tq 2137 Tianbo Qiu HW2

Problem 1

a) $q(\phi) = P(\phi \mid R, U, V) = \dfrac{P(R \mid \phi, U, V)\, P(\phi \mid U, V)}{\int P(R \mid \phi, U, V)\, P(\phi \mid U, V)\, d\phi} = \dfrac{\prod_{(i,j)\in\Omega} P(r_{ij} \mid \phi_{ij})\, P(\phi_{ij} \mid u_i, v_j)}{\int\int\int P(r_{ij} \mid \phi_{ij})\, P(\phi_{ij} \mid u_i, v_j)\, d\phi_{ij} d\ldots}$

$= \prod_{(i,j)\in\Omega} P(\phi_{ij} \mid r_{ij}, u_i, v_j)$

$P(\phi_{ij} \mid r_{ij}, u_i, v_j) = \dfrac{P(r_{ij} \mid \phi_{ij})\, P(\phi_{ij} \mid u_i, v_j)}{\int P(r_{ij} \mid \phi_{ij})\, P(\phi_{ij} \mid u_i, v_j)\, d\phi_{ij}}$

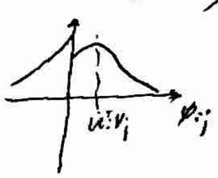
$P(\phi_{ij} \mid u_i, v_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(\phi_{ij} - u_i^T v_j)^2\}$

$P(r_{ij} \mid \phi_{ij}) = \mathbb{1}\{\text{sign}(\phi_{ij})\}$

$= \dfrac{\mathbb{1}\{\text{sign}(\phi_{ij})\}\, P(\phi_{ij} \mid u_i, v_j)}{\int \mathbb{1}\{\text{sign}(\phi_{ij})\}\, P(\phi_{ij} \mid u_i, v_j)\, d\phi_{ij}}$

Thus. $P(\phi_{ij} \mid r_{ij}, u_i, v_j)$ is the truncated normal distribution $TN(u_i^T v_j, \sigma^2)$

$q(\phi) = \prod_{(i,j)\in\Omega} q(\phi_{ij}) = \prod_{(i,j)\in\Omega} \dfrac{\mathbb{1}\{\text{sign}(\phi_{ij})\} \exp\{-\frac{1}{2\sigma^2}(\phi_{ij} - u_i^T v_j)^2\}}{\int \mathbb{1}\{\text{sign}(\phi_{ij})\} \exp\{-\frac{1}{2\sigma^2}(\phi_{ij} - u_i^T v_j)^2\}\, d\phi_{ij}}$



b) $\mathcal{L}(U, V) = \int q(\phi) \ln \dfrac{P(R, U, V, \phi)}{q(\phi)}\, d\phi$

$P(R, U, V, \phi) = P(U, V)\, P(R, \phi \mid U, V) = P(U, V)\, P(R \mid \phi, U, V)\, P(\phi \mid U, V)$

$= P(U)\, P(V)\, P(R \mid \phi)\, P(\phi \mid U, V) = \prod_{i=1}^{N} P(u_i) \prod_{j=1}^{M} P(v_j) \prod_{(i,j)\in\Omega} P(r_{ij} \mid \phi_{ij}) \prod_{(i,j)\in\Omega} P(\phi_{ij} \mid u_i, v_j)$

$\mathcal{L}(U, V) = E_{q(\phi)} \ln P(R, U, V, \phi) - E_{q(\phi)} \ln q(\phi)$

$= E_{q(\phi)} \Big[\sum_{i=1}^{N} \ln p(u_i) + \sum_{j=1}^{M} \ln p(v_j) + \underbrace{\sum_{(i,j)\in\Omega} \ln p(r_{ij} \mid \phi_{ij})}_{0} + \sum_{(i,j)\in\Omega} p(\phi_{ij} \mid u_i, v_j) \Big) + \text{constant}$

