

Д/з. 1. Пусть $\mathcal{H}$ - класс ^(х.ф.) замкнутых крив в $\mathbb{R}^d$:

$$\mathcal{H} := \{ h_{\sigma, \epsilon} : x \mapsto \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{matrix} |x - \sigma| \leq \epsilon \\ \text{и ин. в.} \end{matrix} \} \quad \begin{matrix} \sigma \in \mathbb{R}^d, \\ \epsilon > 0. \end{matrix}$$

Покажем, что $VCdim(\mathcal{H}) \leq d+2$. Пусть, нехай $\exists C \subset \mathbb{R}^d$:

$|C| = d+2$ и C разд. $\mathcal{H}$. Рассмотрим $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^{d+1}$:

$x \mapsto (x, |x|^2)$. Тогда $\varphi(C)$ разбивается классом наиб-

~~разд.~~ просторов L_{d+1} . Пусть, нехай $(y_1, \dots, y_{d+2}) \in$

$\{0, 1\}^{d+2}$ - дов. конечная строка $\exists h_{\sigma, \epsilon} \in \mathcal{H} : h_{\sigma, \epsilon}(C_i) = y_i$

$\forall i$ (где $C = \{C_i\}_{i=1}^{d+2}$), тогда $|C_i - \sigma| \leq \epsilon$ для $y_i = 1$ и

$|C_i - \sigma| > \epsilon$ для $y_i = 0$. Тогда рассмотрим линейную $\langle \tilde{x}, \tilde{w} \rangle + b = 0$,

где $\tilde{w} = (2\sigma_1, \dots, 2\sigma_d, -1)$, $b = -|\sigma|^2 + \epsilon^2$: $\forall i$

$$\langle \varphi(C_i), \tilde{w} \rangle + b = 2\langle C_i, \sigma \rangle - |C_i|^2 - |\sigma|^2 + \epsilon^2 = -|C_i - \sigma|^2 + \epsilon^2,$$

что > 0 для $y_i = 1$ и < 0 для $y_i = 0$. Тогда видно L_d разд.

$\varphi(C)$, где $VCdim(L_d) = d+1$, а $|\varphi(C)| = |C| = d+2$ (от. φ -
инж) $\downarrow$.

Покажем, що $\dim(VC) \geq d+1$. Використаємо n -у канонічну базис $C = \{e_i\}_{i=0}^d$, що y дов. для канонічних просторів: $e_0 = 0$, $e_i = (0, \dots, \underset{i}{1}, \dots, 0) \forall i = 1, d$. Нехай $(y_0, \dots, y_d) \in \{0, 1\}^{d+1}$ - дов. канонічна мінор: $\forall 1 \leq i_1 \leq \dots \leq i_k \leq d, y_{i_1} = \dots = y_{i_k} = 1$, решта $= 0$. Покладемо $v = (0, \underset{i_1}{1}, \dots, \underset{i_2}{1}, \dots, \underset{i_k}{1}, \dots, 0)$ (на інших місцях нулі). Тоді $e_0 - v$ має k (мінус) одиниць, $e_{i_j} - v$, $j = 1, k$ - $k-1$ (мінус) одиниць, і решта $e_{i_j} - v$ - $k+1$ (максимі, 1 од. $i = k-1$). Тому для $\alpha = \sqrt{k}$ маємо $h_{v, \alpha}(e_{i_j}) = 1, \forall j = 1, k, h_{v, \alpha}(e_0) = 1$ і решта нулі, а для $\alpha = \sqrt{k-1}$ - $h_{v, \alpha}(e_{i_j}) = 1, j = 1, k, h_{v, \alpha}(e_0) = 0$ і решта нулі. Це означає, що K розд. C .

