Утверждение 1

Максимизация вариационной нижней оценки

$$\int_{\mathbf{w}} q(\mathbf{w}) \log \frac{p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h})}{q(\mathbf{w})} d\mathbf{w}$$

эквивалентна минимизации расстояния Кульбака–Лейблера между распределением $q(\mathbf{w}) \in \mathcal{Q}$ и апостериорным распределением параметров $p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})$:

$$\hat{q} = \arg\max_{q \in \mathcal{Q}} \int_{\mathbf{w}} q(\mathbf{w}) \log \frac{p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h})}{q(\mathbf{w})} d\mathbf{w} \Longleftrightarrow \hat{q} = \arg\min_{q \in \mathcal{Q}} D_{KL}(q(\mathbf{w}) \| p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})),$$

$$D_{KL}(q(\mathbf{w}) \| p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})) = \int_{\mathbf{w}} q(\mathbf{w}) \log(\frac{q(\mathbf{w})}{p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})})$$

Доказательство

Рассмотрим выражения.

$$\begin{aligned} \arg\max_{q \in \mathcal{Q}} \int_{\mathbf{w}} q(\mathbf{w}) \log \frac{p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h})}{q(\mathbf{w})} d\mathbf{w} &= \\ = \arg\min_{q \in \mathcal{Q}} \int_{\mathbf{w}} -q(\mathbf{w}) \log \frac{p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h})}{q(\mathbf{w})} d\mathbf{w} &= \\ = \arg\min_{q \in \mathcal{Q}} \int_{\mathbf{w}} q(\mathbf{w}) \log \frac{q(\mathbf{w})}{p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h})} d\mathbf{w} & \quad (1) \end{aligned}$$

$$\arg\min_{q \in \mathcal{Q}} D_{KL}(q(\mathbf{w}) \| p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})) = \arg\min_{q \in \mathcal{Q}} \int_{\mathbf{w}} q(\mathbf{w}) \log \frac{q(\mathbf{w})}{p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h})} d\mathbf{w} \quad (2)$$

Тогда утверждение теоремы равнозначно следующему утверждению:

$$p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h}) \sim p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h}) \quad (3)$$

$$p(\mathbf{w}|\mathbf{y}, \mathbf{X}, \mathbf{h}) = \frac{p(\mathbf{y}|\mathbf{w}, \mathbf{X}, \mathbf{h}) p(\mathbf{w}|\mathbf{X}, \mathbf{h})}{\int_{\mathbf{w}} p(\mathbf{y}|\mathbf{w}) p(\mathbf{w}|\mathbf{X}, \mathbf{h}) d\mathbf{w}} \quad (4)$$

$$p(\mathbf{y}, \mathbf{w}, |\mathbf{X}, \mathbf{h}) = p(\mathbf{y}|\mathbf{w}, \mathbf{X}, \mathbf{h}) p(\mathbf{w}|\mathbf{X}, \mathbf{h}) \quad (5)$$

Учитывая, что интеграл в знаменателе (4) из формулы Байеса ялвяется нормировочной константой, то утверждение (3) является верным, ч.т.д.