Gilles Stoltz

Prediction with Expert Advice: Theory and Applications

Lecture #3

I. Recap:

Randomized prediction with expert advice:

- Outcome space $\mathcal{Y}$: sequential outcomes $y_1, y_2, \ldots$
- Decision space $\mathcal{X}$: sequential predictions $\hat{y}_1, \hat{y}_2, \ldots$
- Loss function $\ell : \mathcal{X} \times \mathcal{Y} \to [0,1]$ (takes bounded values!)
- N experts; the j-th expert ($j = 1, \ldots N$) recommends f_{jt} as a prediction for y_t at round t

→ **How the game goes:** [The outcome sequence $y_1, y_2, \ldots$ is fixed in advance.]

For each round $t = 1, 2, \ldots$

(1) experts form their predictions $f_{1t}, \ldots f_{Nt}$ (based on $y_1, \ldots y_{t-1}$)

(2) (a) the forecaster chooses a probability distribution

$$p_t = (p_{1t}, \ldots p_{Nt}) \text{ over the experts,}$$

(b) chooses an expert index I_t at random according to p_t,

(c) and uses the prediction

$$\hat{y}_t = f_{I_t, t}$$

(3) the environment reveals the outcome y_t

(4) the forecaster ensures a loss of $\ell(f_{I_t,t}, y_t)$ and can compute the losses $\ell(f_{jt}, y_t)$ of all experts.

→ **Goal** is to minimize the regret:

$$R_n = \sum_{t=1}^{n} \ell(f_{I_t,t}, y_t) - \min_{j=1,\ldots N} \sum_{t=1}^{n} \ell(f_{jt}, y_t)$$

↪ design strategies such that $R_n = o(n)$.

A good strategy is the following (the exponentially weighted average strategy)

p_1 uniform and for $t \geq 2$, $p_{jt} = \exp\left(-\eta \sum_{s=1}^{t-1} \ell(f_{js}, y_s)\right) \Big/ \sum_{i=1}^{N} \exp\left(-\eta \sum_{s=1}^{t-1} \ell(f_{is}, y_s)\right)$

for some parameter $\eta > 0$ to be tuned by the analysis.

→ <u>Definition</u>: Expected loss at round t:

$$\bar{\ell}(p_t, y_t) = \mathbb{E}\, \ell(f_{I_t, t}, y_t) = \sum_{i=1}^{N} p_{it}\, \ell(f_{it}, y_t)$$

Expected regret till round n:

$$\bar{R}_n = \mathbb{E} R_n = \sum_{t=1}^{n} \bar{\ell}(p_t, y_t) - \min_{j=1,\dots,N} \sum_{t=1}^{n} \ell(f_{jt}, y_t)$$

→ <u>Theorem</u>: For all sequences of outcomes $y_1, y_2, \dots$ the expected regret of the exponentially weighted average forecasted is less than

$$\bar{R}_n \leq \frac{\ln N}{\eta} + \frac{\eta}{2} \sum_{t=1}^{n} \sum_{i=1}^{N} p_{it}\, \ell(f_{it}, y_t)^2.$$

<u>Rk</u>: Tuning of η discussed later on (after the proof).

<u>Proof</u>: We will use the inequalities $\forall x \leq 0,\ e^x \leq 1 + x + x^2/2$ (easy to show by function study) and $\forall u > -1$, $\ln(1+u) \leq u$.

We denote $W_t = w_{1t} + \dots + w_{Nt}$ for $t \geq 1$

where $w_{jt} = \exp\left(-\eta \sum_{s=1}^{t} \ell(f_{js}, y_s)\right)$;

and $W_0 = N$. Thus, $p_{jt} = \dfrac{w_{j,t-1}}{W_{t-1}}$.

First,

$$\ln \frac{W_n}{W_0} = \ln\left(\sum_{j=1}^{N} w_{j,n}\right) - \ln N \geq \max_{j=1,\dots,N} \ln w_{jn} - \ln N$$

$$= -\eta \sum_{t=1}^{n} \ell(f_{jt}, y_t) - \ln N$$

On the other hand, we upper bound $\ln \frac{W_n}{W_0}$ by upper bounding, for each $t = 0, \dots, n-1$, the ratio $\ln \frac{W_{t+1}}{W_t}$:

$$\ln \frac{W_{t+1}}{W_t} = \ln \frac{\sum_{i=1}^N w_{it}\, e^{-\eta \ell(f_{it+1}, y_{t+1})}}{\sum_{j=1}^N w_{jt}} = \ln \sum_{i=1}^N p_{it}\, e^{-\eta \ell(f_{it}, y_t)}$$

$$\leq \ln\left(\sum_{i=1}^N p_{it} \left(1 - \eta \ell(f_{it}, y_t) + \frac{\eta^2}{2} \ell(f_{it}, y_t)^2\right)\right)$$

$$\leq -\eta \underbrace{\sum_{i=1}^N p_{it}\, \ell(f_{it}, y_t)}_{\bar{\ell}(p_t, y_t)} + \frac{\eta^2}{2} \sum_{i=1}^N p_{it}\, \ell(f_{it}, y_t)^2$$

