Lecture 17

RECAP

Online Learning

At every time step t

→ learner receives an input $x_t \in X$.
→ makes prediction $p_t \in Y$.
→ see true label $y_t \in Y$. suffer loss $\ell(p_t, y_t)$.

Think $Y = \{0,1\}$

$$\ell(p_t, y_t) = \mathbb{1}\{p_t \neq y_t\}.$$

Realizability

Def (Mistake bound model) Let $\mathcal{H}$ be a hypothesis class and A be a online learning algo. Given any sequence $S = (x_1, h^*(x_1)), \dots, (x_T, h^*(x_T))$ of T labelled data points where $h^* \in \mathcal{H}$, let $M_A(S)$ be the # mistakes A makes on the sequence S. We denote by $M_A(\mathcal{H})$ to the supremum of $M_A(S)$ over all possible S.

If there exists an algorithm A that satisfies a <u>mistake bound</u> of the form $M_A(\mathcal{H}) \leq B < \infty$, we say $\mathcal{H}$ is online learnable in the mistake bound model.

Alg: Consistent

① Initialize $V_1 = \mathcal{H}$

② for $t = 1, \dots, T$

receive x_t

Choose any $h \in V_t$

predict $p_t = h(x_t)$

receive $y_t = h^*(x_t)$, loss $\mathbb{1}(p_t \neq y_t)$

update $V_{t+1} = \{ h \in V_t : h(x_t) = y_t \}$.

Proposition: Let $\mathcal{H}$ be a finite hypothesis class. The above algorithm gets a mistake bound

$$M_{\text{consistent}}(\mathcal{H}) \leq |\mathcal{H}| - 1.$$

Alg: Halving

① Initialize $V_1 = \mathcal{H}$

② for $t = 1, \dots, T$

receive x_t

predict $p_t = \underset{r \in \{0,1\}}{\operatorname{argmax}} \left| \{ h \in V_t : h(x_t) = r \} \right|$

(if tie, $p_t = 1$)

receive $y_t = h^*(x_t)$, loss $\mathbb{1}(p_t \neq y_t)$

update $V_{t+1} = \{ h \in V_t : h(x_t) = y_t \}$.

Proposition Let $\mathcal{H}$ be a finite hypothesis class. Then halving algorithm satisfies the mistake bound $M_{Halving}(\mathcal{H}) \leq \log_2(|\mathcal{H}|)$.

Proof Whenever the algo. errors, $|V_{t+1}| \leq \frac{|V_t|}{2}$.

If M is # mistakes

$$|V_{T+1}| \leq |\mathcal{H}| \, 2^{-M}$$

as $|V_{T+1}| \geq 1$

$$\Rightarrow \quad M \leq \log_2(|\mathcal{H}|).$$

Littlestone dimension

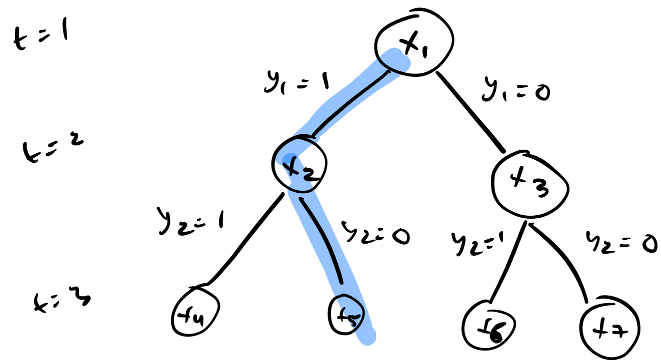
Idea: view online learning as a 2-player game b/w learner & the environment.

At time t:

environment picks x_t
learner picks p_t
enviroment picks $y_t = 1 - p_t$

Q) How to choose x_t to get the learner to make maximum # mistakes, while ensuring realizability?

Strategy for the environment:



Def ($\mathcal{H}$ shattered tree)

A shattered tree of depth d is a sequence of instances $x_1, x_2, \ldots, x_{2^d - 1} \in \mathcal{X}$ s.t. for every path from the root to a leaf $\exists\, h \in \mathcal{H}$ which realizes all the labels along this path.

Def (Littlestone dimension) Littlestone dimension of hypothesis class $\mathcal{H}$ ($\mathrm{Ldim}(\mathcal{H})$) is the maximum T s.t. $\exists$ exists a tree of depth T shattered by $\mathcal{H}$.

