ExplicitAndImplicitFeedback

Xiaolin Shen

April 2018

1 Original EM Derivation

1.1 E-step

1.2 M-step

the updates of the parameters $\{u, v, \theta\}$ as follow:

$$u_k = \frac{\Sigma_{u(d)=u}\gamma(d,k,\Theta^t)}{\Sigma_{s=1}^{K}\Sigma_{u(d)=u}\gamma(d,s,\Theta^t)} \tag{1}$$

2 Stochastic EM Derivation

2.1 E-step

$$\begin{aligned} p(g|d,\Theta^t) &\propto p(g|u,\Theta^t)p(d|g,\Theta^t) \\ p(g_k=1|d,\Theta^t) &\propto u_k^t \frac{w_k^t}{w_k^t+\theta^t l_k^t} \prod_{k'\neq k}[\frac{\theta^t w_{k'}^t}{l_{k'}^t+\theta^t w_{k'}^t}] \end{aligned} \tag{2}$$

First, let's use $\gamma(d,k,\Theta^t)$ to denote the conditional probability $p(g_k=1|d,\Theta^t)$ given parameters in the $t-$th round, when the current session specific favorite aspect is $g_k=1$, defined as follows

$$\begin{aligned} \gamma(d,k,\Theta^t) = \frac{p(d,g|\Theta^t)}{\Sigma_g p(d,g|\Theta^t)} &= \frac{p(g|\Theta^t)p(d|\Theta^t,g)}{\Sigma_g p(g|\Theta^t)p(d|\Theta^t,g)} \\ &= \frac{u_k \Pi_{w\in W^d, v\in L^d} \frac{w_k}{w_k+\theta^t v_k} \Pi_{k'\neq k} \frac{\theta^t w_{k'}}{v_{k'}+\theta^t w_{k'}}}{\Sigma_{k=1}^{K} u_k \Pi_{w\in W^d, v\in L^d} \frac{w_k}{w_k+\theta^t v_k} \Pi_{k'\neq k} \frac{\theta^t w_{k'}}{v_{k'}+\theta^t w_{k'}}} \end{aligned} \tag{3}$$

Note that $\forall d, \Sigma_k \gamma(d,k,\Theta^t) = 1$.

2.2 S-step

then simply add an S-step after the E-step,the value of g for each session d is

$$k \sim u_k^t \frac{w_k^t}{w_k^t + \theta^t l_k^t} \prod_{k' \neq k} [\frac{\theta^t w_{k'}^t}{l_{k'}^t + \theta^t w_{k'}^t}] \tag{4}$$

In the E-step of $t-$th EM round, compute the expectation $Q(\Theta^t) = E_G \ln p(D, G|\Theta)$

$$\begin{aligned} E_G \ln p(D, G|\Theta) &= \Sigma_d \gamma(d, k, \Theta^t) \ln p(d, g|\Theta) \\ &= \Sigma_d \gamma(d, k, \Theta^t) \{\ln u_k + \Sigma_{w \in W_d, v \in V_d} [\ln \frac{w_k}{w_k + \theta v_k} + \Sigma_{k' \neq k} \ln \frac{\theta w_{k'}}{v_{k'} + \theta w_{k'}}]\} \end{aligned}$$

2.3 M-step

2.3.1 For u

first maximize $Q(\Theta^t)$ with respect to U. For each $u \in U$, eliminating constant terms, we have:

$$\begin{aligned} \min -\Sigma_{u(d)=u} \Sigma_{k=1}^K \gamma(d, k, \Theta^t) \ln u_k \\ w.r.t \Sigma_k u_k = 1 \end{aligned} \tag{5}$$

Solving the above Lagrange function Equ. 5, we get

$$u_k = \frac{\Sigma_{u(d)=u} \gamma(d, k, \Theta^t)}{\Sigma_{s=1}^K \Sigma_{u(d)=u} \gamma(d, s, \Theta^t)} \tag{6}$$

2.3.2 For v

for the :

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

we can derive a lower bound for the log-likelihood over the complete data, given the parameters learnt from previous round. hence, we obtain a minorization function of $\tilde{Q}(\Theta^t)$.

$$\begin{aligned} \tilde{Q}(\Theta^t) = \Sigma_d \gamma(d, k, \Theta^t) \Sigma_{w \in W_d, v \in L_d} \{[\ln w_k + 1 - \ln(w_k^t + \theta^t v_k^t) - \frac{w_k + \theta v_k}{w_k^t + \theta^t v_k^t}] + \\ \Sigma_{k' \neq k} [\ln(\theta w_{k'}) + 1 - \ln(v_{k'}^t + \theta^t w_{k'}^t) - \frac{v_{k'} + \theta w_{k'}}{v_{k'}^t + \theta^t w_{k'}^t}]\} \end{aligned}$$