$= -\frac{1}{2c} \sum_{i=1}^{N} u_i^T u_i - \frac{1}{2c} \sum_{j=1}^{M} v_j^T v_j - \sum_{(i,j)\in\Omega} \frac{1}{2\sigma^2} (v_j^T u_i u_i^T v_j - 2 v_j^T u_i E_q[\phi_{ij}]) + \text{constant}.$

$$E_q(\phi_{ij}) = \begin{cases} u_i^T v_j + \sigma \times \dfrac{\Phi'(-u_i^T v_j/\sigma)}{1 - \Phi(-u_i^T v_j/\sigma)}, & \text{if } r_{ij} = 1 \\[2ex] u_i^T v_j + \sigma \times \dfrac{-\Phi'(-u_i^T v_j/\sigma)}{\Phi(-u_i^T v_j/\sigma)}, & \text{if } r_{ij} = -1 \end{cases}$$

(c) $$\frac{\partial \mathcal{L}(U,V)}{\partial V_j} = -\frac{1}{c}V_j - \sum_{(i,j)\in\Omega} \frac{1}{2\sigma^2}\left(2u_i u_i^T V_j - 2u_i \mathbb{E}_q[\phi_{ij}]\right) = 0$$

$$\Rightarrow \quad -\frac{1}{c}V_j - \frac{1}{\sigma^2}\Big(\sum_{(i,j)\in\Omega} u_i u_i^T\Big)V_j + \frac{1}{\sigma^2}\sum_{(i,j)\in\Omega}\left(u_i \mathbb{E}_q[\phi_{ij}]\right)$$

$$\Rightarrow \quad \Big(\frac{1}{c}I + \frac{1}{\sigma^2}\big(\sum_{(i,j)\in\Omega} u_i u_i^T\big)\Big)V_j = \frac{1}{\sigma^2}\sum_{(i,j)\in\Omega}\left(u_i \mathbb{E}_q[\phi_{ij}]\right)$$

$$\Rightarrow \quad V_j = \Big(\frac{1}{c}I + \frac{1}{\sigma^2}\sum_{(i,j)\in\Omega} u_i u_i^T\Big)^{-1}\Big(\frac{1}{\sigma^2}\sum_{(i,j)\in\Omega}(u_i \mathbb{E}_q[\phi_{ij}]\Big)$$

$$\frac{\partial \mathcal{L}(U,V)}{\partial u_i} = 0 \quad \Rightarrow \quad u_i = \Big(\frac{1}{c}I + \frac{1}{\sigma^2}\sum_{(i,j)\in\Omega} V_j V_j^T\Big)^{-1}\Big(\frac{1}{\sigma^2}\sum_{(i,j)\in\Omega}(V_j \mathbb{E}_q[\phi_{ij}]\Big)$$

d) EM algorithm :

1. Initialize $u_i \in \mathbb{R}^d$, $i=1\cdots N$, $V_j \in \mathbb{R}^d$, $j=1\cdots M$, and set σ in some way.

2. For iteration $t=1,\cdots T$

ⓐ E-step : Caculus
$$\mathbb{E}_{(t)}^{(i,j)\in\Omega}(\phi_{ij}) = \begin{cases} u_{i(t-1)}^T V_{j(t-1)} + \sigma \times \dfrac{\Phi'(-u_{i(t-1)}^T V_{j(t-1)}/\sigma)}{1-\Phi(-u_{i(t-1)}^T V_{j(t-1)}/\sigma)}, & \text{if } r_{ij}=1 \\ u_{i(t-1)}^T V_{j(t-1)} + \sigma \times \dfrac{-\Phi'(-u_{i(t-1)}^T V_{j(t-1)}/\sigma)}{\Phi(-u_{i(t-1)}^T V_{j(t-1)}/\sigma)}, & \text{if } r_{ij}=-1 \end{cases}$$

ⓑ M-step : update u_i, V_j using the expectations above in the following equation :

$$u_i^{(t)} = \Big(\frac{1}{c}I + V_{i(t-1)}V_{j(t-1)}^T/\sigma^2\Big)^{-1}\Big(\frac{1}{\sigma^2}V_{j(t-1)}\mathbb{E}_{q(t)}[\phi_{ij}]\Big)$$