Summing over $t = 0, \dots, n-1$:

$$\ln \frac{W_n}{W_0} \leq -\eta \sum_{t=1}^n \bar{\ell}(p_t, y_t) + \frac{\eta^2}{2} \sum_{t=1}^n \sum_{i=1}^N p_{it}\, \ell(f_{it}, y_t)^2$$

Combining with the lower bound and rearranging, we get the desired bound.

$\longrightarrow$ Tuning of η:

* If n know in advance, using $\ell(f_{it}, y_t) \in [0,1]$ we have

$$\bar{R}_n \leq \frac{\ln N}{\eta} + \frac{\eta}{2} n = \sqrt{2n \ln N}$$

for the choice of $\eta = \sqrt{\frac{2 \ln N}{n}}$.

* Otherwise, can choose η in a past-dependent way: p_1 uniform and

for $t \geq 2$,
$$\eta_t = \square \sqrt{\frac{\ln N}{\sum_{s=1}^{t-1} \sum_{i=1}^N p_{is}\, \ell(f_{is}, y_s)^2}} \quad \text{and} \quad p_{jt} = \frac{\exp\left(-\eta_t \sum_{s=1}^{t-1} \ell(f_{js}, y_s)\right)}{\sum_{i=1,\dots,N} \exp\left(-\eta_t \sum_{s=1}^{t-1} \ell(f_{is}, y_s)\right)}$$

Corollary: This fully adaptive strategy ensures that

$$\bar{R}_n \leq \square \sqrt{\left(\sum_{t=1}^n \sum_{i=1}^N p_{it}\, \ell(f_{it}, y_t)^2\right) \ln N}.$$

$\longrightarrow$ High-probability bounds. What about R_n?

Martingale arguments (namely, Bernstein's inequality) show that with probability at least $1-\delta$,

$$R_n \leq R_n + \sqrt{\sum_{t=1}^{n} \sum_{i=1}^{N} p_{it} \; \ell(f_{it}, y_t)^2 \quad \ln 1/\delta}$$

$$\leq \sqrt{\left(\sum_{t=1}^{n} \sum_{i=1}^{N} p_{it} \; \ell(f_{it}, y_t)^2\right) \ln \frac{N}{\delta}}$$

Note: This is nice, since the order of magnitude of the deviations of R_n vs $\overline{R}_n$ matches the one of the bound on $\overline{R}_n$.

II. The multi-armed bandit problem.

In this setting, the forecaster still monitors his own losses $\ell(f_{I_t,t}, y_t)$ but no longer the losses of the experts he didn't follow: $\ell(f_{jt}, y_t)$, $j \neq I_t$.

I.e., step (4) of page 1 is replaced by

(4') the forecaster ensures a loss of $\ell(f_{I_t,t}, y_t)$, which he can see, but has no access to the losses of the experts he didn't follow: $\ell(f_{jt}, y_t)$, $j \neq I_t$.

$\longrightarrow$ Examples:

* [where the name comes from:] In casinos, slot machines are also called one-armed bandit. Imagine you have N such machines, you pick one (the I_t-th of them), pull its arm, and get a given reward. You only monitor your own reward, you do not get to know what you would have got had you played with a different machine.
Equivalently, instead of considering N one-armed bandits, one can consider one machine with N arms, hence the name "multi-armed bandits".

* [dynamic pricing] Goods are to be sold on the Internet, customers connect one by one

to a web site. To customer #t, the good is offered at a price $\hat{y}_t \in \mathcal{X} = \{ 9.99, 14.99, 19.99, \dots, 99.99 \}$. The customer has in mind a maximal price $y_t \in [0, 100]$:

- if $\hat{y}_t \le y_t$, the good is sold but the seller could have increased his price, thus suffering a loss of earnings of $y_t - \hat{y}_t$;
- if $\hat{y}_t > y_t$, the good is not sold, the loss of earnings is y_t.

