HUSZAR

October 18, 2016

We want to demonstrate equation (5) of Ferenc Huszár, 2015

$$
\begin{aligned}
D_{alternative}\left[P \mid\mid Q\right] &= KL\left[P_{x_1} \mid\mid Q_{x_1}\right] + \mathbb{E}_{y\sim P_{x_1}}\mathbb{E}_{z\sim Q_{x_1}} KL\left[P_{x_2|x_1=y} \mid\mid Q_{x_2|x_1=z}\right] \text{ (4)} \\
&= KL\left[P_{x_1} \mid\mid Q_{x_1}\right] + \mathbb{E}_{z\sim Q_{x_1}} KL\left[P_{x_2} \mid\mid Q_{x_2|x_1=z}\right] \text{ (5)}
\end{aligned}
\tag{1}
$$

We study the second right hand side term of the first line (4)

$$
\begin{aligned}
\mathbb{E}_{y\sim P_{X_1}}\mathbb{E}_{z\sim Q_{X_1}} KL\left[P_{X_2|X_1=y} \mid\mid Q_{X_2|X_1=z}\right] &= \mathbb{E}_{z\sim Q_{X_1}} \sum_{y\sim P_{X_1}} P_{X_1}(X_1=y) \sum_{x_2\sim P_{X_2|X_1=z}(\cdot|X_1=y)} P_{X_2|X_1=y}(X_2=x_2 \mid X \\
&= \mathbb{E}_{z\sim Q_{X_1}} \sum_{(y,x2)\sim P_{X_1,X_2}} P_{X_1,X_2}(X_1=y, X_2=x_2)\log\left[\frac{P_{X_2|X_1=y}(X_2=x}{Q_{X_2|X_1=z}(X_2=x}\right. \\
&= \mathbb{E}_{z\sim Q_{X_1}} \sum_{x_2\sim P_{X_2}} P_{X_2}(X_2=x_2) \sum_{y\sim P_{X_1|X_2=x_2}(\cdot|X_2=x_2)} P_{X_1|X_2=x_2}(X_1=y \mid
\end{aligned}
\tag{2}
$$

Now by Bayes' Rule we know that:

$$
P_{X_2|X_1=y}(X_2=x_2 \mid X_1=y) = \frac{P_{X_2}(X_2=x_2)}{P_{X_1}(X_1=y)} P_{X_1|X_2=x_2}(X_1=y \mid X_2=x_2)
$$

So we can bring the terms that are only function of x_2 and z in our first sum to write our previous equation

$$
\begin{aligned}
\mathbb{E}_{y\sim P_{X_1}}\mathbb{E}_{z\sim Q_{X_1}} KL\left[P_{X_2|X_1=y} \mid\mid Q_{X_2|X_1=z}\right] &= \mathbb{E}_{z\sim Q_{X_1}} \sum_{x_2\sim P_{X_2}} P_{X_2}(X_2=x_2)\log\left[\frac{P_{X_2}(X_2=x_2)}{Q_{X_2=x_2|X_1=z}(X_2=x_2 \mid X_1=z)}\right] \\
&\sum_{y\sim P_{X_1|X_2=x_2}(\cdot|X_2=x_2)} P_{X_1|X_2=x_2}(X_1=y \mid X_2=x_2)\log\left[\frac{P_{X_1|X_2=x_2}(X_1=}{P_{X_1}(X_1=}\right.
\end{aligned}
\tag{3}
$$

If this last coefficient

$$
\sum_{y\sim P_{X_1|X_2=x_2}(\cdot|X_2=x_2)} P_{X_1|X_2=x_2}(X_1=y \mid X_2=x_2)\log\left[\frac{P_{X_1|X_2=x_2}(X_1=y \mid X_2=x_2)}{P_{X_1}(X_1=y)}\right] = KL\left[P_{X_1|X_2=x_2} \mid\mid P_{X_1}\right]
$$

is equal to one, we end up with the expression we were looking for

$$\mathbb{E}_{y\sim P_{X_1}}\mathbb{E}_{z\sim Q_{X_1}}KL\left[P_{X_2|X_1=y}\;||\;Q_{X_2|X_1=z}\right]=\mathbb{E}_{z\sim Q_{x_1}}KL\left[P_{x_2}\;||\;Q_{x_2|x_1=z}\right]$$

But I must confess it is not clear to me why this KL-divergence should be equal to one...

```
In [ ]:
```