Bonus Homework 5

3170102587 蒋仕彪

Problem description

In **Probabilistic Latent Semantic Analysis**, we want to calculate parameters θ, π so that

$$\text{maximum } l(\theta, \pi; N) = \sum_{d,w} n(d, w) \log(\sum_{z=1}^{K} P(w|z; \theta)P(z|d; \pi))$$

Apply **Expectation Maximization** to estimate θ, π.

Solution

The **EM** method is an iterative method and can be divided into 2 steps.

Actually we want to estimate two matrices $[N, K]$ and $[K, M]$, denote them as $P(z_k|d_i)$ and $P(w_j|z_k)$. At very beginning, assume we have known these parameters (use random value).

E-step: Use the current results to calculate the posterior probability of latent variables.

$$P(z_k|d_i, w_j) = \frac{P(w_j|z_k) P(z_k|d_i)}{\sum_{l=1}^{K} P(w_j|z_l) P(z_l|d_i)}$$

M-step:

1. We use the posterior in the previous step to "modify" the expectation:

$$E(d_i, w_j, z_k) = \sum_{i=1}^{N} \sum_{j=1}^{M} n(d_i, w_j) \sum_{k=1}^{K} P(z_k|d_i, w_j) \log[P(w_j|z_k) P(z_k|d_i)]$$
$$\text{s.t.} \sum_{j=1}^{M} p(w_j|z_k) = 1, \sum_{k=1}^{K} p(z_k|d_i) = 1$$

2. This function has $N \times K + K \times M$ variables. Use Lagrange Multiplier Approach:

$$\mathcal{H} = E(d_i, w_j, z_k) + \sum_{k=1}^{K} \tau_k \left(1 - \sum_{j=1}^{M} P(w_j|z_k)\right) + \sum_{i=1}^{N} \rho_i \left(1 - \sum_{k=1}^{K} P(z_k|d_i)\right)$$

3. Derivative them, get the equations:

$$\sum_{i=1}^{N} n(d_i, w_j) P(z_k|d_i, w_j) - \alpha_k P(w_j|z_k) = 0$$
$$\sum_{j=1}^{M} n(d_i, w_j) P(z_k|d_i, w_j) - \beta_i P(z_k|d_i) = 0$$

4. Finally we can acquire the new value of parameters:

$$P(w_j|z_k) = \frac{\sum_{i=1}^{N} n(d_i,w_j)P(z_k|d_i,w_j)}{\sum_{m=1}^{M} \sum_{i=1}^{N} n(d_i,w_m)P(z_k|d_i,w_m)}$$
$$P(z_k|d_i) = \frac{\sum_{j=1}^{M} n(d_i,w_j)P(z_k|d_i,w_j)}{n(d_i)}$$

Repeat the above 2 steps until results converge.