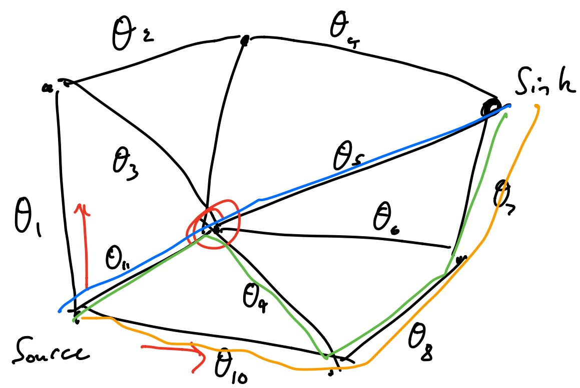

Google Maps

Given action set $\mathcal{X} \subset \mathbb{R}^d$, $|\mathcal{X}| = n$, $\mathcal{X} = \{x_1, \dots, x_n\}$

Adversary chooses $y_{t,i} \in [-1,1]$ $\forall i = 1, \dots, n$, $\forall t \in [T]$

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

Player chooses $I_t \in [n]$, plays $x_{I_t} \in \mathbb{R}^d$

Adversary reveals y_{t, I_t}

($y_{t,i} \equiv$ loss of ith arm)

$$\text{Regret} = \max_{i=1,\dots,n} \sum_{t=1}^{T} y_{t,I_t} - y_{t,i}$$

Stochastic Linear Bandits

$\exists \theta_* \in \mathbb{R}^d$ s.t. $\forall t, i \in [n]$

$$y_{t,i} = \langle x_i, \theta_* \rangle + \xi_t$$

where $\mathbb{E}[\xi_t] = 0$.

Adversarial Linear Bandits

$\forall t$ we have

$$y_{t,i} = \langle x_i, z_t \rangle.$$

EXP3(γ): Exponential Weights for Exploration Exploitation

Input: Time horizon T, n arms, $\eta > 0$, $\gamma \in [0,1]$, $\lambda \in \triangle_n$.
Initialize: Player sets $p_1 = (1/n, \dots, 1/n) \in \triangle_n$. Adversary chooses $\{y_t\}_{t=1}^T \subset [-1,1]^n$.
for: $t = 1, \cdots, T$

Player draws $I_t \sim q_t := (1-\gamma)p_t + \gamma\lambda$ and suffers (and observes) loss $\ell(I_t, y_t) = y_{t,I_t}$

Player computes $\widehat{y}_{t,i}$ where $\mathbb{E}[\widehat{y}_{t,i}|p_t] = y_{t,i}$

Update iterates:

$$\widetilde{p}_{t+1,i} = \exp(-\eta \sum_{s=1}^{t} \widehat{y}_{s,i}) \qquad p_{t,i} = \widetilde{p}_{t+1,i} / \sum_{j=1}^{n} \widetilde{p}_{t+1,j}.$$

for MAB $\hat{y}_{t,i} = \frac{\mathbb{1}\{I_t = i\} y_{t,i}}{q_{t,i}}$

Proposition 7 (EXP3(γ)). *The regret of EXP3(γ) algorithm satisfies*

$$\max_{i\in[n]} \mathbb{E}\left[\sum_{t=1}^{T} y_{t,I_t} - y_{t,i}\right] \leq \frac{\log(n)}{\eta} + 2\gamma T + \frac{1}{\eta}\sum_{t=1}^{T} \mathbb{E}\left[\sum_{i=1}^{n} q_{t,i}\phi(-\eta \widehat{y}_{t,i})\right]$$

where $\phi(x) = e^x - 1 - x \leq x^2$ *for* $|x| \leq 1$.

At time t, player chooses $I_t \sim q_t$ and plays $x_{I_t} \in \mathcal{X}$

and observes $\underline{y_{t,I_t} = x_{I_t}^T z_t}$

Define $\hat{z}_t := Q_t^{-1} x_{I_t} y_{t,I_t}$ where $Q_t := \sum_{i=1}^{n} q_{t,i} x_i x_i^T$

$$\hat{y}_{t,i} = x_i^T \hat{z}_t = x_i^T Q_t^{-1} x_{I_t} y_{t,I_t}.$$

Note $\mathbb{E}[\hat{y}_{t,i} | q_t] = \mathbb{E}[x_i^T Q_t^{-1} x_{I_t} y_{t,I_t} | q_t]$

$$= \mathbb{E}[x_i^T Q_t^{-1} x_{I_t} x_{I_t}^T z_t \mid q_t]$$

$$= x_i^T Q_t^{-1} Q_t z_t = x_i^T z_t = y_{t,i}$$

$$\mathbb{E}\left[x_{I_t} x_{I_t}^T\right] = \sum_{i=1}^{n} \mathbb{P}(I_t = i)\, x_i x_i^T = \sum_{i=1}^{n} q_{t,i}\, x_i x_i^T = Q_t$$

Choose mixing distribution $\lambda = \lambda^*$ (G-optimal design)

$$\lambda^* = \underset{\lambda \in \Delta_X}{\operatorname{argmin}} \max_{j=1,\dots,n} \|x_j\|^2_{\left(\sum \lambda_i x_i x_i^T\right)^{-1}}$$

Value achieved at d.

