Recap Lipschitz bandits

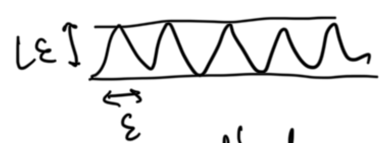
$$|\mu(x) - \mu(x')| \le L \underbrace{\|x - x'\|}_{D(x,x')} \quad \forall\, x, x' \in \mathcal{X}$$

① Fixed discretization $\quad E[R(T)] \asymp T^{\frac{d+1}{d+2}}$ (ignoring $\log T, L$)

simple reduction to multiarm bandit $\qquad$ d - dim of $\mathcal{X}$

② Adaptive discretization $\quad E[R(T)] \asymp \underline{T^{\frac{d'+1}{d'+2}}}$ $\quad$ d' - zooming dim of $\mathcal{X}$

(Activate arm not covered by confidence ball of current active set of arms)



$d' = d$

$d' < d$

max reward { Cumulative regret $\quad \underline{\sum_t} \mu(x^*) - \mu(x_t) \qquad \cdot\, T \cdot \underline{T^{-\frac{1}{d'+2}}}$

Simple regret $\quad \mu(x^*) - \mu(x_t) \qquad \cdot\, T^{-\frac{1}{d'+2}}$ $\Big\}\, d'$

reward everywhere { global $\sqrt{MSE}$ (supervised regression) $\quad \sqrt{\int (\mu(x) - \hat{\mu}_n(x))^2 dx} \qquad \underline{n^{-\frac{1}{d+2}}} \quad - d$

Linear Bandits

Player chooses action x_t

observes $y_t = \langle x_t, \theta_* \rangle + \eta_t \qquad \eta_t$ zero mean subgaussian

$$E[y_t] = \mu_t = \langle x_t, \theta_* \rangle$$

Minimize cumulative regret $\quad E[R(T)] = E\left[\sum_{t=1}^{T} \mu(x^*) - \mu(x_t)\right]$

$$= E\left[\sum_{t=1}^{T} \langle x^*, \theta^* \rangle - \langle x_t, \theta_t \rangle\right]$$

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

$$\hat{\theta}_t = \arg\min_{\theta} \sum_{s=1}^{t-1} (\langle x_s, \theta \rangle - y_s)^2 + \lambda \|\theta\|_2^2 \qquad \text{Ridge regression}$$

$$= (X^T X + \lambda I)^{-1} X^T Y \qquad X = \begin{bmatrix} \vdots \\ x_s^T \\ \vdots \end{bmatrix} \quad Y = \begin{bmatrix} \vdots \\ y_s \\ \vdots \end{bmatrix}$$

$$= V_{t-1}^{-1} X^T Y$$

$$V_0 = \lambda I$$
$$V_t = V_{t-1} + x_t x_t^T$$
$$= \lambda I + \sum_{s=1}^{t} x_s x_s^T$$

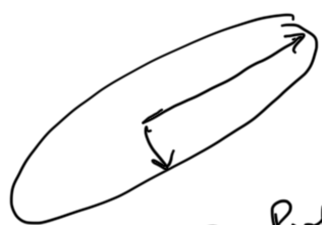
computed sequentially

Confidence set

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

(A1) β_t increasing seq^n ≥ 1 (A2) whp $\theta^* \in C_t$ $\equiv |\hat{\mu}_t - \mu| \le \varepsilon_t$

$\beta_t = d$



→ evecs of V_{t-1} ← covar of data seen

evals of V_{t-1}

≈ Prob of a point being drawn from same distn as prev data

Linear UCB

Play $x_t = \arg\max_x UCB_t(x)$ where $UCB_t(x) = \max_{\theta \in C_t} \langle x, \theta \rangle$

$\equiv \hat{\mu}(x) + \varepsilon_t(x)$

Observe y_t, update $\hat{\theta}_t$.

