Notes to understand the notations

1 Notations in Teh et al.

Paper, see Sec.5 for Inference

1.1 Indices

- j is a document index
- i is an observation index (j, i is observation i in document j)
- k is a word index
- t is a movement mode index

1.2 Notations

- $x_{j,i}$ is observation i in document j (a word index)
- $z_{j,i}$ is the movement mode associated to observation $x_{j,i}$ (a movement mode index)
- $m_{j,k}$ is the number of movement modes in document j that have at least one obervation of word k

2 Notations in Wang et al.

Paper, see Sec.6 for pseudo-algo

2.1 Indices

- j is a document index
- i is an observation index (j, i is observation i in document j)
- k is a movement mode index

2.2 Notations

- $x_{j,i}$ is observation i in document j (a word index)
- $z_{j,i}$ is the movement mode associated to observation $x_{j,i}$ (a movement mode index)
- $t_{j,i}$ is the occurrence in which observation i in doc j is assigned
- $k_{j,t}$ is the movement mode associated to occurrence t in doc j
- $m_{j,k}$ is the number of occurrences in document j that are assigned movement mode k
- $n_{j,t,k}$ is the number of observations in occurrence t of movement mode k
- $\pi_{0,k}$ is the weight of movement mode k in the overall distribution G_0
- $\tilde{\pi}_{c,k}$ is the weight of movement mode k in the distribution related to cluster c: $\tilde{G}_c$

2.3 Algo

Step 1. At step 1. in their algorithm, they assume :

- fixed cluster assignment c_j for document j
- sampling $z_{j,i}$, $\pi_{0,k}$ and $\tilde{\pi}_{c,k}$ is sufficient

Sampling $z_{j,i}$ can be done using Eq.(37) in Teh et al. where we use :

$$f_k^{-(j,i)}(x_{j,i}) \propto \alpha_0 N_k^{-(j,i)}(x_{j,i})$$

where $N_k(w)$ is the number of occurrences of word w assigned to tables which movement mode is k. This value is 0 when k is a new movement mode.

Then $z_{j,i}$ is sampled using :

$$p(z_{j,i} = k|\mathbf{z}^{-(j,i)}, \mathbf{m}, \beta) = (n_{j,\cdot,k}^{-(j,i)} + \alpha_0\beta_k) f_k^{-(j,i)}(x_{j,i})$$

where $n_{j,t,k} = 0$ and $\beta_k = \beta_u$ if k is a new movement mode.

Also, $\pi_{0,k}$ is sampled from a DP according to Eq(36) in Teh et al (β_k in Teh is $\pi_{0,k}$ in Wang) :

$$(\beta_1, \ldots, \beta_K, \beta_u) \sim \text{Dirichlet}(m_{\cdot,1}, \ldots, m_{\cdot,K}, \gamma)$$

Similarly, $\tilde{\pi}_{c,k}$ is sampled using only information from documents assigned to cluster c.

Step 2. At step 2, $z_{j,i}$, $\pi_{0,k}$ and $\tilde{\pi}_{c,k}$ are fixed and we sample cluster assignments c_j using Chinese restaurant process :

Eq(34) in Teh where we operate at the document level instead of observation level :

$$p(c_j = c|\mathbf{c}^{-(j)}) \propto Ndocs_c^{-(j)} f_c^{-(j)}(\mathbf{x}_{j,\cdot})$$

TODO : Pierre, can you elaborate on $f_c^{-(j)}(\mathbf{x}_{j,\cdot})$?

Step 3. Sample beta_clusters based on Eq.(36) adapted at the cluster level.