DATA 37200: Learning, Decisions, and Limits
(Winter 2026)

Lecture 8: The Learning from Experts Problem

Instructor: Frederic Koehler

Reminder I

Homework is due on gradescope by 9 am tomorrow. If you have technical issues with gradescope, email your solutions to me and Joon, the TA (but gradescope is strongly preferred).

Midterm is next week (Thursday, in class). Two-sided cheat sheet, no calculator. I will release a briefy summary sheet reminding what topics we covered in the first part of this class (up to today). The midterm will not require explicit knowledge from Tuesday's class, but all the content is related so showing up on Tuesday is probably still helpful.

Reminder II

Last class we went over the reduction from the contextual bandits problem to online least squares reduction. This means we can focus on *forecasting*/ prediction/learning online for a bit. *We can forget everything about bandits for today.*

Today's lecture is based on Cesa-Bianchi and Lugosi's book *Prediction, Learning, and Games*, Chapter 3.4. I will emphasize a **Bayesian** perspective which I think makes it easier to understand.

Prediction with Expert Advice

The Online Prediction Protocol:

- We have a set of m experts, indexed by $i = 1, \ldots, m$.
- At each time step $t = 1, 2, \ldots, T$:
  1. Each expert i reveals a prediction $f_i(t) \in [0, 1]$.
  2. The learner (us) predicts $\hat{y}(t) \in [0, 1]$.
  3. The environment reveals the outcome $y(t) \in [0, 1]$.
  4. We incur squared loss $(\hat{y}(t) - y(t))^2$.
  5. Experts incur squared loss $(f_i(t) - y(t))^2$.

Goal: Minimize **Regret** relative to the best expert:

$$R_T = \sum_{t=1}^{T} (\hat{y}(t) - y(t))^2 - \min_i \sum_{t=1}^{T} (f_i(t) - y(t))^2$$

Some comments on the setup

- Customary framing here: minimize loss here rather than maximize reward. (Related by negation.)
- This setup is actually harder than what we need for the contextual bandits application from last class.
- In contextual bandits, we assumed the rewards $r(t)$ (corresponds to $y(t)$ here) are generated probabilistically with

$$\mathbb{E}[y(t) \mid x(t)] = f^*(x(t), i^*(t)).$$

  Note $m = |\mathcal{F}|$ in our previous notation, and we required $f \in \mathcal{F}$!
- In some cases, using the fact that the model is "well-specified" makes things easier.
- But for today, it does not help. We can achieve low regret for any sequence of $y(t)$.

ERM is bad (in general)

- Here by ERM I mean: play what worked best in hindsight.
- Last time I sketched why ERM suffers $\Omega(m)$ regret. (Recall $m = |\mathcal{F}|$, we will see this is extremely suboptimal !) This lower bound does hold even in the well-specified setting.
- If responses $y(t)$ are adversarial (general setup), it is possible to show ERM can suffer $\Omega(T)$ regret even when $m = O(1)$. (Why?)
- What if responses $y(t)$ are well-specified? It's no longer obvious to me. Fun question to think about.

Comment on function class size

- Why is $\Omega(m)$ regret bad?
- "Real life" function classes (e.g. linear models, neural nets, decision trees, ...) are very large, generally exponential in number of parameters.
- Linear model in d dimensions is size $e^{\Omega(d)}$, "curse of dimensionality"/Johnson-Lindenstrauss Lemma.
- More concretely, Johnson-Lindenstrauss lemma tells us there are $e^{\Omega(d)}$ vectors in d dimensions such that they are almost orthogonal — their angles are all at least 89 degrees. Tricky to visualize...

Today's Bayesian perspective

- For a statistical model with parameters θ and data X, there are two important perspectives to remember.
- **Frequentist**: the parameters θ are unknown but fixed. Randomness is over the process generating the data $X \sim p(X \mid \theta)$.
- **Bayesian**: probabilities are subjective and model our beliefs about the unknown parameter θ. The purpose of $p(X \mid \theta)$ is to tell us how to update our posterior beliefs using Bayes rule

$$p(\theta \mid X) \propto p(X \mid \theta)p(\theta)$$

based on the data X and our prior beliefs $p(\theta)$.

A Bayesian Perspective

Although the data $y(t)$ is not necessarily random, we model it stochastically to derive our algorithm. Natural idea for a Bayesian.