(A3) $\max_x \langle x_* - x, \theta_* \rangle \le 1$ (A4) $\|x\|_2 \le L$

Thm: If (A1) - (A4) hold, whp (hint: is $\sqrt{\frac{d}{n}}$)

$$E[R(T)] = O(d\sqrt{T}\log T)$$

Proof: $\tilde{\theta}_t = \arg\max_{\theta \in C_t} \langle x_t, \theta \rangle$

$$\langle x_*, \theta_* \rangle \le UCB_t(x_*) \le UCB_t(x_t) = \langle x_t, \tilde{\theta}_t \rangle$$

UCB sampling

Regret at time t $\langle x_* - x_t, \theta_* \rangle = \langle x_*, \theta_* \rangle - \langle x_t, \theta_* \rangle$

$$\leq \langle x_t, \tilde{\theta}_t \rangle - \langle x_t, \theta_* \rangle$$

$$= \langle x_t, \hat{\theta}_t - \theta_* \rangle$$

$$= x_t^T V_{t-1}^{-1/2} \underbrace{V_{t-1}^{1/2} (\tilde{\theta}_t - \theta_*)}$$

Cauchy-Schwartz $\leq \|x_t\|_{V_{t-1}^{-1}} \|\tilde{\theta}_t - \theta_*\|_{V_{t-1}}$

$$\leq \|x_t\|_{V_{t-1}^{-1}} \left(\underbrace{\|\tilde{\theta}_t - \hat{\theta}_t\|_{V_{t-1}}}_{\tilde{\theta}_t \in C_t} + \underbrace{\|\hat{\theta}_t - \theta_*\|_{V_{t-1}}}_{\theta_* \in C_t} \right)$$

whp A2

$$\leq 2\sqrt{\beta_t} \, \|x_t\|_{V_{t-1}^{-1}}$$

$$\langle x_* - x_t, \theta_* \rangle \leq \min(1, 2\sqrt{\beta_t} \, \|x_t\|_{V_{t-1}^{-1}}) \quad \text{A3} \quad \star$$

Elliptic potential lemma

$$\rightarrow \sum_{t=1}^{T} \min(1, \|x_t\|^2_{V_{t-1}^{-1}}) \leq 2 \log \frac{|V_T|}{|V_0|} \quad \approx O(d)$$

$u \geq 0$

$\min(u, 1) \leq 2 \ln(1+u)$

Proof: $\text{LHS} \leq 2 \sum_{t=1}^{T} \log(1 + \|x_t\|^2_{V_{t-1}^{-1}})$

$$V_t = V_{t-1} + x_t x_t^T$$

$$= V_{t-1}^{1/2} \left(I + V_{t-1}^{-1/2} x_t x_t^T V_{t-1}^{-1/2} \right) V_{t-1}^{1/2}$$

$\det(AB) = \det(A)\det(B)$

$$|V_t| = |V_{t-1}| \, |I + \underbrace{V_{t-1}^{-1/2} x_t x_t^T V_{t-1}^{-1/2}}| \quad I + zz^T$$

$$= |V_{t-1}| \left(1 + \|x_t\|^2_{V_{t-1}^{-1}} \right) \cdot \underbrace{1 \cdots 1}_{d-1} \quad \text{---} \#$$

$$|V_T| = |V_0| \underbrace{\prod_{t=1}^{T} (1 + \|x_t\|^2_{V_{t-1}^{-1}})}$$

$$\log \frac{|V_T|}{|V_0|} = \sum_{t=1}^{T} \log(1 + \|x_t\|^2_{V_{t-1}^{-1}})$$

$$|V_0| = |\lambda I| = \lambda^d$$

$$|V_T| = \prod_{i=1}^{d} \lambda_i \le \Big(\frac{1}{d}\underbrace{\mathrm{tr}(V_T)}_{=\sum_i \lambda_i}\Big)^d \le \left(\frac{\overbrace{\mathrm{tr}(V_0)}^{=\lambda d} + TL^2}{d}\right)^d =$$

geom. mean ≤ arith mean

$$\Rightarrow \log \frac{|V_T|}{|V_0|} \le \log\left(\frac{\lambda d + TL^2}{d\lambda}\right)^d = d \log\left(\frac{\lambda d + TL^2}{\lambda d}\right) \approx O(d)$$

$$R(T) = \sum_{t=1}^{T} \langle x_* - x_t, \theta^* \rangle \le \sqrt{\underbrace{\sum_{t=1}^{T} 1^2}_{} \ \sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle^2}$$

$$= \sqrt{T} \cdot \sqrt{\sum_{t=1}^{T} \langle x_* - x_t, \theta_* \rangle^2}$$

$$\overset{\star}{\le} \sqrt{T} \sqrt{\sum_{t=1}^{T} 4\beta_t \min(1, \|x_t\|^2_{V_{t-1}})}$$

