

1. Нехай $X \subset \mathbb{R}^d$, $S = \{(x_i, y_i)\}_{i=1}^m \subset (X \times \{0, 1\})^m$. Пока-
 жати, що існує поліном P_S big координат $x \in X$ такий,
 що його "порогова" версія $h_S(x) = \begin{cases} 0, & P_S(x) \leq 0 \\ 1, & P_S(x) > 0 \end{cases}$
 має функція $h_S(x) = \begin{cases} y_i, & \exists i = \overline{1, m}: x = x_i, (x_i, y_i) \in S; \\ 0 & \text{у інших випадках} \end{cases}$.

Після того $P_S(x)$ повинна бути < 0 для $x = x_i$ таких, що
 $y_i = 0$ та x , що $\notin \{x_i\}_{i=1}^m$ та > 0 для $x = x_i$ з $y_i = 1$
 (нагадаємо, що тут $y_i = f(x_i)$ ознак. функ. x_i). Ви-
 користаємо евід. вираз $\bullet$ між точками: $|x - y|^2 =$
 $= \sum_{i=1}^d (x^i - y^i)^2$ (де $x = (x^1, \dots, x^d)$, $y = (y^1, \dots, y^d)$).

Тоді $P_S(x) = - \prod_{\substack{i = \overline{1, m}: \\ y_i = 1}} |x - x_i|^2$ має потрібну власт.
 це $= 0$ для $\forall x = x_i$ з $y_i = 1$ і < 0 у інших ван.

2. В умовах моделі ML з левією для фікс. іррегуляр

h єкн. помилку $S \mapsto L_S(h)$ можна розглядати як
випадкову величину на X^m ($S \in (X \times Y)^m$), але
тут вона єдин. функ. $\{\pi_i\}_{i=1}^m \in X^m$, тому можна
вважати $S \in X^m$) Показати, що її математичне
співідання (expectation) дорівнює справсній помилці:

$$E_{S \sim \mathcal{D}^m} (L_S(h)) = L_{\mathcal{D}, f}(h)$$

Нагадаємо: випадкова величина - це $\forall$ випадкова
ф-ція $\varphi: X \rightarrow \mathbb{R}$, де $(X, \mathcal{F}, \mathcal{P})$ - ймовірнісний
простір, тобто така, що $\varphi^{-1}(A) \in \mathcal{F} \forall$ борелівської
 $A \subset \mathbb{R}$, її мат. співідання - це інтеграл Лебега

$$E(\varphi) := \int_X \varphi(x) d\mathcal{P}(x)$$

Детально про інтегрування за Лебегом див., наприклад,
В. М. Кадець, Курс функціонального аналізу та теорії міри.

У нас X - це $(X^m, \mathcal{F}^m, \mathcal{D}^m)$, тому

$$E(L_S(h)) = \int_{X^m} L_S(h) d\mathcal{D}^m(S) = \int_{X^m} \frac{|\{i=1, \dots, m \mid h(x_i) \neq f(x_i)\}|}{m} d\mathcal{D}^m(S) =$$

$$= [\text{интеграл}] = \frac{1}{m} \int_{X^m} |\{i \mid h(x_i) \neq f(x_i)\}| d\mathcal{D}^m(S) \in$$

$\mathcal{P}$ -это сиг интегралом монотонно перемещая да суму
характеристических $1_A(S) := \begin{cases} 1, & S \in A \\ 0, & S \notin A \end{cases}, A \subset X^m$

$$\in \frac{1}{m} \int_{X^m} \sum_{i=1}^m 1_{\{S \in X^m \mid h(x_i) \neq f(x_i)\}} d\mathcal{D}^m(S) = [\text{интеграл}] = \frac{1}{m} \sum_{i=1}^m$$

$$\int_{X^m} 1_{\{S \mid h(x_i) \neq f(x_i)\}} d\mathcal{D}^m(S) = \frac{1}{m} \sum_{i=1}^m \mathcal{D}^m(\{S \mid h(x_i) \neq f(x_i)\}) \in$$