2. Для η - невід'ємного на t -му кроці.

$$w_{\eta}^{(t+1)} = w_{\eta}^{(t)} + \eta \sum_i y_i x_i, \text{ де } \eta > 0.$$

При цьому $w_{\underline{z}}^{(0)} = 0 \stackrel{= 2w^{(0)}}{=}$, тому $w_{\underline{z}}^{(1)} = 0 + \sum y_i x_i = \sum w^{(1)}$
 (i тут той самий). Це повториться на $\forall$ кроці:
 $w_{\underline{z}}^{(t)} = \sum w^{(t)}$, тому має ті ж знаки $y_i < w_{\underline{z}}^{(t)}, x_i > =$
 $= \sum y_i < w^{(t)}, x_i >$, отже на t -му кроці у нас не i ;
 $w_{\underline{z}}^{(t+1)} = w_{\underline{z}}^{(t)} + \sum y_i x_i = \sum (w^{(t)} + y_i x_i) = \sum w^{(t+1)}$, отже
 індуктивно маємо потім. Тому $\underline{z}$ -перу. Зупинитися на
 даній S за ті ж T кроків і вийде $w_{\underline{z}}^{(T+1)} = \sum w^{(T+1)}$.
3 Отже, наведено приклад навч. вибірки S з $\mathbb{R}^3 \times \{\pm 1\}$.
 досягається верхня оцінка у $\underline{T}_n$ про нерозривності. По-
 нині, що для цього у дов. $\underline{T}_n$ повинні бути рів-
 ності у наст. місцях:

$\forall t$ перших i , де $y_i < w^{(t)}, x_i > \leq 0$, повинно бути
 так, що $y_i < w^{(t)}, x_i > = 0 \Rightarrow < w^{(t)}, x_i > = 0$ (i
 $y_i < w^{(t)}, x_i > > 0$ для $j < i$); це означає $i = t+1$.

$$R = \sqrt{m+1}$$

$$\forall i: |x_i| = R \quad (R > 1), \text{ мы берем } R: R^2 - 1 \geq m, \text{ наим.}$$

$$\forall i: y_i \langle w^*, x_i \rangle = 1.$$

(и все можно свести для примера к $d = m$).

Будем искать пример задачи вида $x_i = (a, b, y_i)$, где м.ч.
 $a^2 + b^2 + 1 = R^2$. Понемногу $x_1 = (\sqrt{R^2 - 1}, 0, y_1)$ (мы $y_i \in \{\pm 1\}$ -

$\forall$ можно взять все 1). Тогда $w^{(0)} = 0$; $w^{(1)} = 0 + y_1 x_1 =$
 $= (y_1 \sqrt{R^2 - 1}, 0, 1)$. Определим $x_2 = (a, b, y_2)$ так, чтоб

$$0 = \langle w^{(1)}, x_2 \rangle = y_1 a \sqrt{R^2 - 1} + y_2 \Rightarrow a = \frac{-y_1}{y_2 \sqrt{R^2 - 1}}, \quad b = \sqrt{R^2 - 1 - a^2}$$

замечая, что $a^2 = \frac{1}{R^2 - 1} \leq R^2 - 1$ при $R^2 - 1 \geq 1$. Тогда

$$w^{(2)} = w^{(1)} + y_2 x_2 = \left(y_1 \sqrt{R^2 - 1} - \frac{y_1}{\sqrt{R^2 - 1}}, y_2 \sqrt{R^2 - 1 - \frac{1}{R^2 - 1}}, \frac{1}{2} \right),$$

$$|w^{(2)}|^2 = R^2 - 1 + \frac{1}{R^2 - 1} - 2 + R^2 - 1 - \frac{1}{R^2 - 1} + 1 = 2R^2 - 1, \text{ где}$$

и похоже будет за все. Тогда. Индуктивно, пусть

$$w^{(t)} = (\alpha, \beta, t) \quad \text{и} \quad |w^{(t)}|^2 = t R^2 \Rightarrow \alpha^2 + \beta^2 = t(R^2 - t).$$

Значит $x_{t+1} = (a, b, y_{t+1})$ так, что

$$0 = \langle w^{(t)}, x_{t+1} \rangle = \alpha a + \beta b + t y_{t+1} \quad \text{и} \quad a^2 + b^2 = R^2 - 1.$$

Уже заднеги можна зробити дано бак. пер. Π - Φ - Ψ :

$$(a\alpha + b\beta)^2 \leq (a^2 + b^2)(\alpha^2 + \beta^2), \text{ тобто}$$

$$t^2 \leq (R^2 - 1)t + (R^2 - t)$$

$$t \leq R^2 - (t+1)R^2 + t$$

$$R^2 \geq t+1$$

Оскільки $R^2 \geq m+1 \geq t+1$, це дійсно так. Тоді просто проведемо (a, b) довільним $\sqrt{R^2 - 1}$ раз кривим дугам $\frac{-t \text{ y } t+1}{\sqrt{R^2 - 1} \sqrt{t+1} \sqrt{R^2 - t}}$ до ~~до~~