The Model:

- **Prior:** We assume a uniform prior over the experts.

$$\pi_i(0) = \frac{1}{m} \quad \forall i$$

- **Likelihood:** Expert i models the response $y(t)$ as a Gaussian centered at $f_i(t)$ with variance σ^2:

$$P(y(t) \mid f_i(t)) \propto \exp\left(-\frac{(y(t) - f_i(t))^2}{2\sigma^2}\right)$$

where $\sigma^2 > 0$ is a parameter we will tune later.

The Algorithm: Predicting the Posterior Mean

Posterior Update: By Bayes rule, the posterior weight of expert i at time t is:

$$\pi_i(t) \propto \pi_i(t-1) \times P(y(t) \mid f_i(t))$$

and calculating it out we get

$$\pi_i(t) = \frac{\pi_i(t-1) \exp\left(-\frac{(y(t)-f_i(t))^2}{2\sigma^2}\right)}{\sum_{j=1}^{K} \pi_j(t-1) \exp\left(-\frac{(y(t)-f_j(t))^2}{2\sigma^2}\right)}$$

Prediction Strategy: Predict the mean of the posterior distribution:

$$\hat{y}(t+1) = \sum_{i=1}^{K} \pi_i(t) f_i(t+1)$$

"Bayes optimal" prediction.

Exponential Weights Interpretation

Let us define the learning rate $\eta = \frac{1}{2\sigma^2}$.

The cumulative squared loss of expert i is $L_{i,t} = \sum_{\tau=1}^{t} (y(\tau) - f_i(\tau))^2$.

Unrolling the recursion $\pi_i(t) \propto \pi_i(t-1) e^{-\eta (f_i(t) - y(t))^2}$:

$$\pi_i(t) = \frac{e^{-\eta L_{i,t}}}{\sum_{j=1}^{m} e^{-\eta L_{j,t}}}$$

Remark: This Bayesian strategy is exactly the "Exponentially Weighted Forecaster" with learning rate $\eta = 1/2\sigma^2$ using Squared Loss. (Standard terminology.) As we will see, **it works very well for arbitrary data** despite its Bayesian derivation.

Key Property: Exp-Concavity

To analyze the regret, we use the property of **exp-concavity**.

Exp-concavity of Squared Loss

For domains $[0, 1]$ and any outcome $y \in [0, 1]$, the function:

$$G(x) = \exp\left(-\eta(x - y)^2\right)$$

is concave in x provided that $\eta \leq \frac{1}{2}$.

Sidenote: Exp-concavity of squared loss allows us to achieve a tight $O(\log K)$ bound; for some other losses, one can only hope for $O(\sqrt{T})$ regret (e.g. L^1 loss $|x - y|$ is not exp-concave).

(To take advantage of this, we will pick $\sigma^2 \geq 1$ since $\eta = 1/2\sigma^2$).

Main Result (Theorem 3.2)

Regret Bound for Squared Loss

If we choose $\eta = 1/2$ (i.e., $\sigma^2 = 1$), the regret of the Exponentially Weighted Forecaster satisfies:

$$\sum_{t=1}^{T} (\hat{y}(t) - y(t))^2 - \min_i \sum_{t=1}^{T} (f_i(t) - y(t))^2 \leq 2 \log m$$

Note:

- The bound is independent of the time horizon T.
- It grows only logarithmically with the number of experts m.

Proof via Potential Functions (Step 1)

We define **Potential Function** Φ_t as a "softmax" of $-L_{i,t}$ ($-\Phi_t$ is a "softmin" of $L_{i,t}$):

$$\Phi_t = \frac{1}{\eta} \log \left(\sum_{i=1}^{m} e^{-\eta L_{i,t}} \right)$$

Consider the potential difference $\Phi_t - \Phi_{t-1}$:

$$\Phi_t - \Phi_{t-1} = \frac{1}{\eta} \log \left(\frac{\sum_{i=1}^{m} e^{-\eta L_{i,t-1}} e^{-\eta (f_i(t) - y(t))^2}}{\sum_{j=1}^{m} e^{-\eta L_{j,t-1}}} \right)$$