In total, the loss is

$$\ell(\hat{y}_t, y_t) = y_t - \hat{y}_t \, \mathbb{1}_{[\hat{y}_t \le y_t]}.$$

The regret takes a simple form:

$$R_n = \sum_{t=1}^{n} \ell(\hat{y}_t, y_t) - \min_{j \in \mathcal{X}} \sum_{t=1}^{n} \ell(j, y_t)$$

$$= \sum_{t=1}^{n} \left(-\hat{y}_t \, \mathbb{1}_{[\hat{y}_t \le y_t]} \right) - \min_{j \in \mathcal{X}} \sum_{t=1}^{n} \left(-j \, \mathbb{1}_{[j \le y_t]} \right)$$

Defining $\tilde{\ell}(j, y_t) = -j \, \mathbb{1}_{[j \le y_t]}$

we can compute $\tilde{\ell}(\hat{y}_t, y_t)$ (but not the $\tilde{\ell}(j, y_t)$ for $j \ne \hat{y}_t$: we only know the relative positions of y_t and $\hat{y}_t$).

* [other example:]

time needed to go from A to B by car using different possible paths

$\longrightarrow$ Key idea: <u>ESTIMATE</u> unobserved losses.

For all $i = 1, \dots N$:

$$\hat{\ell}(f_{i,t}, y_t) = \begin{cases} 0 & \text{if } I_t \ne i \\ \dfrac{\ell(f_{I_t,t}, y_t)}{p_{i,t}} & \text{if } I_t = i \end{cases}$$

(recall that I_t drawn according to p_t).

These are indeed estimators (one only needs p_t, I_t and $\ell(f_{I_t,t}, y_t)$ to compute them).

These are actually good estimators, because they are unbiased:

$$\mathbb{E}\,\hat{\ell}(f_{it}, y_t) = \mathbb{E}\Big[\ \mathbb{E}\big[\hat{\ell}(f_{it}, y_t)\ \big|\ \underbrace{I_1, \dots I_{t-1}}\big]\Big]$$

$$= \mathbb{E}\Big[\ \frac{\ell(f_{it}, y_t)}{p_{it}}\ \mathbb{E}\big[\mathbb{1}_{\{I_t = i\}}\ \big|\ I_1 \dots I_{t-1}\big]\Big]$$

we condition to fix the value of f_t

$$= \frac{\ell(f_{it}, y_t)}{p_{it}} \times p_{it} = \ell(f_{it}, y_t).$$

$\longrightarrow$ ALGORITHM: exponentially weighted averages on estimated losses.

$p_1 = (1/N, \dots 1/N)$ is the uniform distribution and given a parameter $\eta > 0$, for $t \geq 2$, p_t is given by

$$p_{jt} = \exp\Big(-\eta \sum_{s=1}^{t-1} \hat{\ell}(f_{js}, y_s)\Big) \Big/ \sum_{i=1}^{N} \exp\Big(-\eta \sum_{s=1}^{t-1} \hat{\ell}(f_{is}, y_s)\Big)$$

$\longrightarrow$ ANALYSIS:

* Note that $\displaystyle\sum_{i=1}^{N} p_{it}\, \hat{\ell}(f_{it}, y_t) = \ell(f_{I_t,t}, y_t)$ by definition of the estimators; similarly,

$$\sum_{i=1}^{N} p_{it}\, \hat{\ell}(f_{it}, y_t)^2 = \frac{1}{p_{I_t,t}}\ \ell(f_{I_t,t}, y_t) \leq 1/p_{I_t,t}$$

and thus,

$$\mathbb{E}\Big[\sum_{i=1}^{N} p_{it}\, \hat{\ell}(f_{it}, y_t)^2\Big] \leq \sum_{i=1}^{N} p_{it}\ 1/p_{it} = N.$$

* If we apply the theorem of page 2, then we get:

$$\sum_{t=1}^{n} \sum_{i=1}^{N} p_{it}\, \hat{\ell}(f_{it}, y_t) - \min_{j=1,\dots N} \sum_{t=1}^{n} \hat{\ell}(f_{jt}, y_t) \leq \frac{\ln N}{\eta} + \eta \sum_{t=1}^{n} \sum_{i=1}^{N} p_{it}\, \hat{\ell}(f_{it}, y_t)^2$$

that is, in view of the first equality above:

$$\forall j = 1, \dots N, \quad \sum_{t=1}^{n} \ell(f_{I_t,t}, y_t) \leq \sum_{t=1}^{n} \hat{\ell}(f_{jt}, y_t) + \frac{\ln N}{\eta} + \eta \sum_{t=1}^{n} \sum_{i=1}^{N} p_{it} \times \hat{\ell}(f_{it}, y_t)^2$$