(A1) β_t is increasing ≥ 1

$$= O\left(\sqrt{T\beta_T \, 8d \log \frac{d\lambda + TL^2}{d\lambda}}\right)$$ (A4)

$$\lesssim \tilde{O}\left(\sqrt{T\beta_T d}\right) \approx \tilde{O}(d\sqrt{T})$$

b/c (A2) holds if $\beta_T \approx d$

Cumulative regret $\tilde{O}(d\sqrt{T}) = T \cdot \frac{d}{\sqrt{T}}$

Supervised regression $\sqrt{MSE} = O\left(\sqrt{\frac{d}{n}}\right)$

$\sqrt{d}$ price for exploration (regret incurred) over a continuous space

if linear reward over K actions (no continuous action space)

cumulative regret $\sqrt{\min(d, K)\, T}$

$\sqrt{KT}$

$$UCB_t(x) = \max_{\theta \in C_t} \langle x, \theta \rangle \qquad C_t = \{\theta : \|\theta - \hat\theta_t\|^2_{V_{t-1}} \le \beta_t\}$$

$$UCB'_t(x) = \underbrace{\langle \hat\theta_t, x \rangle}_{\substack{\text{reward estimate} \\ \hat\mu_t(x)}} + \underbrace{\sqrt{\beta_t}\, \|x\|_{V_{t-1}^{-1}}}_{\substack{\text{deviation} \\ \varepsilon_t(x)}}$$

implementation preference

$$\langle x, \theta \rangle = \langle x, \hat\theta_t \rangle + \langle x, \theta - \hat\theta_t \rangle$$

$$\le \langle x, \hat\theta_t \rangle + \|x\|_{V_{t-1}^{-1}} \|\theta - \hat\theta_t\|_{V_{t-1}}$$

$$\le \langle x, \hat\theta_t \rangle + \|x\|_{V_{t-1}^{-1}} \sqrt{\beta_t} \qquad \text{if } \theta \in C_t$$

$$\Rightarrow \underbrace{\max_{\theta \in C_t} \langle x, \theta \rangle}_{= UCB_t(x)} \le \langle x, \hat\theta_t \rangle + \sqrt{\beta_t}\, \|x\|_{V_{t-1}^{-1}} = UCB'_t(x)$$

(A2) whp $\theta^* \in C_t$ if $\beta_t \approx d$

$$\|\theta^* - \hat\theta_t\|_{V_{t-1}} \le \sqrt{\beta_t}$$

Martingale Bernstein (Freedman)

Proof:
$$Z_t = \|\hat\theta_t - \theta^*\|^2_{V_{t-1}} = (\hat\theta_t - \theta^*)^T V_{t-1} (\hat\theta_t - \theta^*)$$

whp $\ge 1-\delta$ $\quad Z_t \le \beta_t = O(d)$

Step 1
$$Z_{t+1} \le Z_t + 2\eta_t \frac{x_t^T(\hat\theta_t - \theta^*)}{1 + x_t^T V_{t-1}^{-1} x_t} + \eta_t^2 \frac{x_t^T V_{t-1}^{-1} x_t}{1 + x_t^T V_{t-1}^{-1} x_t}$$

To show

$$\Rightarrow Z_t \le Z_1 + 2 \sum_{s=1}^{t} \eta_s \frac{x_s^T(\hat\theta_s - \theta^*)}{1 + x_s^T V_{s-1}^{-1} x_s} + \sum_{s=1}^{t} \eta_s^2 \frac{x_s^T V_{s-1}^{-1} x_s}{1 + x_s^T V_{s-1}^{-1} x_s}$$

Term 1
$$Z_1 = \|\hat\theta_1 - \theta^*\|^2_{V_0} \le \|\theta^*\|^2 = O(d)$$

$\hat\theta_1 = 0$ initialization

Term 2 $\quad M_s = \eta_s \dfrac{x_s^T(\hat{\theta}_s - \theta^*)}{1 + x_s^T V_{s-1}^{-1} x_s} \mathbb{1}\{Z_s \le \beta_s \ \forall s \ge 1\}$ (HW)

is a martingale diff^n seq^n & use Freedman's ineq.

