CS281B/Stat241B. Statistical Learning Theory.
Lecture 12.

Wouter M. Koolen

- Tuning Fixed Share for low adaptive regret (See lecture 11 notes).
- Normalised Maximum Likelihood
- Universal Portfolios

Normalised Maximum Likelihood

Today we look at *probability forecasting*.

Log loss measures mismatch of probability prediction P and outcome x:

$$\operatorname{loss}(P, x) \;=\; -\ln P(x)$$

Log-loss game

Protocol:

- For $t = 1, 2, \ldots$
  - Learner chooses a distribution P_t on outcomes $\mathcal{X}$.
  - Adversary reveals outcome $x_t \in \mathcal{X}$.
  - Learner incurs the **log loss** $-\ln P_t(x_t)$.

Joint vs predictive viewpoint

Correspondence between

- Joint distribution $P(x_1, \ldots, x_T)$ and
- Prediction strategy $P(x_t | x_1, \ldots, x_{t-1})$

Log loss telescopes:

$$\sum_{t=1}^{T} - \ln P(x_t | x_1, \ldots, x_{t-1}) \;=\; - \ln P(x_1, \ldots, x_T)$$

Reference strategies : probabilistic model

This lecture: compete with a family of distributions

$$\{P_\theta(x_1, \ldots, x_T) \mid \theta \in \Theta\}$$

Regret w.r.t. Θ

$$R_T \;=\; \underbrace{\sum_{t=1}^{T} - \ln P_t(x_t)}_{\text{Loss of Learner}} - \underbrace{\inf_{\theta \in \Theta} - \ln P_\theta(x_1, \ldots, x_T)}_{\text{Loss of best parameter } \theta \in \Theta}$$

Crucial new ingredient: Learner has *white-box* access to "experts" P_θ.

From T rounds to one mega round

Sequential minimax reduces to single-shot

$$\min_{P_1} \max_{x_1} \ldots \min_{P_T} \max_{x_T} R_T \;=\; \min_{P} \max_{x_1,\ldots,x_T} R_T$$

Idea: joint distribution P already encodes all sequential behaviour

Minimax game solution

Theorem: The T-round log loss prediction game w. model Θ

$$\min_P \max_{x^T} R_T \quad \left(= \max_Q \min_P \mathop{\mathbb{E}}_{x^T \sim Q} R_T \right)$$

has value given by the *parametric complexity* of model Θ

$$\mathfrak{C}_T(\Theta) \ := \ \ln \left(\sum_{x_1, \ldots, x_T} \sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_T) \right),$$

and both Learner's minimax and Adversary's maximin strategy are the *Normalised Maximum Likelihood* distribution

$$P_T^{\text{NML}}(x_1, \ldots, x_T) \ = \ \frac{\sup_{\theta \in \Theta} P_\theta(x_1, \ldots, x_T)}{\sum_{z_1, \ldots, z_T} \sup_{\theta \in \Theta} P_\theta(z_1, \ldots, z_T)}$$

Minimax analysis

For Learner, P^{NML} equalises the regret over all sequences x^T:

$$-\ln P^{\text{NML}}(x^T) - \inf_\theta -\ln P_\theta(x^T) \;=\; \mathfrak{C}_T(\Theta)$$

For Adversary, we need to do some work:

$$\begin{aligned}
&\max_Q \min_P \mathop{\mathbb{E}}_{x^T\sim Q}\left[-\ln P(x^T) - \inf_\theta -\ln P_\theta(x^T)\right] \\
=\;& \max_Q \mathop{\mathbb{E}}_{x^T\sim Q}\left[-\ln Q(x^T) - \inf_\theta -\ln P_\theta(x^T)\right]
\end{aligned}$$

To find maximin distribution Q introduce Lagrange multiplier λ

$$= \min_\lambda \max_Q \mathop{\mathbb{E}}_{x^T\sim Q}\left[-\ln Q(x^T) - \inf_\theta -\ln P_\theta(x^T)\right] + \lambda\left(1 - \sum_{x^T} Q(x^T)\right)$$

At the optimal Q the derivative of the objective vanishes, i.e.

$$0 \;=\; -\ln Q(x^T) + 1 - \lambda - \inf_\theta -\ln P_\theta(x^T)$$

from which it follows that

$$Q(x^T) \;=\; \frac{\sup_\theta P_\theta(x^T)}{\sum_{x^T} \cdots} \;=\; P^{\text{NML}}(x^T)$$

Plugging this in, we find that the value again equals $\mathfrak{C}_T(\Theta)$.

