Machine Learning 1 - Exercise 5

Fränz Beckius (374057)
Ivan David Aranzales Acero (399364)
Janek Tichy (584200)
Jeremias Eichelbaum (358685)
Alessandro Schneider (364988)

November 18, 2018

1 Discrete EM: Coin Tosses from Multiple Distributions

1.a

We rewrite $Q(\Theta, \Theta^{old})$ as:

$$\begin{aligned}
Q(\Theta, \Theta^{old}) &= \sum_{z \in \{heads, tails\}} P(Z = z|X = x, \Theta^{old}) log[P(X = x, Z = z|\Theta)] && (1)\\
&= E_{Z|X,\Theta^{old}} log[p(Z, X|\Theta)] && (2)\\
&= E_{Z|X,\Theta^{old}} log \prod_{i=0}^{N} p(z_i, x_i|\Theta) && (3)\\
&= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})[log\lambda + x_i log(p_1) + (1 - x_i)log(1 - p_1)] && (4)\\
&+ (1 - p(z_i = 1|x_i, \Theta^{old}))[log\lambda + x_i log(p_2) + (1 - x_i)log(1 - p_2)] && (5)
\end{aligned}$$

And we rewrite $p(z_i|x_i, \Theta^{old})$ as:

$$\begin{aligned}
p(z_i = 1|x_i, \Theta^{old}) &= \frac{p(x_i|z_i, \Theta^{old})p(z_i = 1|\Theta^{old})}{p(x_i|\Theta^{old})} && (6)\\
&= \frac{\lambda p_{1old}^{x_i}(1 - p_{1old})^{1-x_i}}{\lambda p_{1old}^{x_i}(1 - p_{1old})^{1-x_i} + (1 - \lambda)p_{2old}^{x_i}(1 - p_{2old})^{1-x_i}} && (7)\\
& && (8)
\end{aligned}$$

We now solve:

$$\lambda^{new} = \frac{\partial Q(\Theta, \Theta^{old})}{\partial \lambda} \tag{9}$$

$$= \sum_{i=0}^{N} \frac{p(z_i = 1|x_i, \Theta^{old})}{\lambda} + \frac{1}{1-\lambda} + \frac{p(z_i = 1|x_i, \Theta^{old})}{1-\lambda} \tag{10}$$

$$= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})(1-\lambda) + \lambda + p(z_i = 1|x_i, \Theta^{old})\lambda \tag{11}$$

$$= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old}) \tag{12}$$

$$= \sum_{i=0}^{N} \frac{\lambda p_{1old}^{x_i}(1-p_{1old})^{1-x_i}}{\lambda p_{1old}^{x_i}(1-p_{1old})^{1-x_i} + (1-\lambda) p_{2old}^{x_i}(1-p_{2old})^{1-x_i}} \tag{13}$$

$$P_1^{new} = \frac{\partial Q(\Theta, \Theta^{old})}{\partial P_1} \tag{14}$$

$$= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})(\frac{x_i}{p_1} - \frac{1-x_i}{1-p_1}) \tag{15}$$

$$= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})(x_i - x_i * p_1 - p_1 + xi * p_1) \tag{16}$$

$$= \sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})(x_i - p_1) \tag{17}$$

$$p_1 = \frac{\sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old}) x_i}{\sum_{i=0}^{N} p(z_i = 1|x_i, \Theta^{old})} \tag{18}$$

$$P_2^{new} = \frac{\partial Q(\Theta, \Theta^{old})}{\partial P_2} \tag{19}$$

$$= \sum_{i=0}^{N} (1 - p(z_i = 1|x_i, \Theta^{old}))(\frac{x_i}{p_2} - \frac{1-x_i}{1-p_2}) \tag{20}$$

$$= \sum_{i=0}^{N} (1 - p(z_i = 1|x_i, \Theta^{old}))(x_i - x_i * p_2 - p_2 + x_i * p_2 \tag{21}$$

$$= \sum_{i=0}^{N} (1 - p(z_i = 1|x_i, \Theta^{old}))(xi - p_2) \tag{22}$$

$$p_2 = \frac{\sum_{i=0}^{N} (1 - p(z_i = 1|x_i, \Theta^{old})) x_i}{\sum_{i=0}^{N} (1 - p(z_i = 1|x_i, \Theta^{old}))} \tag{23}$$

$$\tag{24}$$