The BTL Model

Xiaolin Shen

August 5, 2018

1 The Standard BTL Model –

Base on the standard BTL model:

$$p(i \succ j|u) = \frac{ui}{ui + uj}$$
$$p(i \prec j|u) = \frac{uj}{uj + ui}$$

Suppose that there are K underlying aspects. for each pair, $p(< w, v >) = p^k(w \succ v)$. then the probability of generating a session observation d is defined as:

$$p(d|V, U) = \Pi_{w \in W^d, v \in L^d} \Pi_{k=1}^{K} [\frac{u_k w_k}{u_k w_k + u_k v_k}] \tag{1}$$

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

$$\begin{aligned} \gamma(d, k, \Theta^t) &= \frac{p(d, g|\Theta^t)}{\Sigma_g p(d, g|\Theta^t)} = \frac{p(g|\Theta^t) p(d|\Theta^t, g)}{\Sigma_g p(g|\Theta^t) p(d|\Theta^t, g)} \\ &= \frac{u_{g^k} \Pi_{w \in W^d, v \in L^d} \Pi_{k=1}^{K} \frac{u_k w_k}{u_k w_k + u_k v_k}}{\Sigma_{k=1}^{K} u_{g^k} \Pi_{w \in W^d, v \in L^d} \Pi_{k=1}^{K} \frac{u_k w_k}{u_k w_k + u_k v_k}} \end{aligned} \tag{2}$$

Note that $\forall d, \Sigma_k \gamma(d, k, \Theta^t) = 1$, g^k denote u's favorite aspect.

1.1 E-step

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

$$\begin{aligned} E_G \ln p(D, G|\Theta) &= \Sigma_d \Sigma_{k=1}^{K} \gamma(d, k, \Theta^t) \ln p(d, g|\Theta) \\ &= \Sigma_d \Sigma_{k=1}^{K} \gamma(d, k, \Theta^t) \{\ln u_{g^k} \\ &+ \Sigma_{w \in W_d, v \in V_d} [\ln \frac{u_k w_k}{u_k w_k + u_k v_k}]\} \end{aligned} \tag{3}$$

1.2 M-step

1.2.1 For u

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

By setting the partial derivative of $\frac{\partial \tilde{Q}(u_k,\Theta^t)}{\partial u_k} = 0$, we have:

$$\Sigma_d \Sigma_{k=1}^K \Sigma_{w \in W_d, v \in V_d} (\frac{1}{u_k} + \frac{w_k}{u_k w_k} - \frac{v_k + w_k}{u_k w_k + u_k v_k}) = 0 \tag{4}$$

$$u_k = \Sigma_{u(d)=u} \gamma(d, k, \Theta^t) \tag{5}$$

1.2.2 For v

secondly, maximize $Q(\Theta^t)$ with respect to 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 $\tilde{Q}(\Theta^t)$.

$$\tilde{Q}(\Theta^t) = \Sigma_d \Sigma_k \gamma(d, k, \Theta^t) \Sigma_{w \in W_d, v \in L_d} \{[\ln(u_k w_k) + 1 - \ln(u_k^t w_k^t + u_k^t v_k^t) - \frac{u_k w_k + u_k v_k}{u_k^t w_k^t + u_k^t v_k^t}]\} \tag{6}$$

$\tilde{Q}(\Theta^t)$ 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:

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

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

where:

$|L_d|$ is the number of objects being dominanted in d,
$\alpha(u, w, v, k) = u_k^t w_k^t + u_k^t v_k^t$.

By setting the partial derivative of $\frac{\partial \tilde{Q}(v_k,\Theta^t)}{\partial v_k} = 0$, we have:

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