NONNEGATIVE COLLABORATIVE TOPIC MODEL

1. GENERATIVE PROCESS

Document model:

- Draw topics $\beta_{vk} \sim \text{Gamma}(a, b)$.
- Draw document topic weights $\theta_{dk} \sim \text{Gamma}(c, d)$.
- Draw document per-topic word counts $z_{dv,k} \sim \text{Poisson}(\theta_{dk}\beta_{vk})$.
- Set total word counts $w_{dv} = \sum_k z_{dvk}$.

Recommendation model:

- Draw user weights $x_{uk} \sim \text{Gamma}(e, f)$.
- Draw document user weights $\epsilon_{dk} \sim \text{Gamma}(g, h)$.
- Draw document topic scaling $a_d \sim \text{Gamma}(r, s)$.
- Draw $y^a_{ud,k} \sim \text{Poisson}(x_{uk}a_d\theta_{dk})$.
- Draw $y^b_{ud,k} \sim \text{Poisson}(x_{uk}a_d\epsilon_{dk})$.
- $r_{ud} = \sum_k y^a_{ud,k} + y^b_{ud,k}$

2. VARIATIONAL INFERENCE

The model is conjugate. Marginalizing out the auxiliary variables, y and z,

$$w_{dv} \sim \text{Poisson}(\theta_d^T \beta_v)$$

$$r_{ud} \sim \text{Poisson}(x_u^T a_d(\theta_d + \epsilon_d)). \tag{1}$$

3. COMPLETE CONDITIONALS

$$\theta_{dk}|\beta, x, a, z, y, w, r \sim \text{Gamma}(c + \sum_v z_{dv,k} + \sum_u y^a_{ud,k}, d + \sum_v \beta_{vk} + a_d \sum_u x_{uk})$$

$$\beta_{vk}|\theta, z, w \sim \text{Gamma}(a + \sum_d z_{dv,k}, b + \sum_d \theta_{dk})$$

$$x_{uk}|\theta, y, r \sim \text{Gamma}(e + \sum_d y^a_{ud,k} + \sum_d y^b_{ud,k}, f + \sum_d a_d(\theta_{dk} + \epsilon_{dk}))$$

$$\epsilon_{dk}|x, y, r \sim \text{Gamma}(g + \sum_u y^b_{ud,k}, h + a_d \sum_u x_{uk})$$

$$a_d|\theta, x, r, s \sim \text{Gamma}(r + \sum_u y_{ud}, h + \sum_k (\theta_{dk} + \epsilon_{dk}) \sum_u x_{uk})$$

(2)

4. UPDATES FOR THE MULTINOMIALS

The complete conditionals for the z_{dv} are straightforward to write down:

$$z_{dv}|\theta_d, \beta, w_{dv} \sim \text{Mult}(w_{dv}, \frac{\theta_d \cdot \beta_v}{\sum_k \theta_{dk}\beta_{vk}})$$

(3)

The complete conditional for vector $y_{ud} = \{y^a_{ud,0}, y^a_{ud,1}, .., y^a_{ud,k-1}, y^b_{ud,0}, y^b_{ud,1}, .., y^b_{ud,k-1}\}$ of size $2K$ is

$$y_{ud}|r_{ud}, x_u, \theta_d, \epsilon_d, a_d \sim \text{Mult}\Big(r_{ud}, \tau_{ud}) \tag{4}$$

where τ_{ud} is given by

$$\tau_{ud,k} = \begin{cases} x_{uk}a_d\theta_{dk} & \text{if } k < K, \\ x_{uk}a_d\epsilon_{dk} & \text{if } K \leq k < 2K. \end{cases} \tag{5}$$

5. BATCH ALGORITHM WITH PARALLEL UPDATES

Repeat until convergence:

(1) For each document, update the variational multinomial for z_{dv}.

(2) For each user/document such that $r_{ud} > 0$, update the variational multinomials for y^a_{ud} and y^b_{ud}.

(3) In parallel update the following with full natural gradients computed using the multinomials from steps 1 and 2. Use a a step size of one.

- Update all document topic weight parameters
- Update topic parameters.
- Update user weights, document user weights and document topic scaling weights

6. ALTERNATE STOCHASTIC ALGORITHM

Repeat until convergence:

(1) Sample a document d.

(2) Update the variational multinomial for all z_{dv} in the document.

(3) For each user u rating the document, such that $r_{ud} > 0$, update the variational multinomials for y^a_{ud} and y^b_{ud}.

(4) In parallel update the parameters of sampled document and users with scaled natural gradients.

- Update document d topic weight parameters
- Update topic parameters.
- Update user weights, document user weights and document topic scaling weights of sampled users and document.