Sequential Learning

General Setting

at each round $t = 1, \dots, T$:

1) agent chooses an action $a_t \in \mathcal{A}$, depending on available information (history of observations, experts recommendation, context ...)

2) agent receives some reward $r_t(a_t) \in \mathbb{R}$ and
loss $\ell_t(a_t)$
observes some feedback (e.g $r_t(a_t)$, or extra information ...)

<u>Goal</u>: maximise cumulated reward, or knowledge of the environment
minimise cumulated loss

→ applications in online recommendation, advertisement placement, medical trials, dynamic pricing, etc.

Lecture #1: learning with experts

Expert setting 1

At each round $t = 1, \dots, T$:

- experts output forecast f_{jt}, $j \in \{1, \dots, N\}$ number of experts
- agent aggregates experts forecast: $\hat{y}_t = \sum_{j=1}^{N} p_{jt} f_{jt}$

where $p_t = (p_{1t}, \dots, p_{Nt}) \in A$ = standard N-simplex

- agent observes true value y_t and suffers loss $\ell(\hat{y}_t, y_t)$

No stochastic model: the sequence y_t is chosen arbitrarily.

Goal: minimise cumulative loss $\sum_{t=1}^{T} \ell(\hat{y}_t, y_t)$

equivalently, minimise the regret

$$R_T = \sum_{t=1}^{T} \ell(\hat{y}_t, y_t) - \inf_{p \in A} \sum_{t=1}^{T} \ell\Big(\sum_{j=1}^{N} \underline{p_j} f_{jt}, y_t\Big)$$

p fixed for every t.

Ex: square loss: $\ell(\hat{y}_t, y_t) = (\hat{y}_t - y_t)^2$
absolute loss: $\ell(\hat{y}_t, y_t) = |\hat{y}_t - y_t|$

we will assume ℓ_t convex in $\hat{y}_t$.

$$\underbrace{\sum_{t=1}^{T} \ell(\hat{y}_t, y_t)}_{\text{cumulative loss}} = \underbrace{\inf_{p \in \Delta} \sum_{t=1}^{T} \ell\Big(\sum_{j=1}^{N} p_j f_j^t, y_t\Big)}_{\text{leading term (of order } T)} + \overbrace{R_T}^{\text{we will get this in } o(T) \text{ (second order term)}}$$

Expert setting 2

At each round $t = 1, \dots, T$:

- agent picks $p_t \in \Delta$ (N-simplex)
- true loss function $\ell_t : \Delta \longrightarrow \mathbb{R}$ is revealed and agent suffers loss $\ell_t(p_t)$

Regret $\quad R_T = \sum_{t=1}^{T} \ell_t(p_t) - \inf_{p \in \Delta} \sum_{t=1}^{T} \ell_t(p)$

Cover setting 1:

- $\ell_t(q) = \ell\Big(\sum_{j=1}^{N} q_j f_j^t, y_t\Big)$
- f_j^t are not observed beforehand here (harder)

In the following, consider setting 2 with linear loss:

$$\ell_t(q) = \sum_{j=1}^{N} q_j \ell_{jt} \qquad \text{with} \quad \ell_{jt} = \ell_t(\delta_j) \in [0, 1]$$

Exponentially weighted average predictor (EWA)

Algorithm : learning rate $\eta > 0$

$p_1 = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)$

For $t \geq 2$:

$$p_{jt} = \frac{e^{-\eta \sum_{s=1}^{t-1} l_{js}}}{\sum_{k=1}^{N} e^{-\eta \sum_{s=1}^{t-1} l_{ks}}}$$

Theorem: for any sequence of linear loss functions $(l_{1t}, \dots, l_{Nt})_t \in \left([0,1]^N\right)^{\mathbb{N}}$, EWA($\eta$) satisfies:

$$R_T \leq \frac{\ln N}{\eta} + \eta \frac{T}{8} \qquad (*)$$

Proof: **Hoeffding's lemma**: for random variables $X \in [0,1]$

$\forall \eta \in \mathbb{R}, \quad \ln \mathbb{E}[e^{\eta X}] \leq \eta \mathbb{E}[X] + \frac{\eta^2}{8}$