$\tilde{Q}(\Theta^t)$ can be seperated for each item v. Considering only the $k-$th component v_k, $\tilde{Q}(v_k, \Theta^t)$ involves two terms, one of which is relevant to observations

$d \in W(v)$ where v acts as skyline object, the other is relevant to observations $d \in L(v)$ where v acts as comparisons, $\tilde{Q}(v_k, \Theta^t) = \tilde{Q}^1(v_k, \Theta^t) + \tilde{Q}^2(v_k, \Theta^t)$. Removing all constants and irrelevant terms for v_k, we have the following minorizing function:

$$\tilde{Q}^1(v_k, \Theta^t) = \Sigma_{d \in W(v)} |L_d| \ln v_k - v_k \Sigma_{d \in W(v)} \Sigma_{v' \in L_d} [\frac{\gamma(d, k, \Theta^t)}{\alpha(v, v', k, \Theta^t)} + \Sigma_{k' \neq k} \frac{\theta^t \gamma(d, k, \Theta^t)}{\alpha(v', v, k, \Theta^t)}]$$

$$\tilde{Q}^2(v_k, \Theta^t) = -v_k \Sigma_{d \in L(v)} \Sigma_{v' \in W_d} [\frac{\theta^t \gamma(d, k, \Theta^t)}{\alpha(v', v, k, \Theta^t)} + \Sigma_{k' \neq k} \frac{\gamma(d, k, \Theta^t)}{\alpha(v, v', k, \Theta^t)}]$$

where:

$|L_d|$ is the number of objects being dominanted in d,
$\alpha(v, v', k, \Theta^t) = v_k^t + \theta^t {v'}_k^t$.
By setting the partial derivative of $\frac{\partial \tilde{Q}(v_k, \Theta^t)}{\partial v_k} = 0$, we have:

$$\begin{aligned}\frac{1}{v_k} =& \frac{\Sigma_{d \in W(v)} \Sigma_{v' \in L_d} [\frac{\gamma(d,k,\Theta^t)}{\alpha(v,v',k,\Theta^t)} + \Sigma_{k' \neq k} \frac{\theta^t \gamma(d,k,\Theta^t)}{\alpha(v',v,k,\Theta^t)}]}{\Sigma_{d \in W(v)} |L_d|} \\ &+ \frac{\Sigma_{d \in L(v)} \Sigma_{v' \in W_d} [\frac{\theta^t \gamma(d,k,\Theta^t)}{\alpha(v',v,k,\Theta^t)} + \Sigma_{k' \neq k} \frac{\gamma(d,k,\Theta^t)}{\alpha(v,v',k,\Theta^t)}]}{\Sigma_{d \in W(v)} |L_d|}\end{aligned}$$

More concrete:
if $\gamma(d, k, \Theta^t)$'s k equals to v_k's k:

$$\frac{1}{v_k} = \frac{\Sigma_{d \in W(v)} \Sigma_{v' \in L_d} [\frac{\gamma(d,k,\Theta^t)}{\alpha(v,v',k,\Theta^t)}]}{\Sigma_{d \in W(v)} |L_d|} + \frac{\Sigma_{d \in L(v)} \Sigma_{v' \in W_d} [\frac{\theta^t \gamma(d,k,\Theta^t)}{\alpha(v',v,k,\Theta^t)}]}{\Sigma_{d \in W(v)} |L_d|}$$

else:

$$\frac{1}{v_k} = \frac{\Sigma_{d \in W(v)} \Sigma_{v' \in L_d} [\frac{\theta^t \gamma(d,k,\Theta^t)}{\alpha(v',v,k,\Theta^t)} |_{k'=k}]}{\Sigma_{d \in W(v)} |L_d|} + \frac{\Sigma_{d \in L(v)} \Sigma_{v' \in W_d} [\frac{\gamma(d,k,\Theta^t)}{\alpha(v,v',k,\Theta^t)} |_{k'=k}]}{\Sigma_{d \in W(v)} |L_d|}$$

2.3.3 For θ

Fix u, v, update θ by:

Fix $v \in V$ and $u \in U$, rearranging $Equ.9$, we have the solution for $\frac{\partial \tilde{Q}(\Theta^t)}{\partial \theta} = 0$ as:

$$\theta = \frac{(K-1) \Sigma_d |W_d| |L_d|}{\Sigma_d \Sigma_k \gamma(d, k, \Theta^t) \Sigma_{w,v} [\frac{v_k}{\alpha(w,v,k,\Theta^t)} + \Sigma_{k' \neq k} \frac{w_{k'}}{\alpha(v,w,k',\Theta^t)}]} \tag{7}$$

Note that :
in the $\Sigma_d \Sigma_k \gamma(d, k, \Theta^t)$ only summing these d which $\gamma(d, k, \Theta^t)$'s k equals to d's k.