(до $\forall$ типу $\int_X 1_A d\mathcal{P} = \mathcal{P}(A)$). Пусть $\forall i$ монотонно сиг $\mathcal{D}^m$ -
из фактов $X \times \dots \times \{x \in X \mid h(x) \neq f(x)\} \times \dots \times X$, тогда

$$\text{за факт. } \mathcal{D}^m = 1 \\ \in \frac{1}{m} \sum_{i=1}^m \mathcal{D}(X) \times \dots \times \mathcal{D}(\{x \in X \mid h(x) \neq f(x)\}) \times \dots \times \mathcal{D}(X) = \\ = \frac{1}{m} \cdot m L_{\mathcal{D}, f}(h) = 1$$

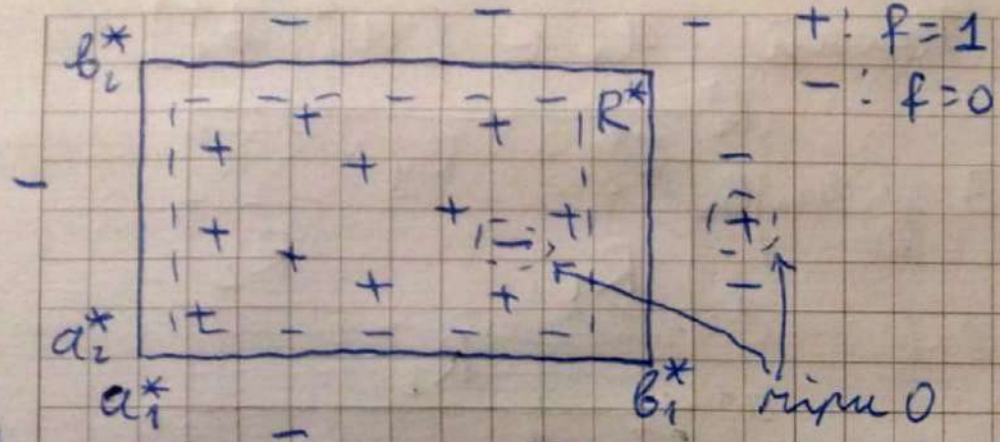
3. Показати, що клас $\mathcal{H}_{\text{мес}}^2$, що задані прямокутниками, створює функції паралельні до осей координат

$$h_{(a_1, b_1, a_2, b_2)}(x^1, x^2) := \begin{cases} 1, & a_1 \leq x^1 \leq b_1 \text{ і } a_2 \leq x^2 \leq b_2; \\ 0 & \text{у іншому випадку.} \end{cases}$$

$\mathcal{H}_{\text{мес}}^2 := \{h_{(a_1, b_1, a_2, b_2)} \mid a_1 \leq b_1, a_2 \leq b_2\}$ (для $X = [a, b] \times [c, d] \subset \mathbb{R}^2$ або $X = \mathbb{R}^2$) є PAC навчанням.

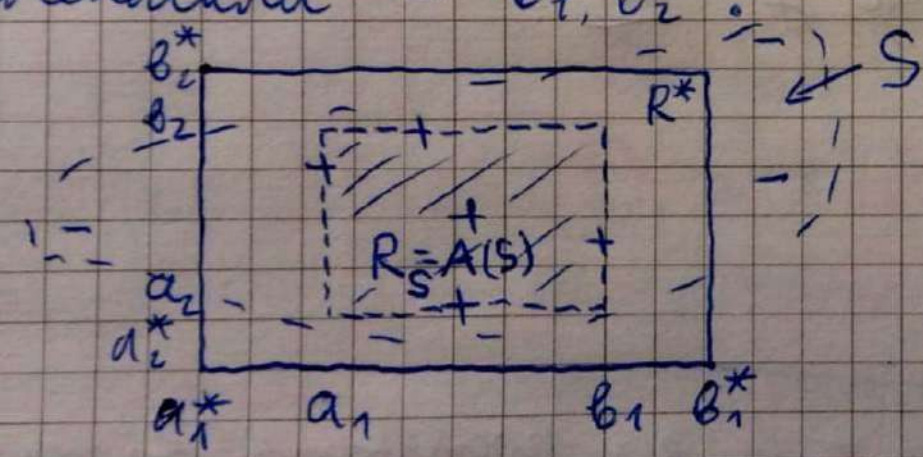
а. Нехай $\mathcal{H}_{\text{мес}}^2$ задовольняє прип. реалізації для деяких $\mathcal{D}, f$. Що буде результатом роботи $\text{ERM}_{\mathcal{H}_{\text{мес}}^2}$?

