Derivation of EM algorithm

The complete log-likelihood including missing data $\{z_i\}$ for the proposed model is

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} I(z_{i,j} = k) \left(\sum_{l=1}^{L} \log f_{k,l,x_{i,j,l}} + \log q_{i,k} \right).$$

Here, we introduce the variable for conditional probability for $z_{i,j}$ given the parameters and the mutation features $\boldsymbol{x}_{i,j}$,

$$\theta_{i,k,m} = \Pr(z_{i,j} = k | x_{i,j} = m, \{f_{k,l}\}, \{q_i\})$$

Note that this conditional probability just depends on the value of mutation feature $\mathbf{m} = (m_1, \dots, m_L)$, not on the index j. Then, the expected complete log-likelihood augmented by Lagrange multipliers is calculated as

$$\sum_{i=1}^{I} \sum_{\boldsymbol{m}} g_{i,\boldsymbol{m}} \sum_{k=1}^{K} \theta_{i,k,\boldsymbol{m}} \left(\sum_{l=1}^{L} \log f_{k,l,m_{l}} + \log q_{i,k} \right) + \sum_{k=1}^{K} \sum_{l=1}^{L} \tau_{k,l} \left(1 - \sum_{p=1}^{M_{l}} f_{k,l,p} \right) + \sum_{i=1}^{I} \rho_{i} \left(1 - \sum_{k=1}^{K} q_{i,k} \right).$$

Differentiating it leads to following stationary equations:

$$\sum_{i=1}^{I} \sum_{\mathbf{m}: m_l = p} g_{i,\mathbf{m}} \theta_{i,k,\mathbf{m}} - \tau_{k,l} f_{k,l,p} = 0, \ (p = 1, \dots, M_l, k = 1, \dots, K, \ l = 1, \dots, L).,$$

$$\sum_{\mathbf{m}} g_{i,\mathbf{m}} \theta_{i,k,\mathbf{m}} - \rho_i q_{i,k} = 0, \ (k = 1, \dots, K, i = 1, \dots, I).$$

Then, by eliminating Lagrange multipliers, updating rules can be obtained.