(proven in lecture 2)

so for any t:

$$\ln\Big(\sum_j p_{jt}\, e^{-\eta l_{jt}}\Big) \leq -\eta \sum_j p_{jt}\, l_{jt} + \frac{\eta^2}{8}$$

$$\sum_j p_{jt}\, l_{jt} \leq -\frac{1}{\eta} \underbrace{\ln\Big(\sum_j p_{jt}\, e^{-\eta l_{jt}}\Big)}_{= \ln\left(\frac{\sum_j e^{-\eta \sum_{s=1}^{t} l_{js}}}{\sum_k e^{-\eta \sum_{s=1}^{t-1} l_{js}}}\right)} + \frac{\eta}{8}.$$

by def of p_{jt}

Summing over t gives a telescopic sum for the ln:

$$\sum_{t=1}^{T} \sum_j p_{jt}\, l_{jt} \leq -\frac{1}{\eta} \ln\Big(\sum_j e^{-\eta \sum_{s=1}^{T} l_{js}}\Big) + \frac{1}{\eta}\ln(N) + \frac{\eta T}{8}.$$

Moreover,

$$\ln\Big(\sum_j e^{-\eta \sum_{s=1}^{T} l_{js}}\Big) \geq \ln\Big(\max_j e^{-\eta \sum_{s=1}^{T} l_{js}}\Big)$$

$$\geq -\eta \min_j \sum_{s=1}^{T} l_{js}.$$

So, finally:

$$\sum_{t=1}^{T} \sum_{j} p_{jt}\, l_{jt} \leq \min_{j} \sum_{s=1}^{T} l_{js} + \frac{\ln N}{\eta} + \frac{\eta T}{8} \quad \square$$

Corollary of the bound (*):

Taking $\eta = \sqrt{\dfrac{8 \ln N}{T}}$: $\boxed{R_T \leq \sqrt{\dfrac{T \ln N}{2}}}$

Questions:

1) why not just choose the best expert up to t?

Consider "Follow the leader" (FTL) strategy:

$$p_1 = \left(\frac{1}{N}, \dots, \frac{1}{N}\right)$$

$$p_{jt} = 0 \quad \text{if} \quad \sum_{s=1}^{t-1} l_{js} > \min_{k} \sum_{s=1}^{t-1} l_{ks}.$$

Equivalent to EWA with $\eta = \infty$.

$\rightarrow$ EWA is a smoothed version of FTL (η is a smoothing factor)

Claim: There exist sequences $(\ell_{1t}, \dots, \ell_{Nt})_t$ such that

$$\exists \delta > 0, \forall T, \ R_T(FTL) \geq \delta T$$

Example with $N = 2$:

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
ℓ_{1t}	0	1	0	1	0
ℓ_{2t}	1/2	0	1	0	1

agent gets loss $= T - 3/4$

$$\min\left(\sum_{t=1}^{T} \ell_{1t}, \sum_{t=1}^{T} \ell_{2t}\right) = \frac{T}{2} - \frac{1}{2}$$

$\square$

2) Can we be more ambitious and hope to get

$$\sum_{t=1}^{T} \min_k \ell_{kt} \ ?$$

Claim: No strategy satisfies for all sequences $(\ell_{1t}, \dots, \ell_{Nt})_t \in ([0,1]^N)^{\mathbb{N}}$:

$$\sum_{t,j} p_{jt} \ell_{jt} - \sum_t \min_k \ell_{kt} = o(T)$$

→ exercise at home

3) The corollary tunes $\eta = \sqrt{\frac{8 \ln N}{T}}$

What if we don't know T? $\longrightarrow$ doubling trick

(see exercise session #2)

4) What if the losses are not in $[0,1]$ but $[m, M]$?

$\longrightarrow$ we can just rescale the observed losses as

$\tilde{\ell}_t = \frac{\ell_t - m}{M - m}$ or equivalently, run EWA with $\tilde{\eta} = \frac{\eta}{M-m}$

But what if m and M are unknown?