(α, β) . Отримавши

$$w^{(t+1)} = w^{(t)} + y_{t+1} x_{t+1} = (\alpha + y_{t+1} a, \beta + y_{t+1} b, t+1) \text{ довільним.}$$

$$\begin{aligned} |w^{(t+1)}|^2 &= \alpha^2 + \beta^2 + 2y_{t+1}(a\alpha + b\beta) + a^2 + b^2 + (t+1)^2 = t(R^2 - 1) - \\ &- 2t + R^2 - 1 + (t+1)^2 = (t+1)R^2 \end{aligned}$$

Тому дійсно для заданого побудови за m кроків.
(Для цього ще перевіряємо, що $y_{t+1} < w^{(t+1)}$, $x_{t+1} > 0$:

дійсно при $t=0$ це $y_1(y_1(R^2 - 1) + y_1) = R^2 > 0$, при $t=1$:

$$y_2 \left(-\frac{1}{y_2} + \frac{1}{y_2(R^2 - 1)} + y_2 \left(R^2 - 1 - \frac{1}{R^2 - 1} \right) + 2y_2 \right) = \frac{1}{R^2 - 1} (-R^2 + 1 + 1 + (R^2 - 1)^2 - 1 + 2R^2 - 2) =$$

$$1 + R^2 - 1 = R^2 > 0, \quad \forall t:$$

$$y_{t+1}(a\alpha + b\beta + y_{t+1}(a^2 + b^2) + y_{t+1}(t+1)) = \cancel{-x} + R^2 - \cancel{x} + \cancel{x} + \cancel{x} = R^2 > 0$$

1. Значимость оценки на $VCdim(L(B, T))$ с оценкой $\begin{pmatrix} \text{max sign} \\ \text{sign} \\ \underline{O(dT \ln dT)} \end{pmatrix}$
a. Показать, что $\forall B: T \geq 1 \quad VCdim(L(B, T)) \geq VCdim(B)$.

Далее, use max, до $B \subset L(B, T): \forall h \in B \quad h =$
 $= \text{sign}^o(1 \cdot h_1 + 0 \cdot h_2 + \dots + 0 \cdot h_T)$.

b. Значимая минимизация B , что $VCdim(L(B, T)) = VCdim(B) \forall T$.

Use max и минимизация функции:

- $B = \{h \equiv 1, h \equiv -1\}$ має $VCdim = 1$. Типу звару $\forall T$
 $\forall h = \text{sign}^o(\sum_{t=1}^T w_t h_t) \in L(B, T)$ мене мінімає оцінку
не саме знач. б. якщо показав, маємо $VCdim(L(B, T)) = 1$.
- Або раціонали, нехай X кінц., і B - yci $X \rightarrow \{\pm 1\}$.

Могли би нас макс константу вапирність $|X| = n$.
 Але могли її має $L(B, T) \supset B \ (\forall T)$.

Є. Показати, що для класу $\mathcal{H}_{DS}$ нерівно рівень на $\mathbb{R}^d$

$$\lfloor \log_2 d \rfloor \leq VCdim(\mathcal{H}_{DS}) \leq \frac{1}{16} + 2 \log_2 d$$

(зокрема, при $d=1$ $VCdim(\mathcal{H}_{DS1}) = 2$ знач. $\neq$ числ. менше).
 Бо $\uparrow$ не повнотр. $\neq \mathbb{R}^1$.

— Верхня оцінка: згадаймо, що

$$\mathcal{H}_{DS} = \{h_{\vec{j}, \theta, b} : x \mapsto b \cdot \text{sign}(x_{\vec{j}} - \theta)\}_{\vec{j}=1, d, \theta \in \mathbb{R} \cup \{\pm \infty\}, b = \pm 1} =$$

$$= \bigcup_{\vec{j}=1}^d (\mathcal{H}_{\vec{j}} := \{h_{\vec{j}, \theta, b}\}_{\theta, b}). \text{ Класи } \mathcal{H}_{\vec{j}} \text{ "ізоморфні"}$$

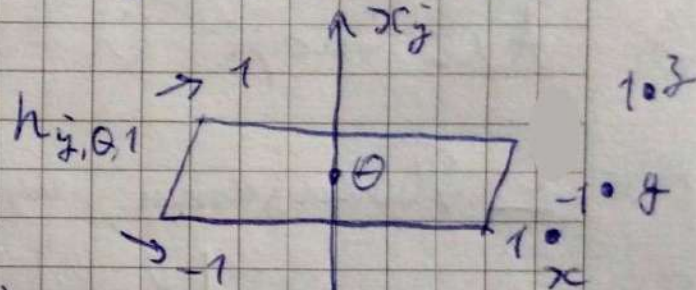
$\mathcal{H}_{DS1}$ і тому має VC -вир. 2 . Дійсно, 2 -ел.

