UCB for linear bandits.

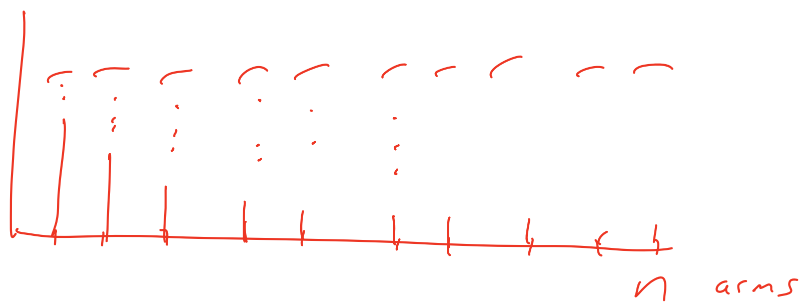
Setting: Given $\mathcal{X} \subset \mathbb{R}^d$. At time t, player chooses $x_t \in \mathcal{X}$

Observes $y_t = \langle x_t, \theta_* \rangle + \eta_t$, η_t is mean-zero sub-Gaussian

$$x_* = \underset{x \in \mathcal{X}}{\operatorname{argmax}} \langle x, \theta_* \rangle.$$

We wish to minimize regret $\mathbb{E}\left[\sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle\right]$.

Recall UCB for MAB. Pulled arm w/ highest UCB.



Suppose we observe $\{(x_s, y_s)\}_{s=1}^{t}$ and compute

$$\hat{\theta}_t = \underset{\theta}{\operatorname{argmin}} \sum_{t=1}^{T} (\langle x_t, \theta \rangle - y_t)^2 + \lambda \|\theta\|_2^2$$

$$= (X^T X + \lambda I)^{-1} X^T y$$

$$= V_t^{-1} \sum_{s=1}^{t} x_s y_s \qquad \text{where } V_t = X^T X + \lambda I = I + \sum_{s=1}^{t} x_s x_s^T$$

Fact: $\exists\, C_t \subset \mathbb{R}^d$ s.t. $\theta_* \in C_t$ $\forall t$. w.p. $\geq 1 - \delta$.

$$C_t = \{\theta : \|\theta - \hat{\theta}_{t-1}\|_{V_{t-1}}^2 \leq \beta_{t-1}\} \quad \text{for some } \beta_1 \leq \beta_2 \leq \ldots$$

UCB Plays $x_t = \arg\max_{x \in \mathcal{X}} UCB_t(x)$ $\quad UCB_t(x) := \max_{\theta \in C_{t-1}} \langle x, \theta \rangle$

Observes y_t, updates $\hat{\theta}_{t+1}$

$$\sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle .$$

Note
$$\begin{aligned} \langle x_*, \theta_* \rangle &\leq UCB_t(x_*) . \\ &\leq UCB_t(x_t) \\ &= \langle \breve{\theta}_t, x_t \rangle \quad \text{where } \breve{\theta}_t := \arg\max_{\theta \in C_t} \langle \theta, x_t \rangle \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad \langle x_* - x_t, \theta_* \rangle &= \langle x_*, \theta_* \rangle - \langle x_t, \theta_* \rangle \\ &\leq \langle x_t, \breve{\theta}_t \rangle - \langle x_t, \theta_* \rangle \\ &= \langle x_t, \breve{\theta}_t - \theta_* \rangle \\ &= x_t^T V_t^{-1/2} V_t^{1/2} (\breve{\theta}_t - \theta_*) \\ &\leq \|x_t\|_{V_t^{-1}} \cdot \|\breve{\theta}_t - \theta_*\|_{V_t} \quad \text{(Cauchy-Schwartz)} \\ &= \|x_t\|_{V_t^{-1}} \cdot \|\breve{\theta}_t - \hat{\theta}_t + \hat{\theta}_t - \theta_*\|_{V_t} \\ &\leq \|x_t\|_{V_t^{-1}} \cdot \left(\|\breve{\theta}_t - \hat{\theta}_t\|_{V_t} + \|\hat{\theta}_t - \theta_*\|_{V_t} \right) \quad \text{(Triangle inequality)} \\ &\leq 2\sqrt{\beta_t} \, \|x_t\|_{V_t^{-1}} \end{aligned}$$

Assume $\max_x \langle x_* - x, \theta_* \rangle \leq 1$

$$\Rightarrow \quad \langle x_* - x_t, \theta_* \rangle \leq \min\{1, 2\sqrt{\beta_t} \|x_t\|_{V_t^{-1}}\}$$

$$\leq 2\sqrt{\beta_t} \cdot \min\{1, \|x_t\|_{V_t^{-1}}\}$$

Assume $\max_x \|x\|_2 \leq L$.

"Eliptic potential Lemma" $\quad V_0 := \lambda I$

$$\sum_{t=1}^{T} \min\{1, \|x_t\|^2_{V_t^{-1}}\} \leq 2 \log\left(\frac{|V_T|}{|V_0|}\right) \leq 2 d \log\left(\frac{d\lambda + TL^2}{d\lambda}\right)$$

$$\Rightarrow \text{Regret} \quad \sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle \leq \sqrt{\left(\sum_{t=1}^{T} 1^2\right)\left(\sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle^2\right)}$$

$$\leq \sqrt{T \cdot 4\beta_T \cdot 2d \log\left(\frac{d\lambda + TL^2}{d\lambda}\right)}$$

$$= \sqrt{8 d T \beta_T \log\left(\frac{d\lambda + TL^2}{d\lambda}\right)}$$

Claim: We can take $\sqrt{\beta_t} \leq \|\theta_*\| \sqrt{\lambda} + \sqrt{2\log(1/\delta) + d\log\left(\frac{d\lambda + TL^2}{d\lambda}\right)}$

take $\delta = 1/T$

$$\Rightarrow \mathbb{E}\left[\sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle\right] \leq c\, d\sqrt{T} \log\left(\frac{d\lambda + TL^2}{d\lambda}\right)$$

Plays $x_t = \underset{x \in \mathcal{X}}{\arg\max}\ UCB_t(x)$ $\qquad UCB_t(x) := \max_{\theta \in C_{t-1}} \langle x, \theta \rangle$

$$C_t = \{ \theta : \|\theta - \hat{\theta}_{t-1}\|^2_{V_{t-1}} \le \beta_{t-1} \}$$

Then $$UCB_t(x) = \langle \hat{\theta}_{t-1}, x \rangle + \sqrt{\beta_t}\, \|x\|_{V_{t-1}^{-1}}$$