Discussion

- Continuous case: densities w.r.t. reference measure
- Efficient implementation tricky.
- NML strategy depends on final horizon T.
- Infinite parametric complexity $\Rightarrow P^{\text{NML}}$ undefined.
- Sometimes success with *Sequential NML* (SNML), also known as *Last Step Minimax*. In round t play minimiser of

$$\min_{P_t} \max_{x_t} - \ln P_t(x_t) - \inf_{\theta \in \Theta} - \ln P_\theta(x_1, \ldots, x_t)$$

Universal portfolios

Best fixed weights

Back to the mix loss game without outcomes.

In lecture 9 we minimised regret w.r.t. the best expert.

Today we consider a more ambitious reference class:

Definition: Regret w.r.t. the best fixed weights

$$R_T = \sum_{t=1}^{T} \underbrace{- \ln \left(\sum_{k=1}^{K} w_t^k e^{-\ell_t^k} \right)}_{\text{Mix loss of Learner}} - \inf_{\boldsymbol{w}} \sum_{t=1}^{T} \underbrace{- \ln \left(\sum_{k=1}^{K} w^k e^{-\ell_t^k} \right)}_{\text{Mix loss of fixed weights } \boldsymbol{w}}$$

Investment interpretation

Loss ℓ^k is negative log return of stock k, i.e. $\ell_t^k = -\ln \frac{\text{price}_{t+1}^k}{\text{price}_t^k}$

Weight vector $\boldsymbol{w}$ is portfolio: fraction w^k of capital invested in stock k.

Mix loss is negative log return of portfolio:

$$-\ln \sum_{k=1}^{K} w_t^k e^{-\ell_t^k} \; = \; -\ln \sum_{k=1}^{K} w_t^k \frac{\text{price}_{t+1}^k}{\text{price}_t^k}$$

Strategy playing fixed weights $\boldsymbol{w}$ called *constantly rebalanced portfolio*

Strategy with low regret w.r.t. best $\boldsymbol{w}$ called *universal* portfolio

Intuition

Mix loss is convex in $\boldsymbol{w}$, so minimiser typically in interior of simplex

stock	daily return							
A	$\frac{1}{2}$	2	$\frac{1}{2}$	2	$\frac{1}{2}$	2	$\frac{1}{2}$	2
B	3	$\frac{1}{3}$	3	$\frac{1}{3}$	3	$\frac{1}{3}$	3	$\frac{1}{3}$

Best single stock: overall return 1.

Best portfolio:

$$\max_p \left\{ 4 \ln\left(p\tfrac{1}{2} + (1-p)3\right) + 4 \ln\left(p2 + (1-p)\tfrac{1}{3}\right) \right\}$$

results in $p = \frac{1}{2}$ and per-round log return

$$0.357 \approx \ln(1.43)$$

Technology

"Experts" are constantly rebalanced portfolios $\boldsymbol{w}$.

We again have white-box experts.

Close $\boldsymbol{w}$ have close loss.

First idea:

- Run AA on finely discretised simplex.
- Regret: overhead of AA w.r.t. best discretistation point plus overhead of best discretisation point w.r.t. best portfolio
- Balancing act.

Second idea:

- Run AA with density on simplex (discretise infinitely fine).
- Balancing act only in analysis.

