Recap: Hedge

$$L_{t-1}(i) = \sum_{s=1}^{t-1} \ell_s(i) \qquad w_t(i) = e^{-\eta L_{t-1}(i)}$$

$$W_t = \sum_i w_t(i) \qquad p_t(i) = \frac{w_t(i)}{W_t}$$

$$R_T = \sum_{t=1}^{T} \ell_t(i) p_t(i) - \sum_{t=1}^{T} \ell_t(k_T^*) \le \frac{\ln K}{\eta} + \frac{\eta}{2} \sum_t \sum_i \ell_t(i)^2 p_t(i)$$

This inequality holds for $\ell_t(i) \geq 0$

If $\ell_t(i) \in [0,1]$ then $R_T \le \sqrt{T \ln K}$

Bandit setting

For $t = 1, 2, \ldots$
- Draw $I_t \sim p_t$
- Pay $\ell_t(I_t)$
- Observe $\ell_t(I_t)$
- Update $p_t \to p_{t+1}$

Cannot compute $L_t(i) = \sum_{s=1}^{t} \ell_s(i)$

Trick: importance weighted estimates

$$\hat{\ell}_t(i) = \begin{cases} \ell_t(i)/p_t(i) & \text{if } I_t = i \\ 0 & \text{otherwise} \end{cases}$$

Exp3 is Hedge run w/ $\hat{\ell}_t(i)$ instead of $\ell_t(i)$

Note : $p_t(i)$ are now random

Fix $I_1, \dots I_T$ and do Hedge analysis using $\hat{\ell}_t \geq 0$

$$k_T^* = \underset{i=1,\dots K}{\operatorname{argmin}} \sum_{t=1}^{T} \ell_t(i) \qquad \hat{R}_T = \sum_t \hat{\ell}_t(i) p_t(i) - \sum_t \hat{\ell}_t(k_T^*)$$

$$\hat{R}_T \leq \frac{\ln K}{\eta} + \frac{\eta}{2} \sum_t \sum_i \hat{\ell}_t(i)^2 p_t(i) \qquad \circledast$$

$\mathcal{F}_{t-1}$ = σ-algebra generated by $I_1, \dots I_{t-1}$

$$\mathbb{E}\left[\hat{\ell}_t(i) \middle| \mathcal{F}_{t-1}\right] = \frac{\ell_t(i)}{p_t(i)} \underbrace{\mathbb{P}(I_t = i | \mathcal{F}_{t-1})}_{\rightarrow p_t(i)} + 0 = \ell_t(i)$$

$$\mathbb{E}\left[\hat{\ell}_t(i)^2 \middle| \mathcal{F}_{t-1}\right] = \frac{\ell_t(i)^2}{p_t(i)^2} \underbrace{\mathbb{P}(I_t = i | \mathcal{F}_{t-1})}_{\rightarrow p_t(i)} + 0 \leq \underbrace{\frac{1}{p_t(i)}}_{\rightarrow \ell_t(i) \in [0,1]}$$

$$\mathbb{E}\left[\hat{R}_T\right] = \mathbb{E}\left[\sum_t \ell_t(i) p_t(i)\right] - \sum_t \ell_t(k_T^*)$$
$$= \mathbb{E}\left[\sum_t \ell_t(I_t)\right] - \sum_t \ell_t(k_T^*) \quad \text{regret } \checkmark$$

From $\circledast$

$$\mathbb{E}\left[\hat{R}_T\right] \leq \frac{\ln K}{\eta} + \frac{\eta}{2} \mathbb{E}\left[\sum_t \sum_i \mathbb{E}\left[\hat{\ell}_t(i)^2 \middle| \mathcal{F}_{t-1}\right] p_t(i)\right]$$

$$\le \frac{\ln K}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_t \sum_i \frac{1}{p_t(i)}\, p_t(i)\right] = \frac{\ln K}{\eta} + \frac{\eta}{2} KT$$

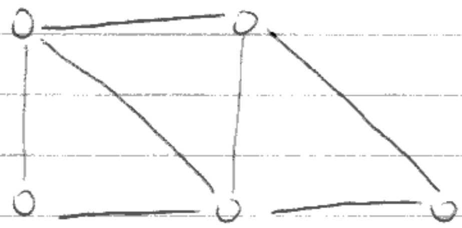
By tuning η, regret is bounded as $\sqrt{KT\ln K}$

We lose $\sqrt{K}$ w.r.t. case when ℓ_t is fully observable

Tuning $\eta_t \simeq \sqrt{\frac{K\ln K}{t}}$ gives same bound (up to constants) uniformly over time

Interpolating experts & bandits

View actions as nodes of an undirected graph $G=(V,E)$



If $I_t = i$, then we observe the losses of $N_i = \{i\} \cup \{j \in V : (i,j) \in E\}$

Experts: G is a clique, Bandits: G has no edges

$$\hat{\ell}_t(i) = \begin{cases} \ell_t(i)/Q_t(i) & \text{if } I_t \in N_i \quad (\ell_t(i) \text{ is observed}) \\ 0 & \text{otherwise} \end{cases}$$

$$Q_t(i) = \mathbb{P}\left(I_t \in N_i \mid \mathcal{F}_{t-1}\right)$$

As before: $\mathbb{E}\left[\hat{\ell}_t(i) \mid \mathcal{F}_{t-1}\right] = \ell_t(i)$

Also: $\mathbb{E}\left[\hat{\ell}_t(i)^2 \mid \mathcal{F}_{t-1}\right] \le 1/Q_t(i)$