Lemma For any online learning algorithm A,

$$M_A(\mathcal{H}) \geq \mathrm{Ldim}(\mathcal{H}).$$

Examples

① Let $\mathcal{H}$ be a finite hypothesis class. Then $\mathrm{Ldim}(\mathcal{H}) \le \log_2(|\mathcal{H}|)$.

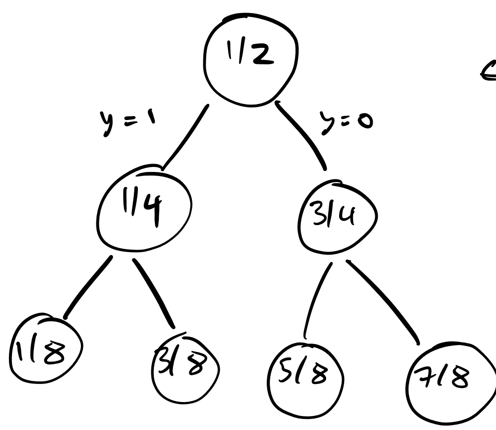
$\rightarrow$ Any tree which is shattered by $\mathcal{H}$ must have $|\mathcal{H}| \geq$ # leaves in the tree.

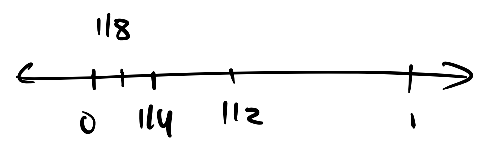
② $\mathcal{X} = [0,1]$

$\mathcal{H} = \{ x \mapsto \mathbb{1}(x < a),\ a \in [0,1] \}$

(thresholds on $[0,1]$)

$\mathrm{Ldim}(\mathcal{H}) = \infty$





Corollary: Cannot learn thresholds in the online learning model.

Lemma: There exists an algo A with $M_A(\mathcal{H}) \le \mathrm{Ldim}(\mathcal{H})$.

Proof

Algo: Standard optimal algorithm(SOA)

① Initialize $V_1 = \mathcal{H}$

② for $t = 1, \dots, T$

receive x_t

for $r \in \{0,1\}$ let $V_t^{(r)} = \{h \in V_t : h(x_t) = r\}$

predict $p_t = \arg\max_{r \in \{0,1\}} \text{Ldim}(V_t^{(r)})$

receive $y_t = h^*(x_t)$, loss $\mathbb{1}(p_t \neq y_t)$

update $V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$.

Claim Whenever algo. makes a mistake $\text{Ldim}(V_{t+1}) = \text{Ldim}(V_t) - 1$.

Proof by contradiction If $\text{Ldim}(V_{t+1}) = \text{Ldim}(V_t)$ though algorithm has made mistake.

w.l.o.g. algo predicts 0. Then $y_t = 1$ since algo. made a mistake.

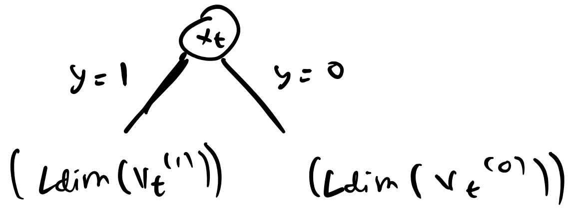
$$\text{Ldim}(V_{t+1}) = \text{Ldim}(V_t^{(1)}) = \text{Ldim}(V_t)$$

also $$\text{Ldim}(V_t^{(0)}) \geq \text{Ldim}(V_t^{(1)}) = \text{Ldim}(V_t)$$

$$\text{Ldim}(V_t^{(0)}) \leq \text{Ldim}(V_t)$$

$$\therefore \quad \text{Ldim}(V_t^{(0)}) = \text{Ldim}(V_t)$$

But then $\text{Ldim}(V_t) = \text{Ldim}(V_t^{(0)}) + 1$

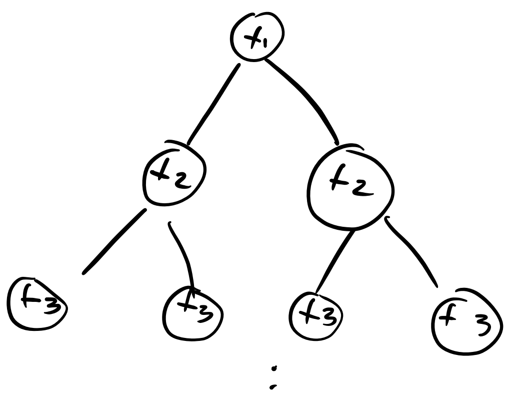


$\Rightarrow \le$

Thm (Compare to $\mathrm{VCdim}(\mathcal{H})$) for any class $\mathcal{H}$, $\mathrm{VCdim}(\mathcal{H}) \le \mathrm{Ldim}(\mathcal{H})$. Further, the gap can be arbitrarily large.

