CSC 581A. Incentives and Machine Learning. Presentation notes. No-Regret and Incentive-Compatible Online Learning.

Kirill Abrosimov and Austin Shouli

March 2026

1 Online Learning: Recap

1.1 Prediction with Expert Advice

Let's consider a basic Online Learning Problem, Prediction with Expert Advice.

We have:

- K experts
- Action space $\mathcal{A}$
- Outcome spave $\mathcal{Y}$
- Loss Function $l : \mathcal{A} \to \mathcal{Y}$

Protocol: At each round $t = 1, 2, \ldots, T$:

- Each expert $i \in [K]$ reveals their advice $f_{j,t}$
- Learner (Algorithm) plays an action $a_t \in \mathcal{A}$
- Nature plays and reveals an outcome $y_t \in \mathcal{Y}$ [Could be Stochastic, or Adversarial]
- Learner suffers loss $\hat{l}_t = l(a_t, y_t)$ [based on what they played, and the actual outcome of nature]

Regret in round n:

$$R_n = \hat{l}_t - \min_{i \in [K]} \ell_{i,t}.$$

Cumulative regret:

$$R_T = \sum_{t=1}^{T} L_t - \min_{i \in [K]} \sum_{t=1}^{T} \ell_{i,t}.$$

Goal:

- We want to design an algorithm that achieves sub-linear regret
- This means that our algorithm will eventually converge to perform as well as the best expert
- Sub-linear regret is equivalent to No-regret

1.2 Exponential Weighted Average Forecaster

Let's look at an algorithm for PEA that achieves sub-linear regret:

This algorithm is known as:

- Exponential Weighted Average Forecaster (EWA)
- The Hedge Algorithm
- Multiplicative Weights Update (MWU)

[There are subtle differences but for our case they are functionally equivalent]

Algorithm:

Each forecaster / expert is assigned an initial weight $w_{i,t} = 1$

In each round $t = 1, 2, \ldots, T$:

- Each expert i makes a prediction $f_{i,t}$
- Probability distribution p among experts, such that $p_{j,t} = \frac{w_{j,t}}{W_t}$, $\quad W_t = \sum_{i=1}^{K} w_t^{(i)}$
- Learner makes a prediction $a_t = \sum_{j=1}^{K} p_{j,t} f_{j,t}$
- Learner incurs loss $\hat{l_t}$
- Experts incur loss $l_{i,t}$
- Each expert's loss is updated such that:

$$w_{i,t+1} = \begin{cases} w_{i,t}, \text{if } f_i, t = y_t & \text{[ie. experts prediction was correct]} \\ e^{-\eta l_{i,t}}, \text{if } f_i, t \neq y_t & \text{[ie. experts prediction was incorrect]} \end{cases}$$

1.3 EWA Regret Analysis

Define the potential:

$$W_t = \sum_{i=1}^{K} w_{i,t}.$$

Initially,

$$W_1 = K.$$

Upper bound. We have:

$$W_{t+1} = \sum_i w_{i,t} e^{-\eta \ell_{i,t}} = W_t \sum_i \pi_{i,t} e^{-\eta \ell_{i,t}}.$$

Using $e^{-x} \leq 1 - x + x^2/2$ for $x \in [0, 1]$,

$$\sum_i \pi_{i,t} e^{-\eta \ell_{i,t}} \leq 1 - \eta L_t + \frac{\eta^2}{2}.$$

Thus,

$$W_{t+1} \leq W_t \exp\left(-\eta L_t + \frac{\eta^2}{2}\right).$$

Iterating,

$$W_{T+1} \leq K \exp\left(-\eta \sum_{t=1}^{T} L_t + \frac{\eta^2 T}{2}\right).$$

Lower bound. For any expert i:

$$w_{i,T+1} = \exp\left(-\eta \sum_{t=1}^{T} \ell_{i,t}\right).$$

Since $W_{T+1} \geq w_{i,T+1}$,

$$W_{T+1} \geq \exp\left(-\eta \sum_{t=1}^{T} \ell_{i,t}\right).$$

Combine bounds.

$$\exp\left(-\eta \sum_{t=1}^{T} \ell_{i,t}\right) \leq K \exp\left(-\eta \sum_{t=1}^{T} L_t + \frac{\eta^2 T}{2}\right).$$

Taking logarithms:

$$\sum_{t=1}^{T} L_t - \sum_{t=1}^{T} \ell_{i,t} \leq \frac{\ln K}{\eta} + \frac{\eta T}{2}.$$

Choosing

$$\eta = \sqrt{\frac{2 \ln K}{T}},$$

we obtain

$$R_T \leq \sqrt{2T \ln K}.$$

2 Wagering

2.1 Prove WSWM is incentive compatible

Proof

$$\Gamma_i^{WSWM}(\vec{p}, \vec{w}, r) = w^{(i)} \cdot (S(p_i, r) - \sum_{j \in [K]} \frac{w_i}{W} \cdot S(p_j, r))$$

Consider the term

$$\sum_{j \in [K]} \frac{w_j}{W} \cdot S(p_j, r) = \frac{w_i}{W} S(p_i, r) + \sum_{J \neq i} \frac{w_j}{W} S(p_j, r)$$

[1st term is expert i's weighted-score, second term is all others]

$$\text{Let } \sum_{J \neq i} \frac{w_j}{W} S(p_j, r) = C(r) \quad \text{[Second term above]}$$

