Machine Learning 1 Homework 07 Theory Part

15.12.2015

The object function is:

$$Q\left(\theta,\theta^{old}\right) = \sum_{z\in\{heads,tails\}^N} P(Z=z|\mathcal{X}=x,\theta^{old})\log P(\mathcal{X}=x,Z=z|\theta)$$

$$= \sum_{z\in\{heads,tails\}^N} P(Z=z|\mathcal{X}=x,\theta^{old})\sum_{i=1}^{N}\left(\log P(Z=z^{(i)}|\theta)+\sum_{j=1}^{m}\log P(X_j=x_j^{(i)}|Z=z^{(i)},\theta)\right)$$

We also know that:

$$\sum_{i=1}^{N}\left(\log P(Z=z^{(i)}|\theta)+\sum_{j=1}^{m}\log P(X_j=x_j^{(i)}|Z=z^{(i)},\theta)\right)$$
$$= \sum_{zishead}[\log\lambda + h(x_i)\log p_1 + t(x_i)\log(1-p_1)] + \sum_{zishead}[\log(1-\lambda)+h(x_i)\log p_2 + t(x_i)\log(1-p_2)]$$

$h(x_i)$ is the number of heads when tossing coin M times, $t(x_i)$ is the number when tossing hcoin M times. Of course that $h(x_i)+t(x_i)=M$

maximize this, we get:

$$\hat{\lambda} = \frac{1}{N}\sum_{i=1}^{N} R_1(i)$$
$$\hat{p_1}^{new} = \frac{\sum_{i=1}^{N} R_1(i)h(x_i)}{M\sum_{i=1}^{N} R_1(i)}$$
$$\hat{p_2}^{new} = \frac{\sum_{i=1}^{N} R_2(i)h(x_i)}{M\sum_{i=1}^{N} R_2(i)}$$

where:

$$\theta^{old} = (\lambda, p_1, p_2)$$
$$R_1(i) = \frac{\lambda p_1^{h(x_i)}(1-p_1)^{t(x_i)}}{\lambda p_1^{h(x_i)}(1-p_1)^{t(x_i)} + (1-\lambda)p_2^{h(x_i)}(1-p_2)^{t(x_i)}}$$
$$R_2(i) = \frac{\lambda p_2^{h(x_i)}(1-p_2)^{t(x_i)}}{\lambda p_1^{h(x_i)}(1-p_1)^{t(x_i)} + (1-\lambda)p_2^{h(x_i)}(1-p_2)^{t(x_i)}}$$