Доказательство утверждения 2

Выполнил: Плетнев Никита

Преобразуем данное выражение, поделённое на λ (постоянный множитель никак не влияет на оптимизацию):

$$\frac{1}{\lambda}\mathbb{E}_{q(w)}\log p(y|X,w) - \mathrm{D}_{KL}(q(w)||p(w|y,X,h)) = \mathbb{E}_{q(w)}\frac{m}{\lambda}\mathbb{E}\log p(y_i|x_i,w) - \mathrm{D}_{KL}(q(w)||p(w|y,X,h)).$$

Усиленный закон больших чисел гарантирует, что выборочное среднее независимых одинаково распределённых случайных величин почти наверное стремится к их мат. ожиданию. Поэтому если внутреннее мат. ожидание в этом выражении заменить на среднее по подвыборке, получится эквивалентное при больших $\frac{m}{\lambda}$ выражение. Так же можно поступить и со средним в определении KL-дивергенции: взять подвыборку вместо выборки. Получим эквивалентное выражение:

$$\mathbb{E}_{q(w)}\frac{m}{\lambda}\frac{\lambda}{m}\sum_{i=1}^{\frac{m}{\lambda}}\log p(y_i|x_i,w) - \mathrm{D}_{KL}(q(w)||p(w|\widehat{y},\widehat{X},h)) = \mathbb{E}_{q(w)}\log p(\widehat{y}|\widehat{X},w) - \mathrm{D}_{KL}(q(w)||p(w|\widehat{y},\widehat{X},h))$$

— а это и есть вариационная оценка обоснованности для подвыборки мощности $\frac{m}{\lambda}$.

Получили, что оптимизация исходного выражения эквивалентна оптимизации вариационной оценки обоснованности для случайной подвыборки мощности $\frac{m}{\lambda}$, поскольку является её пределом при увеличении мощности подвыборки. Что и требовалось доказать.