Теорема 1. *Пусть*

1. *Заданы компактные множества* $\mathbf{U}_{\mathbf{h}} \subset \mathbb{H}, \mathbf{U}_{\boldsymbol{\theta}_{\mathbf{w}}} \subset \boldsymbol{\theta}_{\mathbf{w}}, \mathbf{U}_{\boldsymbol{\theta}_{\boldsymbol{\Gamma}}} \subset \boldsymbol{\theta}_{\boldsymbol{\Gamma}}$.

2. *Вариационное распределение* $q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)$ *является абсолютно непрерывным и унимодальным на* $U_{\boldsymbol{\theta}}$*. Его мода и матожидание совпадают:*

$$\boldsymbol{mode}\ q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right) = \mathsf{E}_{q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)}\mathbf{w}.$$

3. *Априорное распределение* $p\left(\mathbf{w}|\boldsymbol{\Gamma}, \mathbf{h}, \boldsymbol{\lambda}\right)$ *является абсолютно непрерывным и унимодальным на* $U_{\mathbf{h}}$*. Его мода и матожидание совпадают и не зависят от гиперпараметров* $\mathbf{h}$ *на* $\mathbf{U}_{\mathbf{h}}$ *и структуры* Γ *на* $\mathbf{U}_{\boldsymbol{\theta}_{\boldsymbol{\Gamma}}}$*:* $\mathsf{E}_{p(\mathbf{w}|\boldsymbol{\Gamma}, \mathbf{h}, \boldsymbol{\lambda})}\ \mathbf{w} = \boldsymbol{mode}\ p\left(\mathbf{w}|\boldsymbol{\Gamma}_1, \mathbf{h}_1, \boldsymbol{\lambda}\right) = \boldsymbol{mode}\ p\left(\mathbf{w}|\boldsymbol{\Gamma}_1, \mathbf{h}_2, \boldsymbol{\lambda}\right) = \mathbf{m}$ *для любых* $\mathbf{h}_1, \mathbf{h}_2 \in \mathbf{U}_{\mathbf{h}}, \Gamma_1, \Gamma_2 \in \mathbf{U}_{\boldsymbol{\Gamma}}$.

4. *Параметры модели* $\mathbf{w}$ *имеют конечные вторые моменты по маргинальным распределениям:* $\int_{\Gamma} q_{\boldsymbol{\Gamma}}\left(\boldsymbol{\Gamma}|\boldsymbol{\theta}_{\boldsymbol{\Gamma}}\right) q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)_{\mathbf{w}} d\Gamma, \quad \int_{\Gamma} q_{\boldsymbol{\Gamma}}\left(\boldsymbol{\Gamma}|\boldsymbol{\theta}_{\boldsymbol{\Gamma}}\right) p\left(\mathbf{w}|\boldsymbol{\Gamma}, \mathbf{h}, \boldsymbol{\lambda}\right) d\Gamma$ *при любых* $\boldsymbol{\theta}_{\mathbf{w}} \in \mathbf{U}_{\boldsymbol{\theta}_{\mathbf{w}}}, \boldsymbol{\theta}_{\boldsymbol{\Gamma}} \in \mathbf{U}_{\boldsymbol{\theta}_{\boldsymbol{\Gamma}}}, \mathbf{h} \in \mathbf{U}_{\mathbf{h}}$.

5. *Вариационное распределение* $q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)$ *является липшицевым по* $\mathbf{w}$ *с параметров* L.

6. *Значение* $q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)$ *не равно нулю при любых* $\boldsymbol{\theta} \in \mathbf{U}_{\boldsymbol{\theta}}, \Gamma \in \mathbb{G}$.

