Week 3

Lecture 12

Full information Regret

$$R(n, \mathcal{H}) = \sup_{h} \sup_{(x,y)} \left[\sum_{t=1}^{n} |\hat{y}_t - y_t| - \sum (h(x_t) - y_t) \right] \quad \text{——— (1)}$$

After prediction, when the information comes we get to know the output for all hypothesis in $\mathcal{H}$.

Eq (1) holds for any sequence, any hypothesis

The bound for Regret from eq (1) $\leq \sqrt{2 \log|\mathcal{H}| n}$

If we write eq. (1) as:

$$R(n, \mathcal{H}) = \sup_{(x,y)} \sup_{h} \left[\sum_{t=1}^{n} |\hat{y}_t - y_t| - \sum (h(x_t) - y_t \right. \quad \text{——— (1a)}$$

The bound still holds but the regret now is dependent on the sequence.

Eq 1a can further be written as:

$$R(n, \mathcal{H}) = \sup_{(x,y)} \mathbb{E}\left[\sum_{t=1}^{n} |\hat{y}_t - y_t| - \inf_h \sum (h(x_t) - y_t)\right] \quad \text{— 1b}$$

because $\hat{y}_t$ is <u>not</u> deterministic

This was the case for Expert Advice setting with weighted majority.

<u>Bandit Setting</u>

When we choose an arm, we incur loss, however we get information only about the arm we played.

<u>Ex</u>: Casino with various slot machines where we get to play only one machine at one time point

<u>Full information ex</u>:

Share market : we get to choose one share to play. but at the end of the day we can get information about all the shares

---

K-actions $[K] = \{1, 2, \ldots K\}$

In round t

Environment assings loss vector $x_t \in [0,1]^K$

Player selects action $I_t \in [K]$

Observes loss from $I_t = x_{tI_t}$

History till round t

$$H_t = \{ I_1, x_{1I_1}, I_2, x_{2I_2}, \dots I_{t-1}, x_{t-1 I_{t-1}}$$

Policy π :

Choose an action in round t given H_t

$$\pi_t(i) = \Pr\{ I_t = i \mid H_t \}$$

Expected
Regret would then be :

$$\mathbb{E}\left[R(n,\pi) \right] = \mathbb{E} \sum_{t=1}^{n} x_{tI_t} - \mathbb{E}\left[\min_i \sum_{t=1}^{n} x_{ti} \right] \qquad \text{—— } 2$$

↑ because $R(n,\pi)$ is random as I_t is random

Since $R(n,\pi)$ is a random quantity hence we will bound the $\mathbb{E}[R(n,\pi)]$

Regret bound :

$$\sup_{x_1, x_2, \dots x_t} \mathbb{E}\left[R(n,\pi) \right]$$
