

# 1 Explicit Feedback

For explicit feedback, we assume that a pair of purchase and click  $\langle w, l \rangle$  is the result that (1)  $w$  is superior than  $l$  on the most pertinent aspect  $g_k = 1$  and (2)  $w$  at least ties with  $l$  on other aspects.

We have the following equation:

$$p(\langle w, l \rangle | \Theta, g) = \prod_{k=1}^K \frac{w_k}{w_k + \theta l_k}^{g_k} \frac{\theta w_k}{l_k + \theta w_k}^{1-g_k}, \quad (1)$$

where  $g$  the hidden variable is sampled from  $u$ ,  $\Theta = \{\theta, w_k, l_k, u\}$  is the parameter space.

The likelihood over all sessions is defined by

$$p(D|\Theta) = \prod_u \prod_{d \in D(u)} \sum_{\Theta} \left( \prod_{w \in W(d)} \prod_{l \in L(d)} p(\langle w, l \rangle | \Theta, g) p(g|u) \right) \quad (2)$$

Due to the summation part, the log-likelihood can not be analytically optimized. To maximize the log-likelihood, we follow the EM framework.

## E-step

$$\begin{aligned} p(g|d, \Theta^t) &\propto p(g|u, \Theta^t) p(d|g, \Theta^t) \\ p(g_k = 1|d, \Theta^t) &\propto u_k^t \frac{w_k^t}{w_k^t + \theta^t l_k^t} \prod_{k' \neq k} \left[ \frac{\theta^t w_{k'}^t}{l_{k'}^t + \theta^t w_{k'}^t} \right] \end{aligned} \quad (3)$$

## M-step

For any positive variables  $x, y$ , we have

$$\ln \frac{y}{x} \geq 1 - \frac{x}{y}$$

Therefore we can derive a lower bound for the log-likelihood over the complete data, given the parameters learnt from previous round.

$$\begin{aligned} l(\Theta, g) &= \sum_u \sum_{d \in D(u)} \log \left\{ u_k \frac{w_k}{w_k + \theta l_k} \prod_{k' \neq k} \left[ \frac{\theta w_{k'}^t}{l_{k'}^t + \theta w_{k'}^t} \right] \right\} \\ &\geq Q^t(\Theta, g) \\ &= \sum_u \sum_{d \in D(u)} \left\{ \log u_k + \left[ 1 - \frac{w_k + \theta l_k}{w_k^t + \theta^t l_k^t} + \log \frac{w_k}{w_k^t + \theta^t l_k^t} \right] \sum_{k' \neq k} \left[ 1 - \frac{l_{k'} + \theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t} + \log \frac{\theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t} \right] \right\} \end{aligned} \quad (4)$$

where  $l_k^t, w_k^t, \theta^t$  are the parameters learnt from  $t - th$  round.

The objective in M-step is  $Q^t = \sum_u \sum_{d \in D(u)} p(g_k = 1|d, \Theta^t) Q^t(\Theta, g)$ . To maximize  $Q^t$  we adopt coordinate descent.

Fix  $v$  for all  $w$  and  $l$  and  $\theta$ , we update  $u$  by. (See the corresponding derivation in our paper)

Fix  $u, \theta$ , we update  $w$  and  $l$  by.

Fix  $u, v$ , we update  $\theta$  by.

**Stochastic S-step** In stochastic EM, we simply add an S-step after the E-step. We draw the value of  $g$  for each session  $d$  by

$$k \sim u_k^t \frac{w_k^t}{w_k^t + \theta^t l_k^t} \prod_{k' \neq k} [\frac{\theta^t w_{k'}^t}{l_{k'}^t + \theta^t w_{k'}^t}]. \quad (5)$$

Next replace the M-step by

$$\sum_u \sum_{d \in D(u)} \{ [1 - \frac{w_k + \theta l_k}{w_k^t + \theta^t l_k^t} + \log \frac{w_k}{w_k^t + \theta^t l_k^t}] \sum_{k' \neq k} [1 - \frac{l_{k'} + \theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t} + \log \frac{\theta w_{k'}}{l_{k'}^t + \theta^t w_{k'}^t}] \} \quad (6)$$

## 2 Implicit Feedback