7. *Точная нижняя грань* $q_{\mathbf{w}}\left(\mathbf{m}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right)$ *не равна нулю при* $\boldsymbol{\theta}_{\mathbf{w}} \in \mathbf{U}_{\boldsymbol{\theta}_{\mathbf{w}}}$ *и* $\Gamma \in \mathbb{G}$*:*

$$\inf_{\Gamma \in \mathbb{G}, \boldsymbol{\theta}_{\mathbf{w}} \in \mathbf{U}_{\boldsymbol{\theta}_{\mathbf{w}}}} q_{\mathbf{w}}\left(\mathbf{m}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right) > 0.$$

Тогда

$$\left|\mathsf{E}_{q_{\boldsymbol{\Gamma}}(\boldsymbol{\Gamma}|\boldsymbol{\theta}_{\boldsymbol{\Gamma}})}\rho(\mathbf{w}|\Gamma, \boldsymbol{\theta}_{\mathbf{w}}, \mathbf{h}, \boldsymbol{\lambda})^{-1} - 1\right| \leq Const \iint_{\Gamma, \mathbf{w}} |\mathbf{w}| \cdot |q_{\mathbf{w}}\left(\mathbf{w}|\boldsymbol{\Gamma}, \boldsymbol{\theta}_{\mathbf{w}}\right) - p\left(\mathbf{w}|\boldsymbol{\Gamma}, \mathbf{h}, \boldsymbol{\lambda}\right)| q_{\boldsymbol{\Gamma}}\left(\boldsymbol{\Gamma}|\boldsymbol{\theta}_{\boldsymbol{\Gamma}}\right) d\mathbf{w} d\Gamma.$$

[1]Занесли модуль под знак интеграла
[2]Условие Лившица из 5
[3]Из 7го условия теоремы
[4]Из 2 и 3 условия
[5]Перепишем математическое ожидание через интеграл и внесем модуль под знак интеграла

Доказательство.

$$
\begin{aligned}
&\left|\mathsf{E}_{q_{\mathbf{\Gamma}}(\mathbf{\Gamma}|\boldsymbol{\theta}_{\boldsymbol{\theta}})}\rho\left(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}},\mathbf{h},\boldsymbol{\lambda}\right)^{-1}-1\right| = \left|\int_{\mathbf{\Gamma}}\left(\frac{q_{\mathbf{w}(\text{mode }q_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}{q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}\right)q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma}-1\right| = \\
&= \left|\int_{\mathbf{\Gamma}}\left(\frac{q_{\mathbf{w}(\text{mode }q_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}-q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}{q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}\right)q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma}\right| \leq^{1} \\
&\leq \int_{\mathbf{\Gamma}}\left|\left(\frac{q_{\mathbf{w}(\text{mode }q_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}-q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}{q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}\right)\right|q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma} \leq^{2} \\
&\leq L\int_{\mathbf{\Gamma}}\left(\frac{\left|\text{mode }q_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}}\right)-\text{mode }p_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda}\right)\right|}{q_{\mathbf{w}(\text{mode }p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}}\right)q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma} \leq^{3} \\
&\leq Const\int_{\mathbf{\Gamma}}\left|\text{mode }q_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}}\right)-\text{mode }p_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda}\right)\right|q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma} =^{4} \\
&= Const\int_{\mathbf{\Gamma}}\left|\mathsf{E}_{q_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}})}\mathbf{w}-\mathsf{E}_{p_{\mathbf{w}}(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda})}\mathbf{w}\right|q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{\Gamma} \leq^{5} \\
&\leq Const\int_{\mathbf{\Gamma}}\int_{\mathbf{w}}\left|\mathbf{w}\right|\left|q_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\boldsymbol{\theta}_{\mathbf{w}}\right)-p_{\mathbf{w}}\left(\mathbf{w}|\mathbf{\Gamma},\mathbf{h},\boldsymbol{\lambda}\right)\right|q_{\mathbf{\Gamma}}\left(\mathbf{\Gamma}|\boldsymbol{\theta}_{\mathbf{\Gamma}}\right)d\mathbf{w}d\mathbf{\Gamma}
\end{aligned}
$$

□