

HUSZAR

October 18, 2016

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

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

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

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

Now by Bayes' Rule we know that:

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

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

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

If this last coefficient

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

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

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

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

In []: