Homework 2

Bohan Niu

October 14, 2018

Problem 1
a)The goal is to find ϕ that:

$$\int p(\mathcal{R}, U, V, \phi) d\phi = p(\mathcal{R}, U, V) \tag{1}$$

By assuming that ϕ_{ij} are independent, we can write:

$$\begin{aligned}
p(\phi|\mathcal{R}, U, V) &= \frac{\prod_{(i,j)\in\Omega} p(r_{ij}|\phi_{ij}, u_i, v_j)p(\phi_{ij}|u_i, v_j)}{\prod_{(i,j)\in\Omega} \int p(r_{ij}|\phi_{ij}, u_i, v_j)p(\phi_{ij}|u_i, v_j)d\phi_{ij}} \\
&= \prod_{(i,j)\in\Omega} \frac{p(r_{ij}|\phi_{ij}, u_i, v_j)p(\phi_{ij}|u_i, v_j)}{\int p(r_{ij}|\phi_{ij}, u_i, v_j)p(\phi_{ij}|u_i, v_j)d\phi_{ij}} \\
&= \prod_{(i,j)\in\Omega} p(\phi_{ij}|r_{ij}, u_i, v_j)
\end{aligned} \tag{2}$$

From EM equation, we can know that:

$$\begin{aligned}
\ln p(\mathcal{R}, U, V) &= \ln p(U, V) + \sum_{(i,j)\in\Omega} \ln p(r_{ij}|u_i, v_j) \\
&= \ln p(U, V) + \sum_{(i,j)\in\Omega} \int q(\phi_{ij}) \ln \frac{p(r_{ij}, \phi_{ij}|u_i, v_j)}{q(\phi_{ij})} d(\phi_{ij}) \\
&+ \sum_{(i,j)\in\Omega} \int q(\phi_{ij}) \ln \frac{q(\phi_{ij})}{p(\phi_{ij}|r_{ij}, u_i, v_j)} d(\phi_{ij})
\end{aligned} \tag{3}$$

In order to make the last fraction in RHS equals to 0, we select q:

$$q(\phi_{ij}) = p(\phi|\mathcal{R}, \mathcal{U}, \mathcal{V}) = \prod_{(i,j)\in\Omega} p(\phi_{ij}|r_{ij}, u_i, v_j) = \prod_{(i,j)\in\Omega} q(\phi_{ij}) \tag{4}$$

In the data, $r_{ij} = \{1, -1\}$, but it is fine for us suppose that $r_{ij} = \{1, 0\}$ when we derive $q(\phi_{ij})$. Thus, we set an indicator for r_{ij}:

$$r_{ij} = \mathbb{1}(\phi_{ij} > 0), \quad \phi_{ij} \sim \text{Normal}(u_i^T v_j, \sigma^2) \tag{5}$$

Then for $r_{ij} = 1$ we can write the joint likelihood:

$$\begin{aligned} p(r_{ij} = 1, \phi_{ij}|u_i, v_j) &= p(r_{ij} = 1|\phi_{ij})p(\phi_{ij}|u_i, v_j) \\ &= \mathbb{1}(\phi_{ij} > 0)(2\pi\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2\sigma^2}(\phi_{ij}-u_i^T v_j)^2} \end{aligned} \tag{6}$$

By integrating equation (6), we can show that it equals to $\Phi(u_i^T v_j/\sigma)$:

$$\begin{aligned} \int p(r_{ij} = 1, \phi_{ij}|u_i, v_j)d\phi_{ij} &= \int_{-\infty}^{\infty} \mathbb{1}(\phi_{ij} > 0)(2\pi\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2\sigma^2}(\phi_{ij}-u_i^T v_j)^2}d\phi_{ij} \\ &= \int_{0}^{\infty} (2\pi\sigma^2)^{-\frac{1}{2}}e^{-\frac{1}{2\sigma^2}(\phi_{ij}-u_i^T v_j)^2}d\phi_{ij} \\ &= P(\phi_{ij} > 0) \end{aligned} \tag{7}$$

Since we now that $\phi_{ij} \sim \text{Normal}(u_i^T v_j, \sigma^2)$, we can write $\phi_{ij} = u_i^T v_j + \sigma s$, where $s \sim \text{Normal}(0, 1)$. Therefore, we turn equation (7) into:

$$\begin{aligned} \int p(r_{ij} = 1, \phi_{ij}|u_i, v_j)d\phi_{ij} &= P(\phi_{ij} > 0) = P(\phi_{ij} = u_i^T v_j + \sigma s > 0) \\ &= P(s \leq -u_i^T v_j/\sigma) = \Phi(u_i^T v_j/\sigma) \end{aligned} \tag{8}$$

