Adversarial Bandits.

For some fixed time horizon T imagine matrix

$$x_{t,i} \qquad \begin{array}{l} t \in \{1, \dots, T\} \\ i \in \{1, \dots, n\} \end{array}$$

Your algorithm chooses I_t based on history

$$R_T = \max_i \sum_{t=1}^{T} x_{t,i} - x_{t,I_t}$$

Suppose I_t is a deterministic choice given history?

I_1 is deterministic. So define $x_{1,i} = \begin{cases} 0 & \text{if } i = I_1 \\ 1 & \text{o.w.} \end{cases}$

I_2 is deterministic given (I_1, X_{1,I_1})

So define $x_{2,i} = \begin{cases} 0 & \text{if } i = I_2 \\ 1 & \text{o.w.} \end{cases}$

Any deterministic algorithm gets $\Omega(T)$ regret against adversarial sequence.

At time t, draw $I_t \sim P_t$, observe X_{t,I_t}

and construct $\hat{x}_{t,i} = \dfrac{\mathbb{1}\{I_t = i\}}{P_{t,i}} X_{t,I_t}$

$$\mathbb{E}[\hat{x}_{t,i} \mid P_t] = \mathbb{E}\left[\frac{\mathbb{1}\{I_t = i\}}{P_{t,i}} X_{t,I_t} \mid P_t \right]$$

$$= \sum_{j=1}^{n} \underbrace{\mathbb{P}(I_t = j)}_{= p_{t,j}} \frac{\mathbb{1}\{j = i\}}{p_{t,i}} X_{t,j}$$

$$= p_{t,i} \cdot \frac{1}{p_{t,i}} \cdot X_{t,i} = x_{t,i}$$

$$S_{t,i} = \sum_{s=1}^{t} \hat{x}_{t,i} \qquad \mathbb{E}[S_{t,i}] = \sum_{s=1}^{t} x_{t,i} \quad \forall i$$

$$\mathbb{V}(\hat{x}_{t,i}) = \mathbb{E}[\hat{x}_{t,i}^2] - \underbrace{\mathbb{E}[\hat{x}_{t,i}]}_{x_{t,i}}{}^2$$

$$= \sum_{j} p_{t,j} \frac{\mathbb{1}\{j = i\}}{p_{t,i}^2} x_{t,j}^2 - x_{t,i}^2$$

$$= \frac{x_{t,i}^2}{p_{t,i}} - x_{t,i}^2 = \frac{x_{t,i}^2 (1 - p_{t,i})}{p_{t,i}}$$

$x_{t,i} =: 1 - \ell_{t,i}$ Assum $x_{t,i} \in [0,1]$

$$\max_i \sum_{t=1}^{T} x_{t,i} - x_{t,I_t} = \max_i \sum_{t=1}^{T} \ell_{t,I_t} - \ell_{t,i}$$

$$\hat{\ell}_{t,i} = \frac{\mathbb{1}\{I_t = i\}}{p_{t,i}} \ell_{t,I_t}$$

EXP3

Adversary chooses $\{\ell_{t,i}\}_{\substack{i=1,\dots,n \\ t=1,\dots,T}}$

$W_0 = \mathbb{1} \in \mathbb{R}^n$, η given

for $t = 1, 2, \dots, T$

$$P_{t,i} = \frac{W_{t-1,i}}{\sum_j W_{t-1,j}} = \frac{W_{t-1,i}}{W_{t-1}} \quad (P_t \in \Delta_n)$$

$I_t \sim P_t$, observe ℓ_{t,I_t}

Set $$\hat{\ell}_{t,i} = \frac{\mathbb{1}\{I_t = i\}}{P_{t,i}} \ell_{t,I_t}$$

$$W_{t,i} = \exp(-\eta \hat{\ell}_{t,i}) W_{t-1,i}$$
$$= \exp\left(-\eta \sum_{s=1}^{t} \hat{\ell}_{s,i}\right)$$