Taking expectations (using unbiasedness and the inequality above), we get

$$\mathbb{E}\left[\sum_{t=1}^{n} \ell(f_{I_t,t}, y_t) \right] \leq \sum_{t=1}^{n} \ell(f_{jt}, y_t) + \frac{\ln N}{\eta} + \frac{\eta n N}{2}$$

that is, $\overline{R}_n = \mathbb{E}\, R_n \leq \frac{\ln N}{\eta} + \frac{\eta}{2} nN.$

In total, we have proven the following.

<u>Theorem.</u> The expected regret of the exponentially weighted average predictor using estimated loss is less than

$$\overline{R}_n \leq \frac{\ln N}{\eta} + \frac{\eta}{2} nN = \sqrt{2 n N \ln N}$$

for the choice $\eta = \sqrt{\frac{2 \ln N}{nN}}$.

<u>Remarks:</u>

- Same remarks as usual concerning the tuning of η (past-dependent η_t or doubling trick if n not known in advance).
- The bound worsens by a factor of $\sqrt{N}$ wrt the classical ("full-information") setting; it can be show to be necessary: $\sqrt{nN}$ lower bounds exist on the regret.

$\longrightarrow$ What about <u>NON-EXPECTED</u> REGRET R_n ?

In the current form, little can be said to relate R_n to $\overline{R}_n$.
We have to resort to some exploitation / exploration trade-off and choose the experts using distributions given by

$$p_{jt} = \underbrace{(1-\gamma) \frac{\exp\left(-\eta \sum_{s=1}^{t-1} \hat{\ell}(f_{js}, y_s)\right)}{\sum_{i=1,\dots N} \exp\left(-\eta \sum_{s=1}^{t-1} \hat{\ell}(f_{is}, y_s)\right)}}_{\text{exploitation term}} + \underbrace{\frac{\gamma}{N}}_{\text{exploration term}}$$

for some parameters
$\eta > 0$ and $\gamma > 0$
(say, $\eta = \sqrt{2 \ln N / m}$)

The expected regret of this forecaster satisfies:

$$\overline{R}_n \leq \square \sqrt{nN \ln N} + \square \gamma n.$$

The deviations of R_n from $\overline{R}_n$ can be bounded this time (using Bernstein's inequality for martingales); the point is that $\mathrm{Var}\, \hat{\ell}(f_{I_t}, y_t)$ can be controlled:

$$\begin{aligned}\mathrm{Var}\, \hat{\ell}(f_{I_t}, y_t) &\leq \mathbb{E}\left[\left(\hat{\ell}(f_{I_t}, y_t)\right)^2 \right] \\ &= \ell(f_{I_t}, y_t)^2 \frac{1}{p_{I_t}^2} \left(p_{I_t} + 0 \right) \\ &\leq 1/p_{I_t} \leq \frac{N}{\gamma}\end{aligned}$$

and thus, with probability at least $1-\delta$,

$$\begin{aligned}R_n &\leq \overline{R}_n + \square \sqrt{\left(\sum_{t=1}^{n} \mathrm{Var}(\ )\right) \ln 1/\delta} \\ &\leq \square \sqrt{nN \ln N} + \underbrace{\square \gamma n + \square \sqrt{\frac{n}{\gamma} \ln 1/\delta}}_{\gamma \sim n^{-1/3} \text{ is the optimal value and gives a } \square n^{2/3} \text{ bound}}\end{aligned}$$

$$R_n \leq \square \sqrt{n^{2/3} \ln 1/\delta} + \square \sqrt{nN \ln N}$$

The orders of magnitude are different.

This can be solved by shifting in an appropriate way the estimators: using $\hat{\ell}(f_{I_t}, y_t) - \beta/p_{I_t}$ instead of $\hat{\ell}(f_{I_t}, y_t)$.

Then, for appropriate choices of η, γ, and β (depending on δ), one gets:

<u>Th</u>: With probability at least $1-\delta$,

$$R_n \leq \square \sqrt{nN \ln (N/\delta)}$$

<u>III</u>. In a <u>stochastic setting</u>, forecasters with better practical performances exist (eg, UCB).

when $y_1, y_2, \ldots$ are realizations of, say, an iid process $Y_1, Y_2, \ldots$