Proof Suppose $\mathrm{VCdim}(\mathcal{H}) = d$.

Let $x_1, \dots, x_d$ be shattered set



$\mathrm{VCdim}(\mathcal{H}) \le \mathrm{Ldim}(\mathcal{H})$.

The gap can be arbitrarily large just because of thresholds.

Aside: Some recent work shows that every $\mathcal{H}$ with finite $\mathrm{Ldim}(\mathcal{H})$ can be learnt privately (and vice versa).

Online learning in the unrealizable case.

Def (Regret) The regret of an algo A relative to hypothesis h when run on a sequence of T examples is:

$$\text{Regret}_A(h, T) = \sup_{(x_1,y_1),\dots,(x_T,y_T)} \left[\sum_{t=1}^{T} |p_t - y_t| - \sum_{t=1}^{T} |h(x_t) - y_t| \right]$$

The regret of A relative to a hypothesis class $\mathcal{H}$ is:

$$\text{Regret}_A(\mathcal{H}, T) = \sup_{h \in \mathcal{H}} \text{Regret}_A(h, t).$$

* If sequence is realizable, this is same as mistake bound.

Can we get sublinear regret ($o(T)$)?

Unfortunately, no.

Consider $\mathcal{H} = \{h_0, h_1\}$

h_0: always predicts 0

h_1: always predicts 1

An adversary can force the # mistakes made by algo to T.

But, the best predictor in hindsight is the majority of $y_1 \dots y_T$ which makes $\leq T/2$ mistakes.

$\Rightarrow$ Regret $\geq T/2$.

To get around this, allow randomized algorithms, environment decides y_t before random coins are flipped.

Setup

At every time step t,

- $\rightarrow$ learner receives $x_t \in \mathcal{X}$.
- $\rightarrow$ learner decides $p_t \in [0,1]$, prob. of label being 1.
- $\rightarrow$ environment "decides" true label $y_t \in \{0,1\}$.
- $\rightarrow$ learner outputs $\hat{y_t} = \begin{cases} 1, & \text{w.p. } p_t \\ 0, & \text{w.p. } 1-p_t \end{cases}$

- $\rightarrow$ Expected loss at time t,

$$\mathbb{P}(\hat{y_t} \neq y_t) = \begin{cases} p_t, & \text{if } y_t = 0 \\ 1-p_t, & \text{if } y_t = 1 \end{cases}$$

$$= |p_t - y_t|$$

$$\text{Regret}_A(\mathcal{H}, T) = \sup_{h \in \mathcal{H}} \sup_{(x_1,y_1)\dots(x_T,y_T)} \left[\sum_{t=1}^{T} |p_t - y_t| - \sum_{t=1}^{T} |h(x_t) - y_t| \right]$$

Can we get sublinear regret? Yes, using Weighted Majority.

Setting: Prediction with "expert advice".
At every time t, learner has to choose one among d experts to predict based on.
We then see true label, and the loss each expert has on that time step, can use that to refine future predictions.

Q) How well can we do compared to the best expert in hindsight?

Alg: Weighted Majority (Multiplicative Weights / Hedge)

1. initialize $w^{(1)} = (1, \dots, 1)$ (d-dimensional)
2. for $t = 1 \dots T$
3. set $\tilde{w}^{(t)} = w^{(t)} / Z_t$ where $Z_t = \sum_i w_i^{(t)}$
4. Choose expert i at random according to $P[i] = \tilde{w}_i^{(t)}$.
5. Receive costs of all experts $v_t \in [0,1]^d$
6. Pay expected cost: $\langle \tilde{w}^{(t)}, v_t \rangle$
7. update: $\forall i$: $w_i^{(t+1)} = w_i^{(t)} e^{-\eta v_{t,i}}$

Thm Assuming $T > 2\log(d)$, the Weighted-Majority algorithm enjoys the bound

$$\sum_{t=1}^{T} \langle w^{(t)}, v_t \rangle - \min_{i \in [d]} \sum_{t=1}^{T} v_{t,i} \le \sqrt{2\log(d) T}$$