Term 3 $\quad \displaystyle\sum_{s=1}^{t} \eta_s^2 \frac{x_s^T V_{s-1}^{-1} x_s}{1 + x_s^T V_{s-1}^{-1} x_s}$ $\qquad |\eta_s| \le 1 \qquad \underline{\dfrac{a}{1+a} \le \min(1, a)}$

$$\le \sum_{s=1}^{t} \min(1, x_s^T V_{s-1}^{-1} x_s)$$

$a < 1$

$a > 1$

$$\le \sum_{s=1}^{t} 2 \ln(1 + x_s^T V_{s-1}^{-1} x_s)$$

$$= 2 \ln \prod_{s=1}^{t} (1 + x_s^T V_{s-1}^{-1} x_s)$$

$$= 2 \ln |V_t| / |V_0| \qquad \text{from } \#$$

$$\le 2 d \ln t$$

$\dfrac{|V_t|}{|V_0|} = O(d \ln t)$ as above.

Step 1 decomposition into 3 terms:

$$Z_{t+1} = (\hat{\theta}_{t+1} - \theta^*)^T V_t (\hat{\theta}_{t+1} - \theta^*)$$

$$= (V_t (\hat{\theta}_{t+1} - \theta^*))^T V_t^{-1} (V_t (\hat{\theta}_{t+1} - \theta^*))$$

$$= (V_{t-1}(\hat{\theta}_t - \theta^*) + \eta_t x_t)^T V_t^{-1} (V_{t-1}(\hat{\theta}_t - \theta^*) + \eta_t x_t)$$

$\because \ V_t (\hat{\theta}_{t+1} - \theta^*) = V_{t-1}(\hat{\theta}_t - \theta^*) + \eta_t x_t$ since

by def^n of $\hat{\theta}_{t+1}$, $\hat{\theta}_t$
$$\begin{cases} V_t \hat{\theta}_{t+1} = \displaystyle\sum_{s=1}^{t} x_s y_s = \sum_{s=1}^{t} x_s (x_s^T \theta^* + \eta_s) = V_t \theta^* + \sum_{s=1}^{t} x_s \eta_s \\ V_{t-1} \hat{\theta}_t = V_{t-1} \theta^* + \displaystyle\sum_{s=1}^{t-1} x_s \eta_s \end{cases}$$

$$V_t = V_{t-1} + x_t x_t^T$$

$$\Rightarrow V_t^{-1} = V_{t-1}^{-1} - \frac{V_{t-1}^{-1} x_t x_t^T V_{t-1}^{-1}}{1 + x_t^T V_{t-1}^{-1} x_t}$$

$$\Rightarrow Z_{t+1} = Z_t - \left(V_{t-1}(\hat{\theta}_t - \theta^*) + \eta_t x_t\right)^T \frac{V_{t-1}^{-1} x_t x_t^T V_{t-1}^{-1}}{1 + x_t^T V_{t-1}^{-1} x_t} \left(V_{t-1}(\hat{\theta}_t - \theta^*) + \eta_t x_t\right)$$

$$= Z_t - \underbrace{\left(V_{t-1}(\hat{\theta}_t - \theta^*)\right)^T \frac{V_{t-1}^{-1} x_t x_t^T V_{t-1}^{-1}}{1 + x_t^T V_{t-1}^{-1} x_t} V_{t-1}(\hat{\theta}_t - \theta^*)}_{\geq 0}$$

$$+ 2\eta_t x_t^T \left(V_{t-1}^{-1} - \frac{V_{t-1}^{-1} x_t x_t^T V_{t-1}^{-1}}{1 + x_t^T V_{t-1}^{-1} x_t}\right) V_t (\hat{\theta}_t - \theta^*)$$

$$+ \eta_t^2 x_t^T \left(V_{t-1}^{-1} - \frac{V_{t-1}^{-1} x_t x_t^T V_{t-1}^{-1}}{1 + x_t^T V_{t-1}^{-1} x_t}\right) x_t$$

$$\leq Z_t + 2\eta_t \left(1 - \frac{x_t^T V_{t-1}^{-1} x_t}{1 + x_t^T V_{t-1}^{-1} x_t}\right) x_t^T (\hat{\theta}_t - \theta^*)$$

$$+ \eta_t^2 \left(x_t^T V_{t-1}^{-1} x_t - \frac{(x_t^T V_{t-1}^{-1} x_t)^2}{1 + x_t^T V_{t-1}^{-1} x_t}\right)$$