Оптимальная оценка для алгоритма экспоненциального взвешивания

В алгоритме экспоненциального взвешивания экспертных решений, порождаемым потенциальной функцией

$$\Phi_\varepsilon(u) = \frac{1}{\varepsilon} \ln \left(\sum_{k=1}^{N} e^{\varepsilon u_k} \right),$$

предсказание осуществляется по формуле

$$\hat{p}_t = \frac{\sum_{k=1}^N w_k(t-1) f_{k,t}}{\sum_{k=1}^N w_k(t-1)},$$

где

$$w_k(t-1) = \frac{e^{\varepsilon R_{k,t-1}}}{\sum_{j=1}^N e^{\varepsilon R_{j,t-1}}} = \frac{e^{\varepsilon(\hat{L}_{t-1} - L_{k,t-1})}}{\sum_{j=1}^N e^{\varepsilon(\hat{L}_{t-1} - L_{j,t-1})}} = \frac{e^{-\varepsilon L_{k,t-1}}}{\sum_{j=1}^N e^{-\varepsilon L_{j,t-1}}},$$

откуда $\hat{p}_t$ переписывается следующим образом:

$$\hat{p}_t = \frac{\sum_{k=1}^N e^{-\varepsilon L_{k,t-1}} f_{k,t}}{\sum_{k=1}^N e^{-\varepsilon L_{k,t-1}}}.$$

В качестве упражнения для алгоритма экспоненциального взвешивания экспертных решений, основываясь на общем аппарате для потенциальных функций, мы установили справедливость следующего утверждения.

Предложение 1. *Предположим, что функция потерь $l(\cdot,\cdot)$ является выпуклой по первому аргументу и принимает значения из отрезка $[0,1]$. Тогда для любых допустимых исходов и раунда $t \geq 1$ общий регрет для алгоритма экспоненциального взвешивания удовлетворяет неравенству*

$$R_{k,t} \leq \frac{\ln N}{\varepsilon} + \frac{\varepsilon t}{2}.$$

Целью этой лекции является получение более точной оценки, которая к тому же является оптимальной. Мы докажем следующее утверждение.

Теорема 1. *Предположим, что функция потерь $l(\cdot,\cdot)$ является выпуклой по первому аргументу и принимает значения из отрезка $[0,1]$. Тогда для любых допустимых исходов и раунда $t \geq 1$ общий регрет для алгоритма экспоненциального взвешивания удовлетворяет неравенству*

$$R_{k,t} \leq \frac{\ln N}{\varepsilon} + \frac{\varepsilon t}{8}.$$

Лемма 1 (неравенство Хёфдинга)**.** *Пусть X — случайная величина, принимающая значения из отрезка $[a,b]$. Тогда для любого $\alpha \in \mathbb{R}$ справедливо неравенство*

$$\ln \mathbb{E}\left[e^{\alpha X}\right] \leq \alpha \mathbb{E} X + \frac{\alpha^2 (b-a)^2}{8}.$$

Доказательство. Поскольку

$$\ln \mathbb{E}\left[e^{\alpha X}\right] = \ln \mathbb{E}\left[e^{\alpha(X-\mathbb{E}X)} \cdot e^{\alpha \mathbb{E}X}\right] = \alpha \mathbb{E}X + \ln \mathbb{E}\left[e^{\alpha(X-\mathbb{E}X)}\right],$$

то без уменьшения общности рассуждений будем предполагать, что $\mathbb{E}X = 0$. Тогда надо установить справедливость неравенства

$$\mathbb{E}\left[e^{\alpha X}\right] \leq \exp \frac{\alpha^2(b-a)^2}{8}. \tag{1}$$

В силу выпуклости для всех $x \in [a, b]$

$$e^{\alpha x} \leq \frac{x-a}{b-a} e^{\alpha b} + \frac{b-x}{b-a} e^{\alpha a}.$$

Поскольку $\mathbb{E}X = 0$, то

$$\mathbb{E}\left[e^{\alpha X}\right] \leq \frac{b}{b-a} e^{\alpha a} - \frac{a}{b-a} e^{\alpha b}.$$

Введем обозначение $p = -a/(b-a)$:

$$\begin{aligned} \mathbb{E}\left[e^{\alpha X}\right] &\leq p e^{\alpha b} + e^{\alpha a}(1-p) = \left(1 - p + p e^{\alpha(b-a)}\right) e^{\alpha a} = \\ &= \left(1 - p + p e^{\alpha(b-a)}\right) e^{-p\alpha(b-a)} \triangleq e^{\varphi(u)}, \end{aligned} \tag{2}$$