мн. $\{(0, \dots, -\frac{1}{\vec{j}}, \dots, 0), (0, \dots, \frac{1}{\vec{j}}, \dots, 0)\}$ розб. $\{h_{\vec{j}, 0, 1}, h_{\vec{j}, 0, -1},$
 $h_{\vec{j}, 2, 1}, h_{\vec{j}, -2, 1}\} \subset \mathcal{H}_{\vec{j}}$, а $\forall$ 3 -ел.-мна $\{x, y, z\}$ —

ні. Якщо серед x, y, z якісь дві мають однакову
 $\vec{j}$ -коорд., то на них $\forall h_{\vec{j}, \theta, b}$ приймає одн. значення.

Якщо не вони повністю різні, скажімо, $x_j < y_j < z_j$,
то $\nexists h_{j,0,1} : x \mapsto 1, y \mapsto -1, z \mapsto 1$.

Покри за заданого розміру в один
з повн. загалом



$$VCdim(\mathcal{H}_{DS} = \bigcup_{j=1}^J \mathcal{H}_j) \leq 4 \cdot 2 \cdot \log_2(2 \cdot 2) + 2 \log_2 d = 16 + 2 \log_2 d.$$

— мисля оцінка. Позначимо $k := \lfloor \log_2 d \rfloor$, т.ч.,
 $2^k \leq d$. Треба довести $k \leq VCdim(\mathcal{H}_{DS})$, щобмо що
 $\mathcal{H}_{DS}$ розділяє якусь k -ел-мн. γ $\mathbb{R}^d$. Побудуємо її
ан-но мнотини γ приклади про $VCdim = \infty$ для $h(x) = \sin \theta x$.
запам'ятаємо цей ел-мн $\{\pm 1\}^k$ як рядки n -ї, внаслідок
ми аналогічно до двійкових чисел:

$$\begin{array}{ccccccc} & 1 & & & & & k \\ 1 & 1 & \dots & & -1 & -1 & \\ & -1 & \dots & & -1 & 1 & \\ & -1 & \dots & & 1 & -1 & \\ & -1 & \dots & & 1 & 1 & \\ & - & \dots & & - & - & \\ 2^k & 1 & \dots & & 1 & 1 & \end{array}$$

Кожен перел $C = \{x_i\}_{i=1}^k \subset \mathbb{R}^d$ - точки, координат яких є стовпчиками y_i $n \times d$ (допускаємо їх нулями для $d > 2^k$).
 Для $\mathcal{H}_{DS}$ побудуємо C : $\forall$ кожній розстановці
 $\text{sign} \in \{\pm 1\}^k$ просто знайдемо цей рядок y $n \times d$:
 якщо його номер $-j$, то $k_{j,0} = 1$ саме так розст. sign .
 Це й потрібно.

д. Покажемо, що $(\text{знову на } \mathbb{R}^d) \forall T \forall C \dim(L(\mathcal{H}_{DS}, T)) \geq \lfloor \frac{1}{2} T \rfloor \lfloor \log_2 d \rfloor$
 (раніше на лекції показали, що при $d=1$ $\text{VCdim}(L(\mathcal{H}_{DS1}, T)) \geq$
 $\geq T$, використовували лемму про розмір d - y_i).

Знову маємо для $k := \lfloor \log_2 d \rfloor$ $d \geq 2^k$. Тепер потрібно
 побудувати C з ℓ ~~елементами~~ ℓ ел-тів, що розст. $L(\mathcal{H}_{DS}, T)$, де
 $\ell = \lfloor \frac{T}{2} \rfloor$, оскільки $T \geq 2\ell$. Для них, $C = \{x_i\}_{i=1}^k$ з номер.
 пункту розглянемо $\hat{C} := \{\bar{y} x_i\}_{\substack{i=1, k \\ \bar{y}=1, \ell}}$, ~~як~~ можна ~~побудувати~~
 побудувати такі $n \times d$, що поміщені на $1, 2, \dots, \ell$.