Recognize the term inside $\log(\cdot)$ as an expectation under the posterior $\pi^{(t-1)}$:

$$\Phi_t - \Phi_{t-1} = \frac{1}{\eta} \log \left(\mathbb{E}_{i \sim \pi^{(t-1)}} \left[e^{-\eta (f_i(t) - y(t))^2} \right] \right)$$

Remark: potential function strategy

- We define **Potential Function** Φ_t as a "softmax" of $-L_{i,t}$ ($-\Phi_t$ is a "softmin" of $L_{i,t}$):

$$\Phi_t = \frac{1}{\eta} \log \left(\sum_{i=1}^{m} e^{-\eta L_{i,t}} \right)$$

- Note that $\Phi_0 = \log(m)/\eta$. Recall $L_{i,t} = \sum_{s=1}^{t} (y(s) - f_i(s))^2$.
- We will show that (1) the potential decreases at every step, (2) the potential goes own a lot only if our prediction is bad, and (3) the size of $-\Phi_t$ is closely related to the *minimum loss*.
- General idea: drop in potential at round t is how "surprised" we are.

Proof via Potential Functions (Step 2)

Using the **exp-concavity** property: The function $x \mapsto e^{-\eta(x-y(t))^2}$ is concave for $\eta \leq 1/2$.

By **Jensen's Inequality**:

$$\mathbb{E}_{i\sim\pi^{(t-1)}}\left[e^{-\eta(f_i(t)-y(t))^2}\right] \leq e^{-\eta(\mathbb{E}[f_i(t)]-y(t))^2}$$

Since our algorithm predicts the mean $\hat{y}(t) = \mathbb{E}_{i\sim\pi^{(t-1)}}[f_i(t)]$:

$$\mathbb{E}_{i\sim\pi^{(t-1)}}\left[e^{-\eta(f_i(t)-y(t))^2}\right] \leq e^{-\eta(\hat{y}(t)-y(t))^2}$$

Substituting this back into the potential difference:

$$\Phi_t - \Phi_{t-1} \leq \frac{1}{\eta}\log\left(e^{-\eta(\hat{y}(t)-y(t))^2}\right) = -(\hat{y}(t)-y(t))^2$$

rhs: how surprised we are

Proof Conclusion

Summing the inequality over $t = 1, \ldots, T$:

$$\Phi_T - \Phi_0 = \sum_{t=1}^{T} (\Phi_t - \Phi_{t-1}) \leq -\sum_{t=1}^{T} (\hat{y}(t) - y(t))^2$$

Rearranging and substituting $\Phi_0 = \frac{\log m}{\eta}$:

$$\sum_{t=1}^{T} (\hat{y}(t) - y(t))^2 \leq \frac{\log m}{\eta} - \Phi_T$$

We bound $-\Phi_T$:

$$-\Phi_T = -\frac{1}{\eta} \log \left(\sum_{i=1}^{K} e^{-\eta L_{i,T}} \right) \leq -\frac{1}{\eta} \log \left(\max_i e^{-\eta L_{i,T}} \right) = \min_i L_{i,T}$$

Final Bound:

$$\sum_{t=1}^{T} (\hat{y}(t) - y(t))^2 - \min_i \sum_{t=1}^{T} (f_i(t) - y(t))^2 \leq \frac{\log m}{\eta}$$

Additional remark

- From analysis, we see we could also have defined the potential function to be

$$\sum_{s=1}^{t}(y(s)-\hat{y}(s))^2+\frac{1}{\eta}\log\left(\sum_{i=1}^{m}e^{-\eta L_{i,t}}\right)$$

  and it would have been non-increasing.
- Note: this is *similar* but not *identical* to the squared loss regret!

Further Discussion

- Good to think about: why it avoids $\Omega(|\mathcal{F}|)$ lower bound arguments?
- From a purely statistical perspective, this is an excellent strategy for any functional class $\mathcal{F}$ (just discretize it if infinite). C.f. [Yang-Barron '99]
- However, it is not very algorithmically practical if $|\mathcal{F}|$ is large (again, it is often exponentially large).
- In real life we care a lot about vectors and linear models. Future class: how to do linear regression in an online fashion?
- CB reminder: Yields $\tilde{O}(T^{2/3}(\log|\mathcal{F}|)^{1/3})$ regret if we plug it into ϵ-greedy. Not that bad, but improvable.