Прип. реалізації: $\exists h^* \in \mathcal{H}_{\text{мес}}^2 : \mathcal{D}(\{x \mid h^*(x) \neq f(x)\}) = 0$,
тобто $\exists$ прямокутник $R^* = [a_1^*, b_1^*] \times [a_2^*, b_2^*]$, що містить усі точки $x = (x^1, x^2)$ з $f(x) = 1$ і не містить точок з $f(x) = 0$, крім, можливо, деякої кінц. точки біля 0 (можу ними нехувати тут і далі нехувати):



Тоді це буде max і $\forall$
 навч. вибірка S : зі шов. 1
 усі "+" будуть всередині R^* , а
 "-" - ззовні, тому $\forall$ приклад.

$h(a_1, b_1, a_2, b_2) \in \mathbb{R}^2$, де $[a_1, b_1] \subset [a_1^*, b_1^*]$, $[a_2, b_2] \subset [a_2^*, b_2^*]$
~~можливо~~, що містить усі "+" з S , реалізує $ERM_{\mathcal{H}}$.
 Після для значеності наш алгоритм повертає найменший
 можливий пр-к, подібний з найменшими константами a_1, a_2
 і найменшими $-b_1, -b_2$:



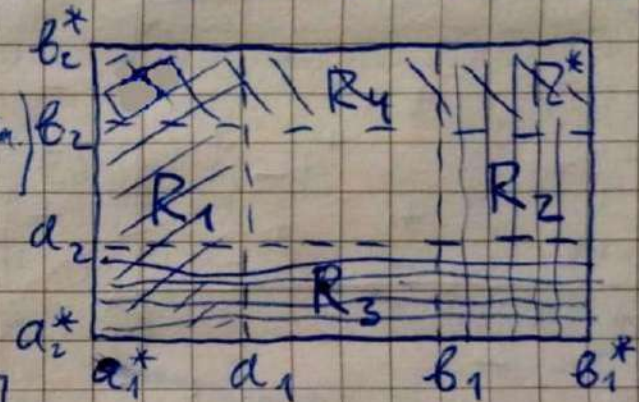
Бо для нього за
 побудовою $L_S(h) = 0$.

б. Тепер покажемо PAC навчальність для такої алг. A .

Омне, нап зам $\varepsilon, \delta \in (0, 1)$. Косан $a_1 \in [a_1^*, b_1^*]$

(де R^* - на буре) маке, що

$$\textcircled{1} \left(\underbrace{[a_1^*, a_1] \times [a_2^*, b_2^*]}_{R_1} \right) = \frac{\varepsilon}{4} \left(\exists \text{ за } \textcircled{2} \right)$$

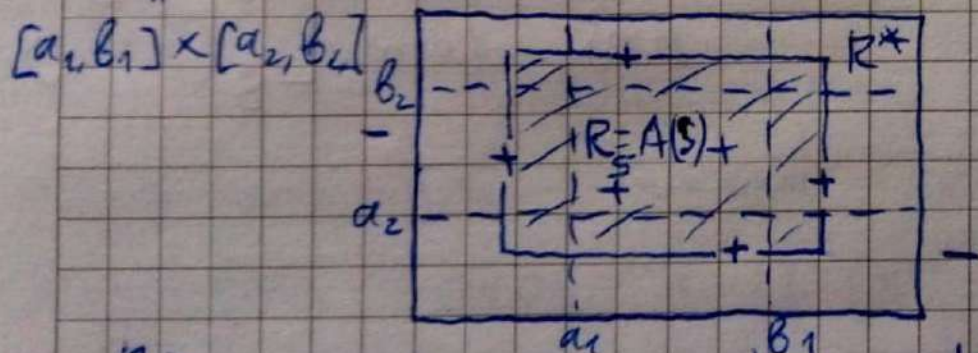


Ан-но суужено $R_2 := [b_1, b_1^*] \times [a_2^*, b_2^*]$,

$R_3 := [a_1^*, b_1^*] \times [a_2^*, a_2]$, $R_4 := [a_1^*, b_1^*] \times [b_2, b_2^*]$ где дееме
 $b_1 \in [a_1^*, b_1^*]$, $a_2, b_2 \in [a_2^*, b_2^*]$: $\textcircled{2}(R_i) = \frac{\varepsilon}{4} \quad \forall i$.