Пусть $\{y_m\}_{m=1}^{kl} \in \{\pm 1\}^{kl}$ - одна произм. миним.,
 $\bar{y}_1$ - номер ряда m -го, из $\text{sign. } (y_1, \dots, y_k)$, $\bar{y}_2$ - из
 $\text{sign. } (y_{k+1}, \dots, y_{2k})$ etc. Тогда.

$$h := \text{sign.} \left(h_{\bar{y}_1, 0, 1} + h_{\bar{y}_2, \frac{3}{2}, 1} + h_{\bar{y}_2, -\frac{3}{2}, 1} + h_{\bar{y}_3, \frac{5}{2}, 1} + 2h_{\bar{y}_3, -\frac{5}{2}, 1} + \dots \right)$$

~~Пусть~~ На $x_1, \dots, x_k$ $h_{\bar{y}_1, 0, 1}$ принимает ± 1 у sign. го $\{y_m\}$,
 равенства - адо $y_{ci} = -1$, адо $y_{ci} = 1$, тогда в сумм 0.

На $2x_1, \dots, 2x_k$ $h_{\bar{y}_1, 0, 1}$ принимает также знач. ± 1 ,
 $h_{\bar{y}_2, \frac{3}{2}, 1} + h_{\bar{y}_2, -\frac{3}{2}, 1}$ - знач. ± 2 у sign. го $\{y_m\}$, сумма равенств
 равн. 0. Тогда знак будет таким, как у $\{y_m\}$.

На $3x_1, \dots, 3x_k$ сумма первых трех принимает знач. sign.
 -3 го 3 , четвертое и пятое - ± 4 у sign. го $\{y_m\}$, сумма
 равенств 0 etc.

Ост. $h \in L(\mathcal{H}_{DS}, 2l-1) \subset L(\mathcal{H}_{DS}, T)$, где g - произм.

2. Нехай $\mathcal{X}$ задані бір. клас. $\forall x: y = \{0, 1\}$ мінка
обирається випадково: $y = 0$ з $P = \frac{1}{2}$ і $y = 1$ з $P = \frac{1}{2}$
(тобто розм. $\oplus$ на $\mathcal{X} \times \mathcal{Y}$ мінки). Розр. алгоритми
ML A , що на $S = \{(x_i, y_i)\}_{i=1}^m$ повертає:

$$h_S = A(S) := \left(\sum_{i=1}^m y_i \right) \bmod 2, \text{ тобто парність } \{y_i\}$$

П.ч. або $h_S = 0$ або $h_S = 1$ - інтегру постійні

Знайти $L_{\oplus}(h_S)$ - крос-валідаційні помилки h_S
для LOO-метода, порівняти їх.

Оскільки половина мінки мають мінку 0, а половина - 1,
 $\forall$ постійна h_S має помилку $\frac{1}{2}$.

Нехай S має парну кількість мінки 1. Тоді $h_S = 0$.
 $\forall i = \overline{1, m}$ LOO-алгоритм відкидає (x_i, y_i) і
надає A на $S'_i = S \setminus \{(x_i, y_i)\}$.

- якщо $y_i = 0$, то S_i має парну кількість 1, тому $h_i = A(S) \equiv 0$. Тоді, $L_{\{x_i, y_i\}}^0(h_i) = 0$

- якщо $y_i = 1$, то S_i має непарну кількість 1, тому $h_i = A(S) \equiv 1$, і $L_{\{x_i, y_i\}}^1(h_i) = 0$

Тоді, крок - вал. помилка $L_S(A) = 0$

Уже ми бачимо, що S має непарну кількість 1. Тоді $h_S = 1$, і $\forall i = \overline{1, m}$:

- якщо $y_i = 0$, то S_i має непарну кількість 1, тому $h_i \equiv 1$, $L_{\{x_i, y_i\}}^0(h_i) = 1$.

- якщо $y_i = 1$, то S_i має парну кількість 1, тому $h_i \equiv 0$, $L_{\{x_i, y_i\}}^1(h_i) = 1$.

Тоді, $L_S(A) = 1$.

У $\forall$ разі, $|L_\emptyset(h_S) - L_S(A)| = \frac{1}{2}$.