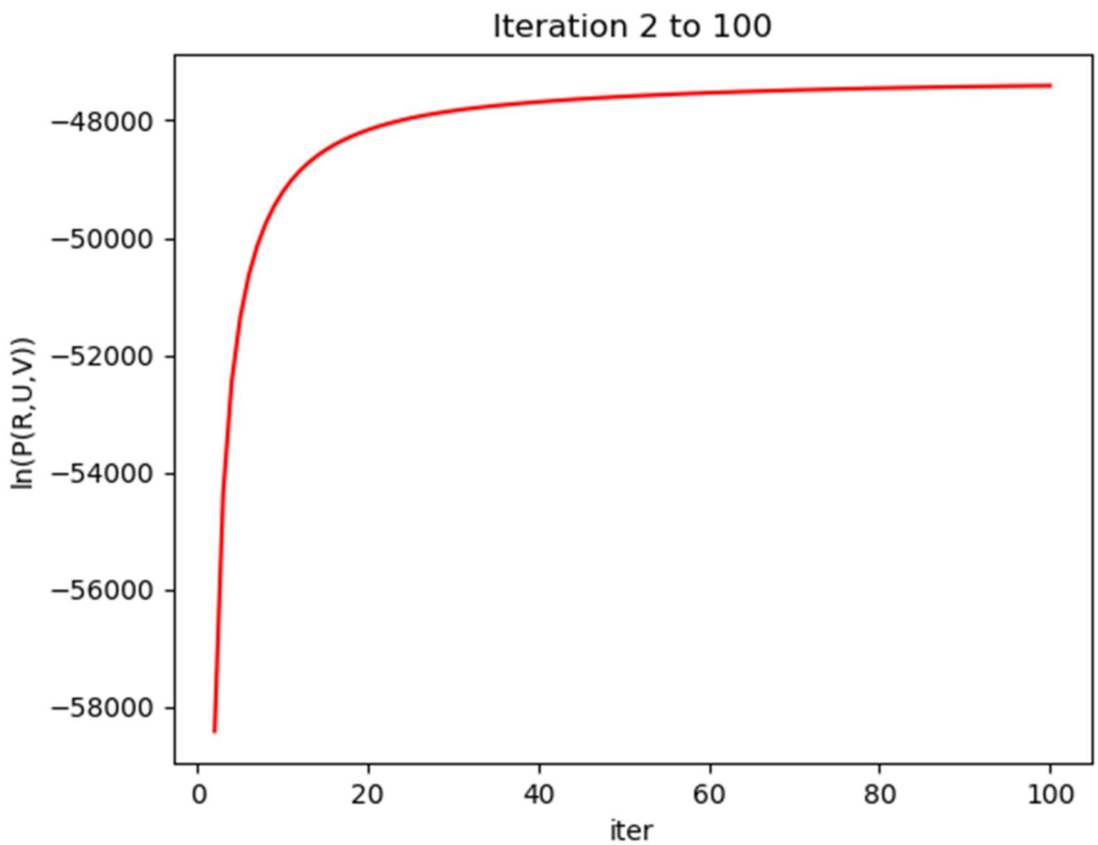
$$V_j^{(t)} = \Big(\frac{1}{c}I + u_{i(t)}u_{i(t-1)}^T/\sigma^2\Big)^{-1}\Big(\frac{1}{\sigma^2}u_{i(t)}\mathbb{E}_{q(t)}[\phi_{ij}]\Big)$$

