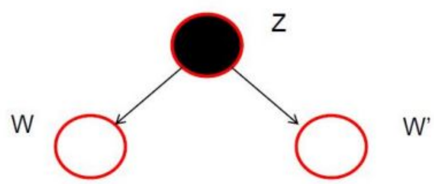
Homework 5

Zhong Yun 2016K8009915009

November 3, 2018

1 Answer Sheet

Theorem 1 对于一个概率图（如图*1*所示），我们可以使用*EM*算法，求解此概率图模型的参数，并达成以下两点目标：
（*1*）写出目标函数，目标函数的下界。
（*2*）*EM*算法推导出的迭代式子。



Proof 1 如图一所示，其结构是由Z引出两个变量ω, ω'。设ω和ω'的分量总数分别为N和M，设Z的分量总数为T。

1. 我们需要明确的目标是：对于图中所示的变量ω, ω'，我们要做的是最大化“所有数据的似然”。即要求如下式子最大化：

$$\prod_{i=1}^{N}\prod_{j=1}^{M} P(\omega_i, \omega'_j)^{n(\omega_i, \omega'_j)}$$

其中，$P(\omega_i, \omega'_j)$可以写成：

$$\begin{aligned} P(\omega_i, \omega'_j) &= \sum_{k=1}^{T} P((\omega_i, \omega'_j) \mid z_k)P(z_k) \\ &= \sum_{k=1}^{T} P(\omega_i \mid z_k)P(\omega'_j \mid z_k)P(z_k) \end{aligned}$$

由于对数函数不改变函数的单调性，所以我们在原函数外加上对数函数是没有问题的，于是问题转化为求$ln\prod_{i=1}^{N}\prod_{j=1}^{M} P(\omega_i, \omega'_j)^{n(\omega_i, \omega'_j)}$ 的最大值。

$$\begin{aligned} ln\prod_{i=1}^{N}\prod_{j=1}^{M} P(\omega_i, \omega'_j)^{n(\omega_i, \omega'_j)} &= \sum_{i=1}^{N}\sum_{j=1}^{M} n(\omega_i, \omega'_j) lnP(\omega_i, \omega'_j) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{M} n(\omega_i, \omega'_j) ln\sum_{k=1}^{T} P(\omega_i \mid z_k)P(\omega'_j \mid z_k)P(z_k) \end{aligned}$$

2. 设目标函数为L，则：

$$L=\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')ln\sum_{k=1}^{T}P(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)$$

以及下面三个限制条件成立：

$$\sum_{i=1}^{N}LP(\omega_i\mid z_k)=1$$

$$\sum_{j=1}^{M}P(\omega_j'\mid z_k)=1$$

$$\sum_{k=1}^{T}P(z_k)=1$$

3. 下面我们使用琴生不等式来估计目标函数的下界。
首先引入一组参量$Q(z_k)$满足$\sum_{k=1}^{T}Q(z_k)=1,Q(z_k)\geq 0$

$$\begin{aligned}L&=\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')ln\sum_{k=1}^{T}P(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)\\&=\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')ln\sum_{k=1}^{T}Q(z_k)\frac{P(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)}{Q(z_k)}\\&\geq\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')\sum_{k=1}^{T}Q(z_k)ln\frac{P(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)}{Q(z_k)}\\&=\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')\sum_{k=1}^{T}Q(z_k)lnP(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)-constant\end{aligned}$$

令目标函数的下界为S，即$S=\sum_{i=1}^{N}\sum_{j=1}^{M}n(\omega_i,\omega_j')\sum_{k=1}^{T}Q(z_k)lnP(\omega_i\mid z_k)P(\omega_j'\mid z_k)P(z_k)$

4. 下面，用拉格朗日乘数法优化目标函数的下界。

$$S*=S+\sum_{k=1}^{T}\alpha_k(1-\sum_{i=1}^{N}LP(\omega_i\mid z_k))+\sum_{k=1}^{T}\beta_k(1-\sum_{j=1}^{M}P(\omega_j'\mid z_k))+\gamma(1-\sum_{k=1}^{T}P(z_k))$$

对自变量$P(\omega_i\mid z_k)$，$P(\omega_j'\mid z_k)$，$P(z_k)$求导，令偏导等于0，如下所示：

$$\sum_{j=1}^{M}\frac{n(\omega_i,\omega_j')Q(z_k)}{P(\omega_i\mid z_k)}-\alpha_k=0$$

$$\sum_{i=1}^{N}\frac{n(\omega_i,\omega_j')Q(z_k)}{P(\omega_j\mid z_k)}-\beta_k=0$$

$$\sum_{i,j=1}^{N,M}\frac{n(\omega_i,\omega_j')Q(z_k)}{P(z_k)}-\gamma=0$$

求解上面三个方程：

$$P(\omega_i\mid z_k)=\frac{\sum_{j=1}^{M}n(\omega_i,\omega_j')Q(z_k)}{\alpha_k}$$

$$P(\omega_j \mid z_k) = \frac{\sum_{i=1}^{N} n(\omega_i, \omega'_j) Q(z_k)}{\beta_k}$$

$$P(z_k) = \frac{\sum_{i,j=1}^{N,M} n(\omega_i, \omega'_j) Q(z_k)}{\gamma}$$

5. 下面我们算$Q(z_k)$。

$$\begin{aligned}
Q(z_k) &= P(z_k \mid (\omega_i, \omega'_j)) \\
&= \frac{P(z_k, \omega_i, \omega'_j)}{P(\omega_i, \omega'_j)} \\
&= \frac{P(z_k) P(\omega_i \mid z_k) P(\omega_j \mid z_k)}{\sum_{k=1}^{T} P(z_k) P(\omega_i \mid z_k) P(\omega_j \mid z_k)}
\end{aligned}$$

综合上面五个部分，我们利用EM算法，写出了目标函数，目标函数的下界以及由算法推导出的迭代式子。