Homework 2 SOLUTIONS

Problem 1.

a)

$$\begin{aligned} p(\phi|U,V,R) &\propto p(R|\phi,U,V)p(\phi|U,V) & (1)\\ &\propto \prod_{(i,j)\in\Omega} p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i,v_j) & (2)\\ &= \prod_{(i,j)\in\Omega} \frac{p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i,v_j)}{\int_{-\infty}^{\infty} p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i,v_j)d\phi_{ij}} & (3)\\ &= \prod_{(i,j)\in\Omega} p(\phi_{ij}|r_{ij},u_i,v_j) & (4) \end{aligned}$$

$$\begin{aligned} p(\phi_{ij}|r_{ij},u_i,v_j) &\propto p(r_{ij}|\phi_{ij})p(\phi_{ij}|u_i,v_j) & (5)\\ &\propto \mathbb{1}(r_{ij} = \mathrm{sign}(\phi_{ij}))N(\phi_{ij}|u_i^T v_j,\sigma^2) & (6)\\ &= \frac{\mathbb{1}(r_{ij} = \mathrm{sign}(\phi_{ij}))N(\phi_{ij}|u_i^T v_j,\sigma^2)}{\int_{-\infty}^{\infty} \mathbb{1}(r_{ij} = \mathrm{sign}(\phi_{ij}))N(\phi_{ij}|u_i^T v_j,\sigma^2)d\phi_{ij}} & (7)\\ &= TN_{r_{ij}}(u_i^T v_j,\sigma^2) & (8) \end{aligned}$$

TN is the truncated normal on $\mathbb{R}_{r_{ij}}$.

b)

$$\begin{aligned} \mathcal{L}(U,V) &= \mathbb{E}[\ln p(R,U,V,\phi)] + \text{const.} & (9)\\ &= \sum_{(i,j)\in\Omega} \mathbb{E}[\ln p(r_{ij}|\phi_{ij})] + \mathbb{E}[\ln p(\phi_{ij}|u_i,v_j) + \sum_{i=1}^{N}\ln p(u_i) + \sum_{j=1}^{M}\ln p(v_j) + \text{const.}\\ &= \sum_{(i,j)\in\Omega} -\frac{1}{2\sigma^2}\mathbb{E}[(\phi_{ij}-u_i^T v_j)^2] - \sum_{i=1}^{N}\frac{1}{2c}u_i^T u_i - \sum_{j=1}^{M}\frac{1}{2c}v_j^T v_j + \text{const.}\\ &= \sum_{(i,j)\in\Omega} -\frac{1}{2\sigma^2}(\mathbb{E}[\phi_{ij}]-u_i^T v_j)^2 - \sum_{i=1}^{N}\frac{1}{2c}u_i^T u_i - \sum_{j=1}^{M}\frac{1}{2c}v_j^T v_j + \text{const.} \end{aligned}$$

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

c)

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

$$\Downarrow$$

$$u_i = \left(c^{-1}I + \sigma^{-2}\sum_{j:(i,j)\in\Omega} v_j v_j^T\right)^{-1}\left(\sigma^{-2}\sum_{j:(i,j)\in\Omega}\mathbb{E}[\phi_{ij}]v_j\right)$$

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

$$\Downarrow$$

$$v_j = \left(c^{-1}I + \sigma^{-2}\sum_{i:(i,j)\in\Omega} u_i u_i^T\right)^{-1}\left(\sigma^{-2}\sum_{i:(i,j)\in\Omega}\mathbb{E}[\phi_{ij}]u_i\right)$$