Proof Similar to AdaBoost.

$$\log\left(\frac{Z_{t+1}}{Z_t}\right) = \log\left(\frac{\sum_i w_i^{(t)} \exp(-\eta v_{t,i})}{Z_t}\right)$$

$$= \log\left(\sum_i \tilde{w}_i^{(t)} \exp(-\eta v_{t,i})\right)$$

Using ① $e^{-a} \le 1 - a + \frac{a^2}{2} \quad \forall a \in (0,1)$

② $\sum \tilde{w}_i^{(t)} = 1$

③ $\log(1-b) \le -b \quad (b \le 1)$

$$\log\left(\frac{Z_{t+1}}{Z_t}\right) \le \log\left(\sum_i \tilde{w}_i^{(t)}\left(1 - \eta v_{t,i} + \frac{\eta^2}{2} v_{t,i}^2\right)\right) \quad ①$$

$$= \log\left(1 - \sum_i \tilde{w}_i^{(t)}\left(\eta v_{t,i} - \frac{\eta^2}{2} v_{t,i}^2\right)\right) \quad ②$$

$$\le -\sum_i \tilde{w}_i^{(t)}\left(\eta v_{t,i} - \frac{\eta^2}{2} v_{t,i}^2\right) \quad ③$$

$$= -\eta \langle \tilde{w}^{(t)}, v_t \rangle + \frac{\eta^2}{2} \sum_i \tilde{w}_i^{(t)} v_{t,i}^2$$

$$\le -\eta \langle \tilde{w}^{(t)}, v_t \rangle + \frac{\eta^2}{2}$$

$$\sum_{t=1}^{T} \log\left(\frac{Z_{t+1}}{Z_t}\right) = \log(Z_{T+1}) - \log(Z_1)$$

$$\le -\eta \sum_{t=1}^{T} \langle \tilde{w}^{(t)}, v_t \rangle + T\frac{\eta^2}{2}$$

$$\therefore \quad \eta \sum_{t=1}^{T} \langle \tilde{w}^{(t)}, v_t \rangle \leq \log(Z_1) - \log(Z_{T+1}) + \frac{T\eta^2}{2}$$

$\log(Z_1) = \log(d)$

Note that $w_i^{(T+1)} = \exp\left(-\eta \sum_t v_{t,i}\right)$

$$\log(Z_{T+1}) = \log\left(\sum_i \exp\left(-\eta \sum_t v_{t,i}\right)\right)$$
$$\geq \log\left(\max_i \exp\left(-\eta \sum_t v_{t,i}\right)\right)$$
$$= -\eta \min_i \sum_t v_{t,i}$$

$$\therefore \quad \eta \sum_{t=1}^{T} \langle \tilde{w}^{(t)}, v_t \rangle \leq \log(d) + \eta \min_i \sum_t v_{t,i} + \frac{T\eta^2}{2}$$

$$\Rightarrow \quad \sum_{t=1}^{T} \langle \tilde{w}^{(t)}, v_t \rangle - \min_i \sum_t v_{t,i} \leq \frac{\log d}{\eta} + \frac{T\eta}{2}$$

Choose $\eta = \sqrt{\frac{2\log(d)}{T}}$, which gives

$$\leq \sqrt{2\log(d)\, T}$$

■

Regret bound for online learning

Thm: Let $\mathcal{H} = \{h_1 \ldots h_d\}$ be finite hypothesis class. Then Weighted-Majority achieves

$$\sum_{t=1}^{T} |p_t - y_t| - \min_{h \in \mathcal{H}} \sum_{t=1}^{T} |h(x_t) - y_t| \leq \sqrt{2\log(|\mathcal{H}|)\, T}.$$

Proof Each experts h_i predicts $h_i(x_t)$ on example x_t.

loss $v_{t,i} = |h_i(x_t) - y_t|$.

Predictions of weighted Majority $p_t = \sum_i w_i^{(t)} h_i(x_t)$.

Expected loss

$$|p_t - y_t| = \Big| \sum_{i=1}^{d} w_i^{(t)} h_i(x_t) - y_t \Big|$$

$$= \sum_{i=1}^{d} \Big| w_i^{(t)} \big(h_i(x_t) - y_t \big) \Big|$$

$$= \langle w^{(t)}, v_t \rangle$$

Then we use the regret bound for weighted-Majority to bound $\langle w^{(t)}, v_t \rangle$.

∎