Substituting back into the wagering mechanism Γ

$$\begin{aligned}
\Gamma_i^{WSWM}(\vec{p}, \vec{w}, r) &= w_i \cdot (S(r,p) - (\frac{w_i}{W} S(r,p) + C(r)) \\
&= w_i S(r,p) - w_i(\frac{w_i}{W} S(r,p) - C(r)) \\
&= S(r,p) w_i (1 - \frac{w_i}{W}) - w_i C(r) \\
\mathbb{E}[\Gamma_i^{WSWM}(\vec{p}, \vec{w}, r)] &= \mathbb{E}[S(r,p) w_i (1 - \frac{w_i}{W}) - w_i C(r)] \\
&= \mathbb{E}[S(r,p) w_i (1 - \frac{w_i}{W})] - \mathbb{E}[w_i C(r)] \\
&= \mathbb{E}[S(r,p)] w_i (1 - \frac{w_i}{W}) - w_i \mathbb{E}[C(r)] \\
&= \mathbb{E}[S(r,p)] w_i (1 - \frac{w_i}{W}) \quad \text{[Drop } \mathbb{E}[C(r)] \text{, it is not dependent on } i \text{]]} \\
&= \mathbb{E}[S(r,p)] \qquad \text{[Drop } w \text{ terms, they are a constant } \geq 0 \text{]}
\end{aligned}$$

We have shown that the expectation of the WSWM wagerin gmechanism Γ is directly dependent on the expectation of a proper scoring rule. We have already shown in class that proper scoring rules are incentive compatible, therefore the WSWM mechanism Γ is incentive compatible.

3 WSU algorithm

$$\pi_{t+1} = \Gamma^{WSWM}(\vec{p_t}, \vec{\pi}_t, r_t)$$

$$\pi_{t+1}^{(i)} = \eta \cdot \Gamma_i^{WSWM}(\vec{p_t}, \vec{\pi}_t, r_t) + (1-\eta) \cdot \pi_t^{(i)}$$

Initialization of $\pi_1^{(i)} = \frac{1}{K}, \forall i$.

$$L_t^{(i)} = l_t^{(i)} - \sum_{j=1}^{K} \pi_t^{(j)} \cdot l_t^{(j)}$$

Updating rule:

$$\pi_{t+1}^{(i)} = \eta \cdot \pi_t^{(i)} \cdot (1 - l_t^{(i)} + \sum_{j=1}^{K} \pi_t^{(j)} \cdot l_t^{(j)}) + (1-\eta) \cdot \pi_t^{(i)} = \eta \cdot \pi_t^{(i)} \cdot (1 - L_t^{(i)}) + (1-\eta) \cdot \pi_t^{(i)} =$$

$$= \pi_t^{(i)} \cdot (\eta \cdot (1 - L_T^{(i)}) - (1-\eta)) = \pi_t^{(i)} \cdot (1 - \eta \cdot L_T^{(i)})$$

3.1 Regret analysis

Theorem: WSU is incentive-compatible and for step size $\eta = \sqrt{\frac{\ln K}{T}}$ yields regret $R \leq 2 \cdot \sqrt{T \cdot \ln K}$.

Proof. For i^* - the best expert.

$$1 \geq \pi_{T+1}^{(i^*)} = \pi_t^{(i^*)} \cdot (1 - \eta \cdot L_T^{(i^*)}) = \pi_1^{(i^*)} \prod_{t=1}^{T} (1 - \eta \cdot L_t^{(i^*)}) = \frac{1}{K} \prod_{t=1}^{T} (1 - \eta \cdot L_t^{(i^*)})$$

Taking the logarithm

$$0 \geq -\ln K + \sum_{t=1}^{T} \ln(1 - \eta \cdot L_t^{(i^*)})$$

We know that $\ln(1-x) \geq -x - x^2; \forall x \leq \frac{1}{2}$

$$0 \geq -\ln K + \sum_{t=1}^{T} (-\eta \cdot L_t^{(i^*)} - \eta^2 \cdot L_t^{2(i^*)}))$$

We can rewrite:

$$-\eta \sum_{t=1}^{T} L_t^{(i^*)} \leq \ln K + \eta^2 \sum_{t=1}^{T} L_t^{2(i^*)} | : \eta$$

$$-\sum_{t=1}^{T} L_t^{(i^*)} \leq \frac{\ln K}{\eta} + \eta \sum_{t=1}^{T} L_t^{2(i^*)}$$

We know that:

$$L_t^{(i^*)} = l_t^{(i^*)} - \sum_{j=1}^{K} \pi_t^{(j)} \cdot l_t^{(j)}$$

And Regret:

$$R = \sum_{t=1}^{T} \sum_{j=1}^{K} \pi_t^{(j)} \cdot l_t^{(j)} - \sum_{t=1}^{T} l_t^{(i^*)} = -\sum_{t=1}^{T} L_t^{(i^*)}$$

Therefore,

$$R \leq \frac{\ln K}{\eta} + \eta \sum_{t=1}^{T} L_t^{2(i^*)} \leq \frac{\ln K}{\eta} + \eta \cdot T$$

Find the best η^*:

$$f(\eta) = \frac{\ln K}{\eta} + \eta \cdot T; \quad f'(\eta) = -\frac{\ln K}{\eta^2} + T = 0 \implies \eta^* = \sqrt{\frac{\ln K}{T}}$$

Using η^*:

$$R \leq \frac{\ln K}{\sqrt{\frac{\ln K}{T}}} + \sqrt{\frac{\ln K}{T}} \cdot T = \sqrt{T \cdot \ln K} + \sqrt{T \cdot \ln K} = 2 \cdot \sqrt{T \cdot \ln K}$$