Machine Learning 1
Exercise 5

Group: BSSBCH

November 20, 2017

Matthias Bigalke, 339547, maku@win.tu-berlin.de
Tolga Buz, 346836, buz_tolga@yahoo.de
Alejandro Hernandez, 395678, alejandrohernandezmunuera@gmail.com
Aitor Palacios Cuesta, 396276, aitor.palacioscuesta@campus.tu-berlin.de
Christof Schubert, 344450, christof.schubert@campus.tu-berlin.de
Daniel Steinhaus, 342563, dany.steinhaus@googlemail.com

Exercise 1

We are searching for

$$\theta^{\text{new}} = \text{argmax}_\theta Q(\theta, \theta^{old})$$

We have

$$P(Z = heads|\theta) = \lambda,\ P(Z = tails|\theta) = 1 - \lambda \tag{1}$$

and for $\mathbf{X} = (X_1, ..., X_m)$

$$P(X_i = heads|Z = heads, \theta) = p_1,\ P(X_i = tails|Z = heads, \theta) = 1 - p_1 \tag{2}$$

$$P(X_i = heads|Z = tails, \theta) = p_2,\ P(X_i = tails|Z = tails, \theta) = 1 - p_2 \tag{3}$$

For $x = (x^{(1)}, ..., x^{(N)})$ and $z = (z^{(1)}, ..., z^{(N)})$ we have:

$$P(\mathcal{X} = x, \mathcal{Z} = z|\theta) = \prod_{i=1}^{N} P(Z = z^{(i)}|\theta) \prod_{j=1}^{m} P(X_j = x_j^{(i)}|Z = z^{(i)}, \theta) \tag{4}$$

With this we get (*):

$$Q(\theta, \theta^{old}) = \sum_{z\in\{h,t\}^N} \underbrace{P(\mathcal{Z} = z|\mathcal{X} = x, \theta^{old})}_{:=q(z)} \log P(\mathcal{X} = x, \mathcal{Z} = z|\theta)$$

$$= \sum_{z\in\{h,t\}^N} q(z) \log P(\mathcal{X}=x, \mathcal{Z}=z|\theta)$$

$$\overset{(4)}{=} \sum_{z\in\{h,t\}^N} q(z) \log \left[\prod_{i=1}^{N} P(Z=z^{(i)}|\theta) \prod_{j=1}^{m} P(X_j = x_j^{(i)}|Z=z^{(i)},\theta)\right]$$

$$\overset{(1)}{=} \sum_{z\in\{h,t\}^N} q(z) \log \left[\lambda^{a_z}(1-\lambda)^{N-a_z} \prod_{i=1}^{N}\prod_{j=1}^{m} P(X_j = x_j^{(i)}|Z=z^{(i)},\theta)\right]$$

$$\overset{(2),(3)}{=} \sum_{z\in\{h,t\}^N} q(z) \log[\lambda^{a_z}(1-\lambda)^{N-a_z} p_1^{b_z}(1-p_1)^{a_z\cdot m-b_z} p_2^{c_z}(1-p_2)^{(N-a_z)\cdot m-c_z}]$$

Where

- a_z is the amount of $z^{(i)} = h$ for a specific $z \in \{h,t\},\ i \in \{1,...,N\}$
- b_z is the amount of $x_j^{(i)} = h,\ j \in \{1,...,m\}$ with corresponding $z^{(i)} = h$
- c_z is the amount of $x_j^{(i)} = h,\ j \in \{1,...,m\}$ with corresponding $z^{(i)} = t$

We should also highlight, that $q(z)$ is completely known, since

$$q(z) = P(\mathcal{Z}=z|\mathcal{X}=x,\theta^{old}) = \frac{P(\mathcal{Z}=z,\mathcal{X}=x|\theta^{old})}{P(\mathcal{X}=x|\theta^{old})} = \frac{P(\mathcal{Z}=z,\mathcal{X}=x|\theta^{old})}{\sum_{z\in\{h,t\}^N} P(\mathcal{Z}=z,\mathcal{X}=x|\theta^{old})}$$

which can be calculated using $(1)-(4)$.

Calculation of $\theta^{\mathbf{new}} = (\hat{\lambda}, \hat{p}_1, \hat{p}_2)$

Calculation of $\hat{\lambda}$:

$$\frac{\partial}{\partial\lambda} Q(\theta,\theta^{old}) \overset{(*)}{=} \frac{\partial}{\partial\lambda} \sum_{z\in\{h,t\}^N} q(z) \log[\lambda^{a_z}(1-\lambda)^{N-a_z} p_1^{b_z}(1-p_1)^{a_z\cdot m-b_z} p_2^{c_z}(1-p_2)^{(N-a_z)\cdot m-c_z}]$$

$$= \sum_{z\in\{h,t\}^N} q(z) \frac{\partial}{\partial\lambda}[a_z\cdot\log(\lambda) + (N-a_z)\log(1-\lambda)]$$

$$= \sum_{z\in\{h,t\}^N} q(z) \left(\frac{a_z}{\lambda} - \frac{N-a_z}{1-\lambda}\right)$$

$$= \sum_{z\in\{h,t\}^N} q(z) \frac{a_z - a_z\lambda - N\lambda + a_z\lambda}{\lambda(1-\lambda)}$$