In that case, we can choose adaptive η_t as:

$$\eta_t = \frac{\ln N}{\sum_{s=1}^{t-1} \delta_s} \quad \text{where } \delta_t = \sum_{j=1}^{N} p_{jt} \ell_{jt} + \frac{1}{\eta_t} \ln \left(\sum_{j=1}^{N} p_{jt} e^{-\eta_t \ell_{jt}} \right)$$

$\longrightarrow$ more on that in exercise session #1

Optimality of the $\sqrt{\frac{T}{2}\ln N}$ bound

(asymptotic lower bound)

Theorem: $$\liminf_{N\to+\infty}\ \liminf_{T\to+\infty}\ \inf_{\text{algorithms}}\ \sup_{\ell_{jt}\in[0,1]} \frac{\sum_{j,t} p_{jt}\ell_{jt} - \min_k \sum_t \ell_{kt}}{\sqrt{\frac{T}{2}\ln N}} = 1$$

The order of the quantifier (and in particular inf and sup here) is important!

Proof: we have already shown with EWA that this limit is ≤ 1.

Now, we lower bound $\sup_{\ell_{jt}}$ by some $\mathbb{E}$ with random j.

Let $\ell_{jt} \overset{iid}{\sim} \mathrm{Ber}\left(\frac{1}{2}\right)$

obviously, $\mathbb{E}\left[\sum_j p_{jt}\ell_{jt}\right] = \frac{1}{2}$ for any p_t and t.

(and so for any algo.)

But we can show:

$$(+)\quad \mathbb{E}\left[\min_k \sum_{t=1}^{T} \ell_{kt}\right] = \frac{T}{2} - \sqrt{\frac{\ln N}{2} T} + o\left(\sqrt{T \ln N}\right)$$

so that

$$\sup_{\ell_{jt} \in [0,1]} \sum_{j,t} p_{jt} \ell_{jt} - \min_k \sum_t \ell_{kt} \geq \mathbb{E}_{\ell_{jt} \sim \mathrm{Ber}(\frac{1}{2})}\left[\sum_{t,j} p_{jt} \ell_{jt} - \min_k \sum_t \ell_{jt}\right]$$

$$\geq \sqrt{\frac{\ln N}{2} T} + o(\sqrt{T}). \quad \blacksquare$$

Proof of (+):

Denote $Z_{kT} = \dfrac{\sum_{t=1}^{T} (\frac{1}{2} - \ell_{kt})}{\frac{1}{2}\sqrt{T}}$

CLT yields: $Z_T = \begin{pmatrix} Z_{1T} \\ \vdots \\ Z_{NT} \end{pmatrix} \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, I_N)$.

for $f_N(x) = \max_{k \leq N} x_k$, we want an asymptotic of $\mathbb{E}[f_N(Z_T)]$.

Note that $\mathbb{E}[f_N(Z_T)^2] \leq \sum_{k=1}^{N} \mathbb{E}[Z_{kT}^2] = N$. so that:

1) $f_N(Z_T)$ is bounded in $\mathbb{L}^2$ norm independtly from T

2) $Z_T \xrightarrow{\mathcal{L}} Z$

These two conditions imply that
$$\lim_{T\to\infty} \mathbb{E}[f_N(Z_T)] = \mathbb{E}[f_N(Z)]$$

(see exercise session #1)

Thus $\lim_{T\to\infty} \mathbb{E}[f_N(Z_T)] = \mathbb{E}[f_N(Z)]$ when $Z \sim N(0, I_N)$

Reminder: we want to show

$$\liminf_{N\to+\infty} \liminf_{T\to\infty} \frac{1}{\sqrt{2\ln N}} \mathbb{E}[\max_{k\leq N} Z_{kT}] \geq 1.$$

i.e $$\liminf_{N\to+\infty} \frac{1}{\sqrt{2\ln N}} \mathbb{E}[\max_{k\leq N} Z_k] \geq 1.$$