ⓒ caculate $\ln P(R,U,V)$ using equation

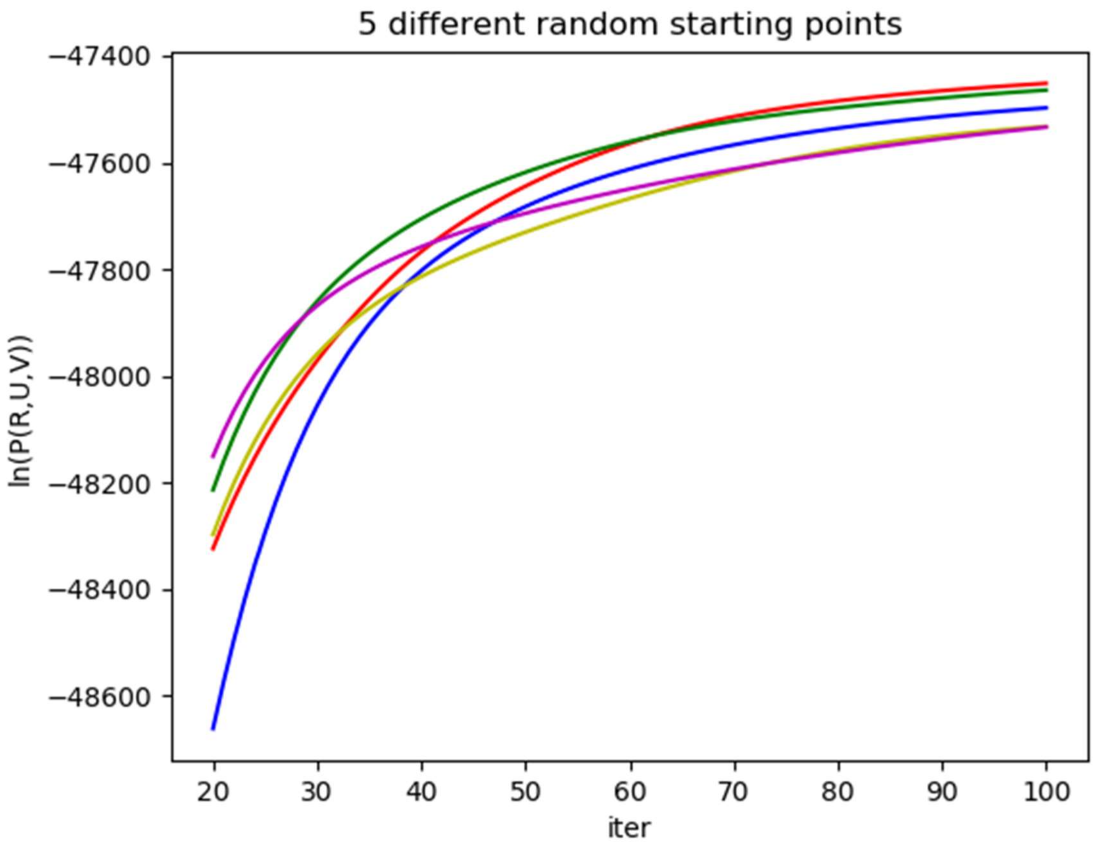
$$\ln P(R,U,V) = -\frac{1}{2c}\sum_{i=1}^{N} u_i^T u_i - \frac{N}{2}\ln((2\pi)^d c) - \frac{1}{2c}\sum_{j=1}^{M} V_j^T V_j - \frac{M}{2}\ln((2\pi)^d c)$$
$$+ \sum_{(i,j)\in\Omega}\begin{cases} \ln \Phi_{(t)}(u_{i(t)}^T V_{j(t)}/\sigma), & \text{if } r_{ij}=1 \\ \ln(1-\Phi(u_{i(t)}^T V_{j(t)}/\sigma)), & \text{if } r_{ij}=-1 \end{cases}$$

Problem 2

(a) Plot_a.py EM_training.py



(b) plot_b.py EM_training.py



(c) EM_rating_test.py

```
C:\Users\tianb\Google Drive\18Fall\BAYES6720\HW2\movies_csv>python EM_rating_v0.3.py
-----Confusion Matrix-----
actual          likes   dislikes
predict
likes           2146    813
dislikes        588     1453
```

-----Confusion Matrix-----

actual   likes   dislikes

predict

likes   2146   813

dislikes   588   1453

We see that 2146 likes are predicted right out of all 2734 true likes, and 1453 dislikes are predicted right out of all 2266 true dislikes.