Косан менер S мае точки z $y_i = 1$ у косному z R_i ,

i $R_S = A(S)$ подуживаним ан у a !



Видн. инемежа $h_S = h(a_1, b_1, a_2, b_2)$ поставать бирни кинки

(0) (маинсе) усім точкам за менсани R^* ма бирни

(1) (найменше) усіх точкам всередині R , але ~~найменше~~ найбільш
 (0) - точкам мінімуму, що містяться у $\bigcup_{i=1}^y R_i$ за
 побудовою, тому $L_{\mathcal{D}, f}(h_S) = \mathcal{D}(R^* \setminus R_S) \leq [\text{монот.}] \leq$
 $\leq \mathcal{D}(\bigcup_{i=1}^y R_i) \leq [\text{сума}] \leq \sum_{i=1}^y \mathcal{D}(R_i) = \varepsilon$. Тому

$$\mathcal{D}^m(\{S \mid L_{\mathcal{D}, f}(h_S) > \varepsilon\}) \leq [\text{монот.}] \leq \mathcal{D}^m(\{S \mid \exists j = \overline{1, y} : \forall i = \overline{1, m} : y_i = 1 \Rightarrow x_i \notin R_j\})$$

$\downarrow$
 $y_i = 0 \Rightarrow x_i \notin R^*$
 $\downarrow$
 сума

$$= \mathcal{D}^m(\bigcup_{j=1}^y \{S \mid \forall i : y_i = 1 \Rightarrow x_i \notin R_j\}) \leq [\text{сума}]$$

$$\leq \sum_{j=1}^y \mathcal{D}^m((X \setminus R_j) \times \dots \times (X \setminus R_j)) \underset{\text{власн. } \mathcal{D}^m}{=} \sum_{j=1}^y (\mathcal{D}(X \setminus R_j))^m = y \cdot (1 - \frac{\varepsilon}{y})^m <$$

$$< y e^{-\frac{\varepsilon}{y} m} \quad (\leq)$$

Тому для $m \geq \frac{y}{\varepsilon} \ln \frac{y}{\delta} \quad (\leq) \delta$, маємо

$$\mathcal{D}^m(\{S \mid L_{\mathcal{D}, f}(h_S) \leq \varepsilon\}) \geq 1 - \delta, \text{ що й потрібно.}$$

4. Показати, що складність вибірки $m_H(\varepsilon, \delta)$ з дов.
 PAC наближеності незростає за ε і δ :

$$\forall \varepsilon_1 \leq \varepsilon_2, \delta \quad m_H(\varepsilon_1, \delta) \geq m_H(\varepsilon_2, \delta)$$

$$\forall \varepsilon, \delta_1 \leq \delta_2 \quad m_H(\varepsilon, \delta_1) \geq m_H(\varepsilon, \delta_2)$$

Дійсно, нехай $\varepsilon_1 \leq \varepsilon_2$. Покі $\forall \delta$ вибірка S розміру
 $m \geq m_n(\varepsilon_1, \delta)$ з імод. $> 1-\delta$ має h_S з нормою
 $L_{\Phi, F}(h) \leq \varepsilon_1 \leq \varepsilon_2$, а отже $m_n(\varepsilon_2, \delta) \leq m_n(\varepsilon_1, \delta)$
 як найменше значення m з такою власт. Для
 $\delta_1 \leq \delta_2$ і $\forall \varepsilon$ все ап-но.

Д/З 1. І загони 3.:

- Ізобразити на $X \subset \mathbb{R}^d$ для заданої вибірки d (задача, знайти $m_{\mathcal{H}}(\epsilon, \delta)$)

- Показати, що час виконання алгоритма пропорційний до $d, \frac{1}{\epsilon}, \ln \frac{1}{\delta}$.

2. Показати, що для $X = \mathbb{R}^2$, $\mathcal{Y} = \{0, 1\}$ клас концентричних кіл з центром в 0, тобто $\mathcal{H} = \{h_{\epsilon} = 1_{\{|x| \leq \epsilon\}} \mid \epsilon \in \mathbb{R}_+\}$, PAC навчаний для $m_{\mathcal{H}}(\epsilon, \delta) \leq \left\lceil \frac{1}{\epsilon} \ln \frac{1}{\delta} \right\rceil$.