For that, define $\quad M_N = \dfrac{(\max_{k\leq N} Z_k)_+}{\sqrt{2\ln N}}$

Since $\quad 0 \geq \mathbb{E}\left[\left(\max_{k\leq N} Z_k\right)_-\right] \geq \mathbb{E}\left[(Z_1)_-\right]$

$$\geq -\sqrt{\mathbb{E}\left[(Z_1)^2 \mathbb{1}_{Z_1\geq 0}\right]}$$

$$= -1/\sqrt{2}$$

we have $\quad \liminf_{N\to\infty} \dfrac{\mathbb{E}\left[\max_{k\leq N} Z_k\right]}{\sqrt{2\ln N}} = \liminf_{N\to\infty} \mathbb{E}[M_N]$

By Fatou's lemma:

$$\liminf_{N\to+\infty} \mathbb{E}[M_N] \geq \mathbb{E}\left[\liminf_{N\to+\infty} M_N\right]$$

For $\varepsilon \in (0,1)$, we have:

$$\mathbb{P}(M_N \leq \sqrt{1-\varepsilon}) = \mathbb{P}\left(\forall k\leq N,\ Z_k \leq \sqrt{2(1-\varepsilon)\ln N}\right)$$

$$= \left(F\left(\sqrt{2(1-\varepsilon)\ln N}\right)\right)^N \qquad F \text{ cdf of } \mathcal{N}(0,1)$$

$$\leq e^{-N(1-F(\cdot))} \qquad 1+u \leq e^u$$

As $\quad 1 - F(x) \underset{x \to \infty}{\sim} \dfrac{e^{-x^2/2}}{x\sqrt{2\pi}}$ (proved with integration by parts, see below)

we have $\quad N(1-F(\cdots)) \sim \dfrac{N e^{-(1-\varepsilon)\ln N}}{2\sqrt{\pi(1-\varepsilon)\ln N}} = \dfrac{N^{\varepsilon}}{2\sqrt{\pi(1-\varepsilon)\ln N}}$

Thus: $\displaystyle\sum_N e^{-N(1-F(\cdots))}$ is summable

$\displaystyle\sum_N \mathbb{P}(M_N \leq \sqrt{1-\varepsilon}) < \infty.$

Borel-Cantelli implies that $\liminf\limits_{N \to +\infty} M_N \geq \sqrt{1-\varepsilon}$ a.s.

this holds for any $\varepsilon > 0$, so that

$\liminf\limits_{N \to +\infty} M_N \geq 1$ a.s.

and thus

$$\liminf_{N \to +\infty} \mathbb{E}[M_N] \geq \mathbb{E}\left[\liminf_{N \to +\infty} M_N\right] \geq 1. \qquad \square$$

Proof of $1-F(x) \underset{x\to\infty}{\sim} \dfrac{e^{-x^2/2}}{x\sqrt{2\pi}}$

$$\sqrt{2\pi}\,(1-F(x)) = \int_x^{+\infty} e^{-u^2/2}\,du = \int_x^{+\infty} u e^{-u^2/2}\,\frac{1}{u}\,du$$

IbP

$$= \left[-\frac{1}{u}e^{-u^2/2}\right]_x^{+\infty} - \int_x^{+\infty} \frac{e^{-u^2/2}}{u^2}\,du$$

$$= \frac{e^{-x^2/2}}{x} - \int_x^{+\infty} \frac{e^{-u^2/2}}{u^2}\,du \qquad (*)$$

2nd IBP:

$$\int_x^{+\infty} \frac{e^{-u^2/2}}{u^2}\,du = \left[-\frac{e^{-u^2/2}}{u^3}\right]_x^{+\infty} - \int_x^{+\infty} \frac{3e^{-u^2/2}}{u^4}\,du$$

$x \geq 0$

$$0 \leq \qquad \leq \frac{e^{-x^2/2}}{x^3}$$

So $\displaystyle\int_x^{+\infty} \frac{e^{-u^2/2}}{u^2}\,du = o\left(\frac{e^{-x^2/2}}{x}\right)$ and $(*)$ concludes $\square$