g, \sigma_g^2 I_d)}{\sum_{g=1}^G \lambda_g \phi_d(z_j; \mu_g, \sigma_g^2 I_d)} \\ P(K_{w_{i,k}} = g | \text{others}) &= \frac{\lambda_g \phi_d(w_{i,k}; \mu_g, \sigma_g^2 I_d)}{\sum_{g=1}^G \lambda_g \phi_d(w_{i,k}; \mu_g, \sigma_g^2 I_d)}\end{aligned}$$

where

$$\begin{aligned}m_g &= \sum_{k: K_{z_j}=g} 1 + \sum_{i: K_{w_{i,k}}=g} 1 \\ s_g^2 &= \frac{1}{d} \left( \sum_{k: K_{z_j}=g} (z_j - \mu_g)^T (z_j - \mu_g) + \sum_{i: K_{w_{i,k}}=g} (w_{i,k} - \mu_g)^T (w_{i,k} - \mu_g) \right) \\ \bar{z}_g &= \frac{1}{m_g} \left( \sum_{k: K_{z_j}=g} z_j + \sum_{k: K_{z_j}=g} z_j \right)\end{aligned}$$

and  $\phi_d$  is the d-dimensional multivariate normal density.

**Overall procedure.** We perform the steps in the following order:

1. Propose and update  $z$  and  $w$ .
2. Propose and update  $\theta$ , and  $\tau$ , sampling  $\sigma_\theta^2$  and  $\sigma_\tau^2$  after each respectively.
3. Sample  $K, \mu_g, \sigma_g^2, \lambda_g$ , and  $\sigma_\mu^2$  from their posterior distributions.