Online Learning with Expert Advice.

- Decision Maker chooses at time t $\underline{x}_t \in \Delta_n$, $(\Delta_n = \{\underline{x} \in \mathbb{R}^n \mid \underline{x} \geq 0, \sum_i x_i = 1\})$ a probability distribution over n options.
- Environment / Adversary reveals a loss vector $\underline{l}_t \in [0,1]^n$.
- Loss of DM at time t is $\underline{x}_t \cdot \underline{l}_t$.

$$\text{Regret} = \sum_{t=1}^{T} \underline{l}_t \cdot \underline{x}_t - \min_{\underline{x} \in \Delta^n} \sum_{t=1}^{T} \underline{x} \cdot \underline{l}_t$$

$$= \sum_{t=1}^{T} \underline{l}_t \cdot \underline{x}_t - \min_{i \in [n]} \sum_{t=1}^{T} l_{t,i}$$

Multiplicative Weight Update Algorithm:

$w_1 = (1, \ldots, 1) \in \mathbb{R}^n$

For each $t = 1, \ldots, T$

- $\underline{x}_t = \frac{1}{Z_t} \cdot \underline{w}_t$ where $Z_t = \sum_i w_{t,i}$. $(*)$.
- $w_{t+1,i} = w_{t,i} \cdot \exp(-\eta\, l_{t,i})$.

Theorem: Regret (MWUA) $= O(\sqrt{T \log n})$.

Interpretation of $(*)$.

$$\boxed{\min_{\underline{x} \in \Delta_n} \sum_{s=1}^{t-1} \underline{x} \cdot \underline{l}_s - \underbrace{\frac{1}{\eta} H(\underline{x})}}$$

$H(\underline{x}) = -\sum_i x_i \ln x_i$

$\underline{x} \geq 0$

$\sum_i x_i = 1$

$$x_i \propto \exp\left(-\eta \sum_{s=1}^{t-1} l_{s,i}\right).$$

$$\text{argsoftmin}\left(+\eta \sum_{s=1}^{t-1} \underline{l}_s\right).$$

$\underline{x} = (x_1, \ldots, x_n)$

$p_i \propto \exp(-x_i)$

Boosting:

Let $\{x_1, \ldots, x_m\}$ as training data, with labels $\{y_1, \ldots, y_m\}$.

DM: Picks distribution $\underline{p}_t$ over $\{1, \ldots, m\}$. (using MWUA).

Weak Learner: guarantees h_t. s.t. $err(h_t; \underline{p}_t) \le \frac{1}{2} - \gamma$.

$$l_{t,i} = \begin{cases} 1 & \text{if } h_t(x_i) = y_i \\ 0 & \text{if } h_t(x_i) \ne y_i \end{cases}.$$

$$\underline{l}_t \cdot \underline{p}_t = 1 - err(h_t; p_t) \ge \frac{1}{2} + \gamma$$

If we play for T rounds, then

$$Loss(DM) = \sum_{t=1}^{T} \underline{l}_t \cdot p_t \ge T \cdot \left(\frac{1}{2} + \gamma\right).$$

MWUA Theorem: Regret $= O(\sqrt{T \log m})$

$$Regret[i] = \sum_{t=1}^{T} \underline{l}_t \cdot \underline{p}_t - \sum_{t=1}^{T} l_{t,i} \le 2\sqrt{T \log m}$$

$$\sum_{t=1}^{T} l_{t,i} \ge T\left(\frac{1}{2} + \gamma\right) - 2\sqrt{T \log m} > T/2$$

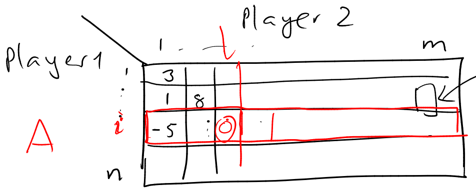
$$= \frac{T}{2} + \underbrace{\gamma T - 2\sqrt{T \log m}}$$ for sufficiently large T.

For every i, a majority of the hypotheses from the set $\{h_1, \ldots, h_T\}$ correctly classify it.

$\cdot H(x) = MAJ(h_1(x), \ldots, h_T(x))$ classifies all m examples correctly.

von Neumann's Min-max Theorem :

Two person zero-sum games.



Payoff to player 1 = − Payoff to player 2

Player 1 goes first: what would player 2 do?

Player 1 picks i, Player 2 picks j that minimizes A_{ij}

$$\underbrace{\max_i \min_j A_{ij}}_{\text{Payoff to player 1 if they go first}} \leq \underbrace{\min_j \max_i A_{ij}}_{\text{Payoff to player 1 if they go second.}}$$

Player 1 picks probability distribution over $\{1, \ldots, n\}$. σ_1
Player 2 " " " over $\{1, \ldots, m\}$. σ_2

$$\text{Payoff to Player 1} = \mathbb{E}_{\substack{i \sim \sigma_1 \\ j \sim \sigma_2}} A_{ij} = \sum_{ij} (\sigma_1)_i (\sigma_2)_j A_{ij}$$

$$\underbrace{\max_{\sigma_1} \min_{\sigma_2} \mathbb{E}_{\substack{i \sim \sigma_1 \\ j \sim \sigma_2}} [A_{ij}]}_{v_1} = \underbrace{\min_{\sigma_2} \max_{\sigma_1} \mathbb{E}_{\substack{i \sim \sigma_1 \\ j \sim \sigma_2}} [A_{ij}]}_{v_2}.$$ (von Neumann's Min-max)

Payoff instead of losses.

Regret = Best payoff in hindsight − Payoff of online algorithm $\leq O(\sqrt{T \log n})$.

Player 1 uses MWUA to pick σ_1^t over $\{1, \ldots, n\}$.

Player 2 picks a strategy that's optimal for σ_1^t.

$$\sigma_2^t \in \arg\min \underbrace{\mathbb{E}_{\substack{i \sim \sigma_1^t \\ j \in \sigma_2^t}} A_{ij}}, \qquad P_{t,i} = \mathbb{E}_{j \sim \sigma_2^t} A_{ij}$$

(i) $\mathbb{E}_{\substack{i \sim \sigma_1^t \\ j \sim \sigma_2^t}} A_{ij} \leq v_1.$

$$\text{Total payoff} = \sum_{t=1}^{T} \mathbb{E}_{\substack{i \sim \sigma_1^t \\ j \sim \sigma_2^t}} A_{ij} \leq T v_1$$

(ii) $\frac{1}{T} \max_\sigma \sum_{t=1}^{T} \sigma \cdot P_t$

$\frac{1}{T} \max_\sigma \sigma \cdot \underbrace{\sum_{t=1}^{T} P_t}$

$\geq v_2$

$\max_\sigma \sum_{t=1}^{T} \sigma \cdot P_t \geq T v_2.$

Player 2 plays $\frac{1}{T} \sum \sigma_2^t$

Regret $\geq T\vartheta_2 - T\vartheta_1$

Regret $\leq O(\sqrt{T \log n})$.

$$\vartheta_2 - \vartheta_1 \leq 2\sqrt{\frac{\log n}{T}} \quad \forall T.$$

$$\Rightarrow \vartheta_2 = \vartheta_1.$$

$\square$.