After showing this, we have proved that we found the right ϕ_{ij}. So finally, we can derive $q(\phi_{ij})$:

$$\begin{aligned} q(\phi_{ij}) = p(\phi_{ij}|r_{ij}, u_i, v_j) &= \frac{p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i, v_j)}{\int p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i, v_j)d\phi_{ij}} \\ &= \frac{\mathbb{1}(\text{sign}(\phi_{ij}) = r_{ij})e^{\frac{1}{2\sigma^2}(\phi_{ij}-u_i^T v_j)^2}}{\int_{-\infty}^{\infty} \mathbb{1}(\text{sign}(\phi_{ij}) = r_{ij})e^{\frac{1}{2\sigma^2}(\phi_{ij}-u_i^T v_j)^2}d\phi_{ij}} \end{aligned} \tag{9}$$

That is, a truncated normal distribution $TN_1(u_i^T v_j, \sigma^2)$ when $r_i j = 1$ and $TN_0(u_i^T v_j, \sigma^2)$ when $r_i j = -1$.

b)Then we can calculate:

$$\mathcal{L}(U, V) = \ln p(U, V) + \sum_{(i,j)\in\Omega} \mathbb{E}_q[\ln p(r_{ij}, \phi_{ij}|u_i, v_j)] + \text{const.} \tag{10}$$

We suppose that $U$and V are independent, so $\ln p(U, V) = \ln p(U) + \ln p(V) = \sum_i \ln p(u_i) + \sum_j \ln p(v_j)$. Then we plug in $p(r_{ij}, \phi_{ij}|u_i, v_j)$ and get:

$$\begin{aligned} \mathcal{L}(U, V) = &\sum_i \ln p(u_i) + \sum_j \ln p(v_j) + \sum_{(i,j)\in\Omega} \mathbb{E}_q[\ln \mathbb{1}(\text{sign}(\phi_{ij}) = r_{ij})] \\ &- \sum_{(i,j)\in\Omega} \mathbb{E}_q[(\phi_{ij} - u_i^T v_j)^2] + \text{const.} \\ &= -\sum_i \frac{1}{2c}u_i^T u_i - \sum_j \frac{1}{2c}u_j^T u_j - \sum_{(i,j)\in\Omega} \frac{1}{2\sigma^2}(u_i^T v_j v_j^T u_i - 2u_i^T v_j \mathbb{E}_q[\phi_{ij}]) + \text{const.} \end{aligned} \tag{11}$$

and $\mathbb{E}_q[\phi_{ij}]$ is:

$$\mathbb{E}_q[\phi_{ij}] = \begin{cases} u_i^T v_j + \sigma \times \frac{\Phi'(-u_i^T v_j/\sigma)}{1-\Phi(-u_i^T v_j/\sigma)} & \text{if } r_{ij} = 1 \\ u_i^T v_j + \sigma \times \frac{-\Phi'(-u_i^T v_j/\sigma)}{\Phi(-u_i^T v_j/\sigma)} & \text{if } r_{ij} = -1 \end{cases} \tag{12}$$

c)Because U, V are independent, so we can derive then separately. We first derive U:

$$u_i = \arg\max_{u_i} \mathcal{L}(U) \Leftrightarrow u_i = (c^{-1}I + \sum_{(i,j)\in\Omega} v_j v_j^T/\sigma^2)^{-1}(\sum_{(i,j)\in\Omega} v_j \mathbb{E}_q[\phi_{ij}]/\sigma) \tag{13}$$

In the same way, we can derive V:

$$v_j = \arg\max_{v_j} \mathcal{L}(V) \Leftrightarrow v_j = (c^{-1}I + \sum_{(i,j)\in\Omega} u_i u_i^T/\sigma^2)^{-1}(\sum_{(i,j)\in\Omega} u_i \mathbb{E}_q[\phi_{ij}]/\sigma) \tag{14}$$

d) In summary, the steps of the EM algorithm are:

1. Initialize U, V, σ and c in some way.
2. For iteration $t = 1, \dots, T$
   (a) E-Step: Calculate the matrix $\mathbb{E}[\phi]$, where:

$$\mathbb{E}_q[\phi_{ij}] = \begin{cases} u_i^T v_j + \sigma \times \frac{\Phi'(-u_i^T v_j/\sigma)}{1-\Phi(-u_i^T v_j/\sigma)} & \text{if } r_{ij} = 1 \\ u_i^T v_j + \sigma \times \frac{-\Phi'(-u_i^T v_j/\sigma)}{\Phi(-u_i^T v_j/\sigma)} & \text{if } r_{ij} = -1 \end{cases}$$