$$= \sum_{z\in\{h,t\}^N} q(z) \frac{a_z - N\lambda}{\lambda(1-\lambda)}$$

and we get:

$$\begin{aligned}
& 0 = \frac{\partial}{\partial \lambda} Q(\theta, \theta^{old}) \\
\Leftrightarrow \quad & 0 = \sum_{z \in \{h,t\}^N} q(z) \left(\frac{a_z - N\hat{\lambda}}{\hat{\lambda}(1-\hat{\lambda})} \right) \\
\Leftrightarrow \quad & \sum_{z \in \{h,t\}^N} q(z) N \hat{\lambda} = \sum_{z \in \{h,t\}^N} q(z) a_z \\
\Leftrightarrow \quad & \hat{\lambda} = \frac{1}{N} \cdot \frac{\sum_{z \in \{h,t\}^N} q(z) a_z}{\sum_{z \in \{h,t\}^N} q(z)}
\end{aligned}$$

Calculation of $\hat{p}_1$:

$$\begin{aligned}
\frac{\partial}{\partial p_1} Q(\theta, \theta^{old}) & \overset{(*)}{=} \frac{\partial}{\partial p_1} \sum_{z \in \{h,t\}^N} q(z) \log[\lambda^{a_z}(1-\lambda)^{N-a_z} p_1^{b_z} (1-p_1)^{a_z \cdot m - b_z} p_2^{c_z} (1-p_2)^{(N-a_z)\cdot m - c_z}] \\
& = \sum_{z \in \{h,t\}^N} q(z) \frac{\partial}{\partial p_1} [b_z \cdot \log(p_1) + (a_z \cdot m - b_z) \log(1-p_1)] \\
& = \sum_{z \in \{h,t\}^N} q(z) \left(\frac{b_z}{p_1} - \frac{a_z \cdot m - b_z}{1-p_1} \right) \\
& = \sum_{z \in \{h,t\}^N} q(z) \frac{b_z - b_z p_1 - a_z m p_1 + b_z p_1}{p_1(1-p_1)} \\
& = \sum_{z \in \{h,t\}^N} q(z) \frac{b_z - a_z m p_1}{p_1(1-p_1)}
\end{aligned}$$

and we get:

$$\begin{aligned}
& 0 = \frac{\partial}{\partial p_1} Q(\theta, \theta^{old}) \\
\Leftrightarrow \quad & 0 = \sum_{z \in \{h,t\}^N} q(z) \frac{b_z - a_z m \hat{p}_1}{\hat{p}_1(1-\hat{p}_1)} \\
\Leftrightarrow \quad & \sum_{z \in \{h,t\}^N} q(z) a_z m \hat{p}_1 = \sum_{z \in \{h,t\}^N} q(z) b_z \\
\Leftrightarrow \quad & \hat{p}_1 = \frac{1}{m} \cdot \frac{\sum_{z \in \{h,t\}^N} q(z) b_z}{\sum_{z \in \{h,t\}^N} q(z) a_z}
\end{aligned}$$

Calculation of $\hat{p_2}$

$$
\begin{aligned}
\frac{\partial}{\partial p_2} Q(\theta, \theta^{old}) &\overset{(*)}{=} \frac{\partial}{\partial p_2} \sum_{z \in \{h,t\}^N} q(z) \log[\lambda^{a_z}(1-\lambda)^{N-a_z} p_1^{b_z}(1-p_1)^{a_z \cdot m - b_z} p_2^{c_z}(1-p_2)^{(N-a_z)\cdot m - c_z}] \\
&= \sum_{z \in \{h,t\}^N} q(z) \frac{\partial}{\partial p_2} [c_z \cdot \log(p_2) + ((N-a_z)\cdot m - c_z) \log(1-p_2)] \\
&= \sum_{z \in \{h,t\}^N} q(z) \left(\frac{c_z}{p_2} - \frac{(N-a_z)\cdot m - c_z}{1-p_2} \right) \\
&= \sum_{z \in \{h,t\}^N} q(z) \frac{c_z - c_z p_2 - (N-a_z) m p_2 + c_z p_2}{p_2(1-p_2)} \\
&= \sum_{z \in \{h,t\}^N} q(z) \frac{c_z - (N-a_z) m p_2}{p_2(1-p_2)}
\end{aligned}
$$

and we get:

$$
\begin{aligned}
& 0 = \frac{\partial}{\partial p_2} Q(\theta, \theta^{old}) \\
\Leftrightarrow\quad & 0 = \sum_{z \in \{h,t\}^N} q(z) \frac{c_z - (N-a_z) m \hat{p}_2}{\hat{p}_2(1-\hat{p}_2)} \\
\Leftrightarrow\quad & \sum_{z \in \{h,t\}^N} q(z)(N-a_z) m \hat{p}_2 = \sum_{z \in \{h,t\}^N} q(z) c_z \\
\Leftrightarrow\quad & \hat{p}_2 = \frac{1}{m} \cdot \frac{\sum_{z \in \{h,t\}^N} q(z) c_z}{\sum_{z \in \{h,t\}^N} q(z)(N-a_z)}
\end{aligned}
$$

Exercise 2

See next page