Model explanation

We have a $n \times m$ matrix X. We want to decompose X into low-rank matrices U and V. Their shapes are $n \times k$ and $k \times m$, where $k <<< m, n$.

The priors of u and v are

$$u_i \sim N(0, \lambda^{-1} I), \quad .i = 1, \cdots, n$$

,

$$v_j \sim N(0, \lambda^{-1}), \quad j = 1, \cdots, m$$

The distribution of the elements of X is,

$$X_{ij} \sim N(u_i^T v_j, \sigma^2), \quad \text{for } (i, j) \in \Omega$$

Ω is the set containing the pairs (i, j) such that M_{ij} is measure i.e. user i has rated object j. Similarly Ω_{u_i} corresponds to the objects rated by user u, and Ω_{v_j} corresponds to the users who have rated object j.

The joint likelihood is given as ,

$$p(X, U, V) = \prod_{(i,j) \in \Omega} p(X_{ij} | u_i, v_j) \times \prod_{i=1}^{n} p(u_i) \times \prod_{j=1}^{m} p(v_j)$$

$$\log p(X, U, V) = \sum_{(i,j) \in \Omega} \log p(X_{ij} | u_i, v_j) + \sum_{i=1}^{n} \log p(u_i) + \sum_{j=1}^{m} \log p(v_j)$$

The MAP is computed by maximizing the function:

$$\mathcal{L} = \sum_{(i,j) \in \Omega} \frac{1}{2\sigma^2} ||M_{ij} - u_i^T v_j||^2 - \sum_{i=1}^{n} \frac{\lambda}{2} ||u_i||^2 - \sum_{j=1}^{m} \frac{\lambda}{2} ||v_j||^2$$

By taking the derivative with respect to u_i and v_j and setting them to 0, we get

$$u_i = \left(\lambda \sigma^2 I + \sum_{j \in \Omega_{u_i}} v_j v_j^T \right)^{-1} \left(\sum_{j \in \Omega_{u_i}} X_{ij} v_j \right)$$

$$v_j = \left(\lambda \sigma^2 I + \sum_{i \in \Omega_{v_j}} u_i u_i^T \right)^{-1} \left(\sum_{i \in \Omega_{v_j}} X_{ij} u_i \right)$$

We then use co-ordinate ascent algorithm to find the solution.