Need to show $|\xi \hat{y}_{t,i}| \leq 1$.

By assumption $|y_{t,i}| = |\langle x_i, z_t \rangle| \leq 1$.

$$\begin{aligned}
|\xi \hat{y}_{t,i}| &= |\xi x_i^T \hat{z}_t| \\
&= |\xi x_i^T Q_t^{-1} x_{I_t} y_{t,I_t}| \\
&\leq |\xi x_i^T Q_t^{-1} x_{I_t}| \cdot |y_{t,I_t}| \\
&\leq \xi \max_{j=1,\dots,n} \|x_j\|^2_{Q_t^{-1}} \\
&\leq \frac{\xi}{\gamma} \max_{j=1,\dots,n} \|x_j\|^2_{\left(\sum_i \lambda_i^* x_i x_i^T\right)^{-1}} \\
&= \frac{\xi d}{\gamma}
\end{aligned}$$

To make ≤ 1 set $\xi = \frac{\gamma}{d}$

Cauchy-schwartz says $\forall x, y \in \mathbb{R}^d \quad \langle x, y \rangle \leq \|x\|_2 \cdot \|y\|_2$

Note $\quad x_i^T Q_t^{-1} x_{I_t} = x_i^T Q_t^{-1/2} Q_t^{-1/2} x_{I_t} = \langle Q_t^{-1/2} x_i, Q_t^{-1/2} x_{I_t} \rangle$

$$\le \| Q_t^{-1/2} x_i \|_2 \cdot \| Q_t^{-1/2} x_{I_t} \|_2$$

$$= \| x_i \|_{Q_t^{-1}} \cdot \| x_{I_t} \|_{Q_t^{-1}}$$

$$\le \max_{j=1,\dots,n} \| x_j \|_{Q_t^{-1}}^2$$

Recall: $Q_t = \sum_{i=1}^n q_{t,i} x_i x_i^T$

$$= (1-\gamma) \sum_{i=1}^n p_{t,i} x_i x_i^T + \gamma \sum_{i=1}^n \lambda_i^* x_i x_i^T$$

For any PSD matrices A, B and vector x we have $x^T (A+B)^{-1} x \le x^T A^{-1} x$.

$$\Rightarrow \| x_j \|_{Q_t^{-1}}^2 = x_j^T Q_t^{-1} x_j$$

$$\le x_j^T \left(\gamma \sum_{i=1}^n \lambda_i^* x_i x_i^T \right)^{-1} x_j$$

Now that (w/ $\eta = \frac{\gamma}{d}$) we have $|\eta \hat{y}_{t,i}| \le 1$ we have

$$\text{Regret} \le \frac{\log(n)}{\eta} + 2\gamma T + \frac{1}{\eta} \sum_{t=1}^T \underbrace{\mathbb{E}\left[\sum_{i=1}^n q_{t,i} (\eta \hat{y}_{t,i})^2 \right]}_{\le \eta^2 d}$$

$$\le \frac{\log(n)}{\eta} + 2\gamma T + \eta T d$$

$$= \mathbb{E}\left[\sum_i q_{t,i} \, \xi^2 \, \hat{y}_{t,i}^2 \right]$$

$$\mathbb{E}\left[\hat{y}_{t,i}^2 \right] = \mathbb{E}\left[\left(x_i^T Q_t^{-1} x_{I_t} y_{t,I_t} \right)^2 \right]$$

$$\leq \mathbb{E}\left[\left(x_i^T Q_t^{-1} x_{I_t} \right)^2 \right]$$

$$= \mathbb{E}\left[x_i^T Q_t^{-1} x_{I_t} x_{I_t}^T Q_t^{-1} x_i \right]$$

$Tr(ABC)$
$= Tr(CAB)$
$= Tr(BCA)$

$$= x_i^T Q_t^{-1} x_i$$

$$\mathbb{E}\left[\sum_i q_{t,i} \, \xi^2 \, \hat{y}_{t,i}^2 \right] \leq \sum_{i=1}^n \xi^2 q_{t,i} x_i^T Q_t^{-1} x_i$$

$$= \xi^2 \sum_{i=1}^n q_{t,i} Tr\left(x_i x_i^T Q_t^{-1} \right)$$

$$= \xi^2 \, Tr\left(\underbrace{\sum_i q_{t,i} x_i x_i^T}_{Q_t} Q_t^{-1} \right)$$

$$= \xi^2 \, Tr(I) = \xi^2 d.$$

$$\text{Regret} \le \frac{\log(n)}{\eta} + 2\gamma T + \frac{1}{\eta}\sum_{t=1}^{T} \mathbb{E}\left[\sum_{i=1}^{n} q_{t,i}(\eta \hat{y}_{t,i})^2\right]$$

$$\le \frac{\log(n)}{\eta} + 2\gamma T + \eta T d$$

Chose $\eta = \frac{\gamma}{d}$

$$= \frac{d\log(n)}{\gamma} + 2\gamma T + \gamma T$$

$$\le \frac{d\log(n)}{\gamma} + 3\gamma T$$

$$\le 2\sqrt{3 d T \log(n)}$$

$\blacksquare$

$\gamma = \sqrt{\frac{d\log(n)}{3T}}$