Thus: $\mathbb{E}\left[\hat{R}_T\right] \leq \frac{\ln K}{\eta} + \frac{\eta}{2}\mathbb{E}\left[\sum_t \underbrace{\sum_i \frac{1}{Q_t(i)} p_t(i)}_{(*)}\right]$

$(*) = \sum_i \frac{p_t(i)}{\sum_{j \in N_i} p_t(j)} \leq \alpha$ for <u>any</u> $\{p_t(1), \dots p_t(K)\}$

↳ independence number of G

Hence: $\mathbb{E}\left[\hat{R}_T\right] \leq \frac{\ln K}{\eta} + \frac{\eta}{2}\alpha T \ln K \rightsquigarrow \sqrt{\alpha T \ln K}$

↳ Tuning η

G = clique (experts) $\overset{\boxed{\alpha = 1}}{\Longrightarrow} \mathbb{E}\left[\hat{R}_T\right] \leq \sqrt{T \ln K}$ ✓

G = edgeless (bandits) $\overset{\boxed{\alpha = K}}{\Longrightarrow} \mathbb{E}\left[\hat{R}_T\right] \leq \sqrt{K T \ln K}$ ✓

<u>Exercise</u>: $G = (V, E)$ undirected graph

$\forall i \in V \quad N_i = \{i\} \cup \{j : (i,j) \in E\}$

Then $\sum_{i \in V} \frac{1}{|N_i|} \leq \alpha$ ← Prove it

Online Convex Optimization

S is a convex subset of a linear space

For $t = 1, 2, \ldots$

- Play $w_t \in S$
- Observe convex loss $\ell_t : S \to \mathbb{R}$
- Pay $\ell_t(w_t)$
- Update $w_t \to w_{t+1} \in S$

Regret $R_T(u) = \sum_{t=1}^{T} \left(\ell_t(w_t) - \ell_t(u) \right) \quad \forall u \in S$

Experts : $w_t = p_t \quad S = \text{simplex} \quad \ell_t(p) = \ell_t^T p$ linear loss

$R_T(q) = \sum_t \left(p_t^T \ell_t - q^T \ell_t \right)$ best q always a corner of the simplex

Follow the leader (FTL)

$w_{t+1} = \underset{w \in S}{\operatorname{argmin}} \sum_{s=1}^{t} \ell_s(w)$ has linear regret in the worst case!

$S = [-1, 1] \quad \ell_1(w) = \frac{w}{2} \quad \ell_t(w) = \begin{cases} -w & t \text{ even} \\ w & t \text{ odd} \end{cases}$

Then $\sum_{s=1}^{t} \ell_s(w) = \begin{cases} -w/2 & t \text{ even} \\ w/2 & t \text{ odd} \end{cases}$

Therefore $w_{t+1} = \begin{cases} 1 & t \text{ even} \\ -1 & t \text{ odd} \end{cases} \Rightarrow \ell_{t+1}(w_{t+1}) = 1$

FTL $\forall t$

Introduce regularization to add stability

Follow the regularized leader (FTRL)

$$w_{t+1} = \underset{w \in S}{\operatorname{argmin}} \left[\eta \sum_{s=1}^{t} \ell_s(w) + \Phi(w) \right]$$

strongly convex function

$\Phi : S \to \mathbb{R}$ is β-s.c. w.r.t. a norm $\|\cdot\|$ if $\forall u, v \in S$

$$\Phi(v) \geq \Phi(u) + \nabla \Phi(u)^T (v-u) + \frac{\beta}{2} \|u - v\|^2$$

- Euclidean norm is 1-sc w.r.t. $\|\cdot\|_2$

  $\to \frac{1}{2}\|\cdot\|_2^2$

- Entropy $\sum_i p_i \ln p_i$ is 1-sc w.r.t. $\|\cdot\|_1$ (p in the simplex)

Linearization trick (using convexity of losses)

$$\ell_t(w_t) - \ell_t(u) \leq \underbrace{\nabla \ell_t(w_t)}_{\nabla_t}{}^T (w_t - u)$$

FTRL w/ linearized losses

$$w_{t+1} = \underset{w \in S'}{\operatorname{argmin}} \left[\eta \sum_{s=1}^{t} \nabla_s^T w + \Phi(w) \right] \qquad \theta_{t+1} = -\eta \sum_{s=1}^{t} \nabla_s$$

$$= \underset{w \in S'}{\operatorname{argmax}} \left[\theta_{t+1}^T w - \Phi(w) \right] = \nabla \Phi^*(\theta_{t+1})$$

$\Phi^*(\theta) = \max_{w \in S'} \left(\theta^T w - \Phi(w) \right)$ is convex dual of Φ

If Φ is sc. then Φ^* is convex and differentiable

Online Mirror Descent (OMD)

$\theta_1 = (0, \dots, 0) \in \mathbb{R}^d$

For $t = 1, 2, \dots$

Play $w_t = \nabla \Phi^*(\theta_t)$ (mirror step)

Pay $\ell_t(w_t)$

Observe $\nabla_t = \nabla \ell_t(w_t)$

Update $\theta_{t+1} = \theta_t - \eta \nabla_t$ (gradient step)

Entropic regularization $\Phi(p) = \sum_i p_i \ln p_i$

$\Phi^*(\theta) = \ln \sum_i e^{\theta_i} \qquad \nabla \Phi^*(\theta)_i = e^{\theta_i} / \sum_j e^{\theta_j}$

Recall $\theta_t = -\eta \sum_{s=1}^{t-1} \nabla_s$ if ℓ_t are linear then OMD $\equiv$ Hedge