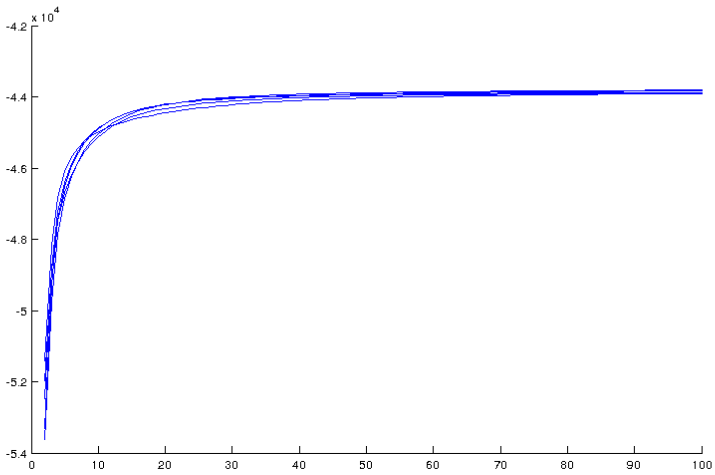
d)
- Initialize U^0 and V^0 in some way.
- For iteration $t = 1, \ldots, T$
- E-Step: Update $q_t(\phi_{ij})$ as above.
- M-Step:
  - For $i = 1, \ldots, N$: Set $u_i^t = \left(c^{-1}I + \sigma^{-2}\sum_{j:(i,j)\in\Omega} v_j^{t-1}v_j^{(t-1)T}\right)^{-1}\left(\sigma^{-2}\sum_{j:(i,j)\in\Omega}\mathbb{E}_{q_t}[\phi_{ij}]v_j^{t-1}\right)$
  - For $j = 1, \ldots, M$: Set $v_j^t = \left(c^{-1}I + \sigma^{-2}\sum_{i:(i,j)\in\Omega} u_i^t u_i^{tT}\right)^{-1}\left(\sigma^{-2}\sum_{i:(i,j)\in\Omega}\mathbb{E}_{q_t}[\phi_{ij}]u_i^t\right)$
- Calculate

$$\mathcal{L}_t = \ln p(R, U^t, V^t)$$

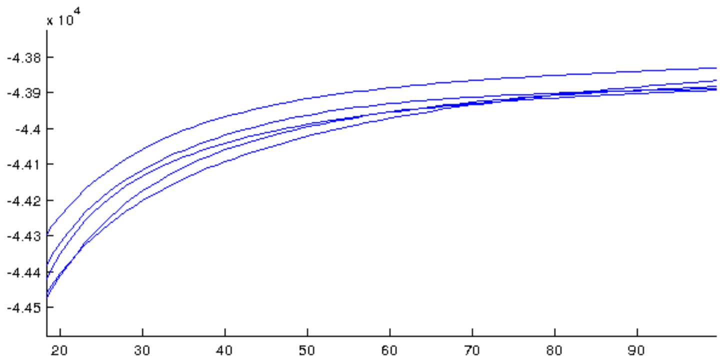
$$= \sum_i \ln p(u_i^t) + \sum_j \ln p(v_j^t) + \sum_{(i,j)\in\Omega} \mathbb{1}(r_{ij}=1)\ln\Phi(u_i^t v_j^{tT}/\sigma) + \mathbb{1}(r_{ij}=-1)\ln(1-\Phi(u_i^t v_j^{tT}/\sigma))$$

Problem 2.

a) Here are 5 to show a few. It must be monotonically increasing.



b) As with Problem 2a, the plot needs to be monotonically increasing.



c) The algorithm likely won't converge to the same results for different initializations. Below I have copied results for 5 different runs to show the range of values. The organization is

$$M = \begin{matrix} \text{correct 1} & \text{truth = 1, predict = -1} \\ \text{truth = -1, predict = 1} & \text{correct -1} \end{matrix}$$

$$M = \begin{matrix} 2124 & 610 \\ 827 & 1439 \end{matrix} \qquad M = \begin{matrix} 2142 & 592 \\ 838 & 1428 \end{matrix} \qquad M = \begin{matrix} 2122 & 612 \\ 829 & 1437 \end{matrix}$$

$$M = \begin{matrix} 2140 & 594 \\ 851 & 1415 \end{matrix} \qquad M = \begin{matrix} 2143 & 591 \\ 825 & 1441 \end{matrix}$$