Cover's *Universal* "algorithm"

Put Dirichlet $(\frac{1}{2}, \ldots, \frac{1}{2})$ prior on simplex: $\pi_1(\boldsymbol{w}) = \frac{\prod_{k=1}^{K} w_k^{-1/2}}{\int \ldots \, \mathrm{d}\boldsymbol{w}}$.

Density in round t:

$$\pi_t(\boldsymbol{w}) = \frac{\pi_1(\boldsymbol{w}) e^{-\sum_{s=1}^{t-1} - \ln \sum_{k=1}^{K} w^k e^{-\ell_t^k}}}{\int \cdots \, \mathrm{d}\boldsymbol{w}}$$

$$= \frac{\pi_1(\boldsymbol{w}) \prod_{s=1}^{t-1} \left(\sum_{k=1}^{K} w^k e^{-\ell_s^k} \right)}{\int \cdots \, \mathrm{d}\boldsymbol{w}}$$

Actual portfolio $\boldsymbol{w}_t$ played:

$$\boldsymbol{w}_t = \int \boldsymbol{w} \pi_t(\boldsymbol{w}) \, \mathrm{d}\boldsymbol{w}$$

Analysis

Cumulative mix loss of Universal again telescopes:

$$\begin{aligned}
&\sum_{t=1}^{T} -\ln\left(\sum_{k=1}^{K} w_t^k e^{-\ell_t^k}\right) \\
&= \sum_{t=1}^{T} -\ln\left(\sum_{k=1}^{K}\left(\int w^k \pi_t(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}\right) e^{-\ell_t^k}\right) \\
&= \sum_{t=1}^{T} -\ln\left(\int\left(\sum_{k=1}^{K} w^k e^{-\ell_t^k}\right)\pi_t(\boldsymbol{w})\,\mathrm{d}\boldsymbol{w}\right) \\
&= -\ln\left(\int \pi_1(\boldsymbol{w})\prod_{t=1}^{T}\left(\sum_{k=1}^{K} w^k e^{-\ell_t^k}\right)\,\mathrm{d}\boldsymbol{w}\right)
\end{aligned}$$

However, we cannot bound using a single $\boldsymbol{w}$.

Analysis

Regret

$$
\begin{aligned}
R_T &= \sup_{\boldsymbol{w}} \ln \frac{\prod_{t=1}^T \left(\sum_{k=1}^K w^k e^{-\ell_t^k}\right)}{\int \pi_1(\boldsymbol{w}) \prod_{t=1}^T \left(\sum_{k=1}^K w^k e^{-\ell_t^k}\right) \, \mathrm{d}\boldsymbol{w}} \\
&\leq \sup_{\boldsymbol{w}} \max_{k_1,\ldots,k_T} \ln \frac{\prod_{t=1}^T w_{k_t} e^{-\ell_t^{k_t}}}{\int \pi_1(\boldsymbol{w}) \prod_{t=1}^T w_{k_t} e^{-\ell_t^{k_t}} \, \mathrm{d}\boldsymbol{w}} \\
&= \sup_{\boldsymbol{w}} \max_{k_1,\ldots,k_T} \ln \frac{\prod_{t=1}^T w_{k_t}}{\int \pi_1(\boldsymbol{w}) \prod_{t=1}^T w_{k_t} \, \mathrm{d}\boldsymbol{w}} \\
&\leq \frac{K-1}{2} \ln T + \ln K
\end{aligned}
$$

The first inequality follows from the log-sum inequality, which is tight when *every round* all but one expert suffer infinite loss.

The last inequality is the regret bound for the KT estimator (see Catoni 2004).
Now the worst-case is when $k_1, \ldots, k_T$ identical.

Discussion

Universal algorithm not efficient. Can integrate in $O(T^{K-1})$ time.

Algorithms based on sampling. Typically $\sqrt{T}$ regret regime.

Can use efficient Online Newton Step, which has $O(\ln T)$ regret provided the expert losses are bounded (shares do not spike/crash too much per round).

Transaction costs?

Adaptive regret? Perhaps slow drift instead of abrupt switches?