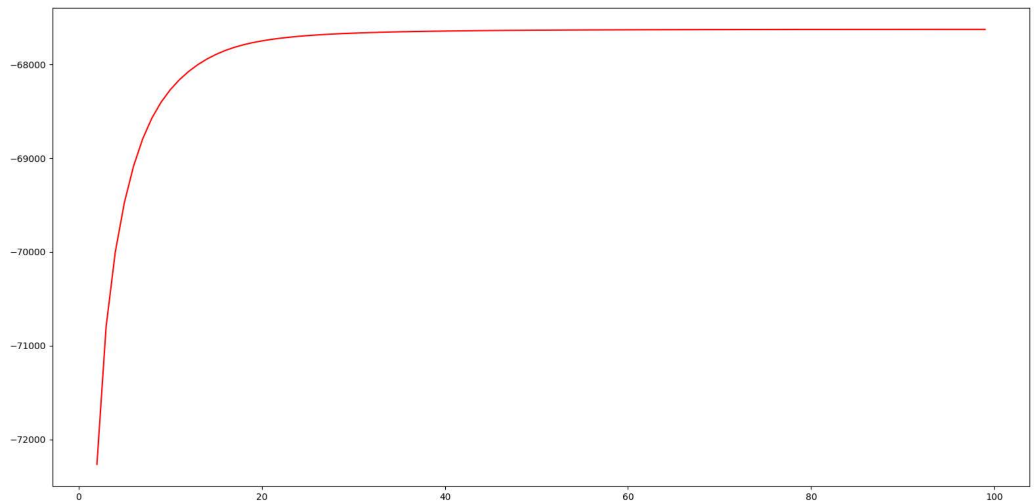
   (b) M-Step: Update the vector u_i and v_j using the expectations above in the following equation

$$\begin{aligned} u_i &= (c^{-1}I + \sum_{(i,j)\in\Omega} v_j v_j^T/\sigma^2)^{-1}(\sum_{(i,j)\in\Omega} v_j \mathbb{E}_q[\phi_{ij}]/\sigma) \\ v_j &= (c^{-1}I + \sum_{(i,j)\in\Omega} u_i u_i^T/\sigma^2)^{-1}(\sum_{(i,j)\in\Omega} u_i \mathbb{E}_q[\phi_{ij}]/\sigma) \end{aligned}$$

   (c)Calculate $\ln p(\mathcal{R}, U, V)$ using the equation

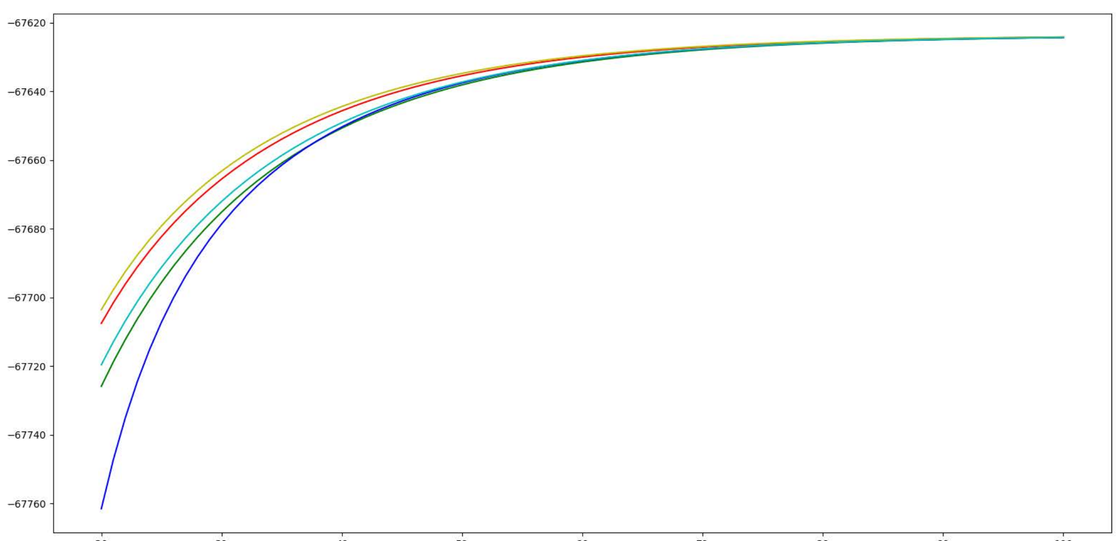
$$\begin{aligned} \ln p(\mathcal{R}, U, V) =& \frac{d}{2}(M+N)\ln(\frac{1}{2c\pi}) - \frac{1}{2c}\sum_i u_i^T u_i - \frac{1}{2c}\sum_j v_j^T v_j \\ &+ \sum_{(i,j)\in\Omega} \frac{1+r_{ij}}{2} \ln \Phi(u_i^T v_j/\sigma) + \sum_{(i,j)\in\Omega} \frac{1-r_{ij}}{2} \ln(1 - \Phi(u_i^T v_j/\sigma)) \end{aligned}$$

Problem 2
a)



−68000
−69000
−70000
−71000
−72000
0
20
40
60
80
100

b)



−67620
−67640
−67660
−67680
−67700
−67720
−67740
−67760
20
30
40
50
60
70
80
90
100

c) confusion matrix

	like	not like
like	2180	554
not like	1115	1151