Theorem Regret of Exp3 for any loss seq in $[0,1]^{T\times n}$ is

$$\max_i \mathbb{E}\left[\sum_{t=1}^{T} \ell_{t,I_t} - \ell_{t,i}\right] \le \sqrt{2nT\log(n)}.$$

$$\sum_{t=1}^{T} \log\left(\frac{W_t}{W_{t-1}}\right) \qquad W_t = \sum_i w_{t,i}$$

$$\sum_{t=1}^{T} \log\left(\frac{W_t}{W_{t-1}}\right) = \log\left(\frac{W_T}{W_{T-1}}\right) + \log\left(\frac{W_{T-1}}{W_{T-2}}\right) + \dots -$$

$$= \log(W_T) - \log(W_{T-1}) + \log(W_{T-1}) - \log(W_{T-2}) + \dots$$

$$= \log(W_T) - \log(W_0)$$

$$= \log\left(\sum_i \exp\left(-\eta \sum_{t=1}^{T} \hat{\ell}_{t,i}\right)\right) - \log(n)$$

$$\geq -\eta \sum_{t=1}^{T} \hat{\ell}_{t,i} - \log(n) \quad \text{for any } i \in [n].$$

$$\log\left(\frac{W_t}{W_{t-1}}\right) = \log\left(\sum_{i=1}^{n} \frac{w_{t,i}}{W_{t-1}}\right)$$

$$= \log\left(\sum_{i=1}^{n} \frac{w_{t-1,i}}{W_{t-1}} \exp(-\eta \hat{\ell}_{t,i})\right)$$

$$= \log\left(\sum_{i=1}^{n} p_{t,i} \exp(-\eta \hat{\ell}_{t,i})\right)$$

$$e^{x} \geq 1 + x \quad \forall x \qquad e^{-x} \leq 1 - x + \frac{x^2}{2}, \quad x \geq 0$$

$$\log(1+x) \leq x$$

$$f(x) \le f(0) + f'(0)x + \sup_{z \in (0,x]} f''(z) \frac{x^2}{2}$$

$$\le \log\left(\sum_{i=1}^{n} p_{t,i} \left(1 - \eta \hat{\ell}_{t,i} + \frac{\eta^2 \hat{\ell}_{t,i}^2}{2} \right) \right)$$

$$= \log\left(1 - \eta \ell_{t,I_t} + \frac{\eta^2}{2} \sum_{i=1}^{n} p_{t,i} \hat{\ell}_{t,i}^2 \right)$$

$$\le -\eta \ell_{t,I_t} + \frac{\eta^2}{2} \sum_{i=1}^{n} p_{t,i} \hat{\ell}_{t,i}^2$$

$$\sum_{i=1}^{n} p_{t,i} \hat{\ell}_{t,i} = \sum_{i} p_{t,i} \frac{\mathbb{1}\{I_t = i\}}{p_{t,i}} \ell_{t,I_t} = \ell_{t,I_t}$$

$$-\eta \sum_{t=1}^{T} \hat{\ell}_{t,i} - \log(n) \le \sum_{t} \log\left(\frac{W_t}{W_{t-1}} \right) \le -\eta \sum_{t=1}^{T} \hat{\ell}_{t,I_t} + \frac{\eta^2}{2} \sum_{t,i} p_{t,i} \hat{\ell}_{t,i}^2$$

divide η on both sides and rearrange

$$\sum_{t=1}^{T} \hat{\ell}_{t,I_t} - \hat{\ell}_{t,i} \le \frac{\log(n)}{\eta} + \frac{\eta}{2} \sum_{t} \underbrace{\sum_{i} p_{t,i} \hat{\ell}_{t,i}^2}$$

$$\mathbb{E}\left[\sum_{t=1}^{T} \ell_{t,I_t} - \ell_{t,i} \right] = \mathbb{E}\left[\sum_{t=1}^{T} \hat{\ell}_{t,I_t} - \hat{\ell}_{t,i} \right]$$

$$\le \frac{\log(n)}{\eta} + \frac{\eta}{2} nT \quad \blacksquare$$

$$\mathbb{E}\left[\sum_i p_{t,i}\,\hat{\ell}_{t,i}^2\right] = \mathbb{E}\left[\sum_i p_{t,i}\,\frac{\mathbb{1}\{I_t=i\}}{p_{t,i}^2}\,\ell_{t,I_t}^2\right]$$

$$= \sum_j p_{t,j}\sum_i p_{t,i}\,\frac{\mathbb{1}\{j=i\}}{p_{t,i}^2}\,\ell_{t,j}^2$$

$$= \sum_i \ell_{t,i}^2 \le n$$