где $u = \alpha(b-a)$, откуда

$$\varphi(u) = -pu + \ln\left(1 - p + p e^{u}\right).$$

Далее,

$$\varphi'(u) = -p + \frac{p e^{u}}{1 - p + p e^{u}} = -p + \frac{p}{(1-p)e^{-u} + p}.$$

Заметим, что $\varphi(0) = \varphi'(0) = 0$. Наконец,

$$\varphi''(u) = \frac{p(1-p)e^{-u}}{\left[(1-p)e^{-u} + p\right]^2} \leq \frac{1}{4}.$$

Теперь из разложения Тейлора следует

$$\varphi(u) = \varphi(0) + u\varphi'(0) + \frac{u^2}{2}\varphi''(\xi) \leq \frac{u^2}{8} = \frac{\alpha^2(b-a)^2}{8}.$$

Используя последнее неравенство в (2), наконец, получаем (1). □

Доказательство теоремы. Введем величину

$$W_t = \sum_{k=1}^{N} w_k(t) = \sum_{k=1}^{N} e^{-\varepsilon L_{k,t}},$$

как обычно полагая, что $W_0 = N$. Идея доказательства состоит в оценивании величин вида

$$\ln \frac{W_t}{W_{t-1}}.$$

Сначала преобразуем это логарифм:

$$\ln \frac{W_t}{W_{t-1}} = \ln \frac{\sum_{k=1}^N e^{-\varepsilon L_{k,t}}}{\sum_{k=1}^N w_k(t-1)} = \ln \frac{\sum_{k=1}^N e^{-\varepsilon l(f_{k,t}, y_t)} e^{-\varepsilon L_{k,t-1}}}{\sum_{k=1}^N w_k(t-1)} = \ln \frac{\sum_{k=1}^N e^{-\varepsilon l(f_{k,t}, y_t)} w_k(t-1)}{\sum_{k=1}^N w_k(t-1)} =$$
$$= \ln \sum_{k=1}^N p_k(t-1) e^{-\varepsilon l(f_{k,t}, y_t)} = \ln \mathbb{E}\left[e^{-\varepsilon l(f_{k,t}, y_t)}\right].$$

Теперь, используя неравенство Хёфдинга, получим оценку сверху:

$$\ln \frac{W_t}{W_{t-1}} = \ln \mathbb{E}\left[e^{-\varepsilon l(f_{k,t}, y_t)}\right] \leq -\varepsilon \mathbb{E}[l(f_{k,t}, y_t)] + \frac{\varepsilon^2}{8} = -\varepsilon \sum_{k=1}^N p_k(t-1) \cdot l(f_{k,t}, y_t) + \frac{\varepsilon^2}{8}.$$

Принимая во внимание выпуклость $l(\cdot, \cdot)$ по первом аргументу, заключаем, что

$$\ln \frac{W_t}{W_{t-1}} = -\varepsilon \sum_{k=1}^N p_k(t-1) \cdot l(f_{k,t}, y_t) + \frac{\varepsilon^2}{8} \leq -\varepsilon l \left(\sum_{k=1}^N p_k(t-1) f_{k,t}, y_t \right) + \frac{\varepsilon^2}{8}.$$

Суммируя по t от 1 до T последнее неравенство получаем

$$\ln \frac{W_T}{W_0} \leq -\varepsilon \hat{L}_t + \frac{\varepsilon^2 t}{8}. \tag{3}$$

С другой стороны, оценка снизу для величины $\ln(W_t/W_0)$ приводит к неравенству

$$\ln \frac{W_T}{W_0} = \ln \left(\sum_{k=1}^N e^{-\varepsilon L_{k,t}} \right) - \ln N \geq \ln \left(\max_{k=1,\ldots,N} e^{-\varepsilon L_{k,t}} \right) - \ln N = -\varepsilon \min_{k=1,\ldots,N} L_{k,t} - \ln N. \tag{4}$$

Сочетая оценки (3) и (4), наконец, получаем неравенство

$$-\varepsilon \min_{k=1,\ldots,N} L_{k,t} - \ln N \leq -\varepsilon \hat{L}_t + \frac{\varepsilon^2 t}{8},$$

откуда следует желаемая оценка

$$\hat{L}_t - \min_{k=1,\ldots,N} L_{k,t} \leq \frac{\ln N}{\varepsilon} + \frac{\varepsilon t}{8}.$$

□

Упражнение 1. *Допустим известен горизонт игры T. Определите значение параметра ε, при котором оценка полученная в теореме является наилучшей.*

Упражнение 2. *Пусть $Y = \{0, 1\}$, $D = [0, 1]$, $l(x, y) = |x - y|$. Докажите, что общие потери алгоритма экспоненциального взвешивания экспертных решений не меньше общих потерь лучшего эксперта.*