Homework 12

李云志 2015K8009929014

2017 年 11 月 22 日

问题描述：y是离散型随机变量，其可能的取值为$c_k,(k=1,2,...,K)$。x是由M个离散型随机变量组成的随机向量，$x=(x^1,x^2,...,x^M)$,其中x^j的可能取值为$a_{jl},(l=1,2,...,L_j)$。
同时存在归一化条件如下：

$$\sum_{k=1}^{K} P(y=c_k)=1$$

$$\sum_{l=1}^{L_j} P(x^j=a_{jl}|y=c_k)=1$$

目标函数如下：

$$-\sum_{i=1}^{n} \log P(y=y_i|x=x_i)$$

我们的目标是找到$P(y=c_k),P(x^j=a_{jl}|y=c_k)$ 最佳赋值，使得该组赋值既满足归一化条件又能使得我们的目标函数最小。

Solution:
原目标等于最大化函数：

$$\begin{aligned}&\sum_{i=1}^{n}\log P(y=y_i|x=x_i)\\=&\sum_{i=1}^{n}\sum_{j=1}^{M}\log P(x_i^j=X_i^j|y=y_i)+\sum_{i=1}^{n}\log P(y=y_i)\end{aligned}$$

对于目标函数第一部分，设训练集中$y=c_k,x^j=a_{jl}$的数据出现了t_{jkl}次，我们记$P(x^j=a_{jl}|y=c_k)$为p_{jkl},则我们的目标函数第一部分可以被转化为：

$$\sum_{k=1}^{K}\sum_{j=1}^{M}\sum_{l=1}^{L_j} t_{jkl}\log p_{jkl}$$

此部分需要满足的约束为：

$$\sum_{l=1}^{L_j} p_{jkl}-1=0$$

设拉格朗日函数为：

$$Lag=\sum_{k=1}^{K}\sum_{j=1}^{M}\sum_{l=1}^{L_j} t_{jkl}\log p_{jkl}-\sum_{k=1}^{K}\sum_{j=1}^{M}\lambda_{jk}(\sum_{l=1}^{L_j} p_{jkl}-1)$$

对上式分别求偏导得到:

$$Lag'_{p_{jkl}} = \frac{t_{jkl}}{p_{jkl}} - \lambda_{jk} = 0$$
$$Lag'_{\lambda_{jk}} = \sum_{l=1}^{L_j} p_{jkl} - 1 = 0$$

解得：

$$p_{jkl} = \frac{t_{jkl}}{\sum_{l=1}^{L_j} t_{jkl}}$$

同理，对于目标函数的第二部分，我们采用同样的方法，设训练集中$y = c_k$的数据出现了t_k次，则我们可以通过约束条件得到拉格朗日函数：

$$Lag = \sum_{k=1}^{K} t_k \log P(y = c_k) - \lambda(\sum_{k=1}^{K} P(y = c_k) - 1)$$

对上式分别求偏导得到:

$$Lag'_{P(y=c_k)} = \frac{t_k}{P(y=c_k)} - \lambda = 0$$
$$Lag'_{\lambda} = \sum_{k=1}^{K} P(y = c_k) - 1 = 0$$

解得:

$$P(y = c_k) = \frac{t_k}{\sum_{k=1}^{K} t_k}$$

综上所述，我们的概率赋值只需要通过数据集计算出频率来即可。