CS861: Theoretical Foundations of Machine Learning Lecture 19 - 16/10/2024
University of Wisconsin–Madison, Fall 2024

Lecture 19: Linear bandits (cont'd), Martingale concentration

Lecturer: Kirthevasan Kandasamy *Scribed by: Lakshika Rathi*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the instructor.*

In this lecture, we will continue discussing `linear bandits`, then discuss `Martingale Concentration`.

We start by Claim 3 in the LinUCB proof.

Claim 3. $\sum_{t=1}^{T} \min\left(1, \|A_t\|_{V_{t-1}}\right) \leq 2d \log\left(1 + \frac{TL^2}{d\lambda}\right)$

Proof Recall that, $V_t = \lambda I + \sum_{s=1}^{t} A_s A_s^T$.

$$\begin{aligned}
\det(V_t) &= \det(V_{t-1} + A_t A_t^T) \\
&= \det(V_{t-1}^{\frac{1}{2}}(I + V_{t-1}^{-\frac{1}{2}} A_t A_t^T V_{t-1}^{-\frac{1}{2}}) V_{t-1}^{\frac{1}{2}}) \\
&= \det(V_{t-1}) \det(I + (V_{t-1}^{-\frac{1}{2}} A_t)(V_{t-1}^{-\frac{1}{2}} A_t)^T) \text{ , since } \det(AB) = \det(A)\det(B) \\
&= \det(V_{t-1})(1 + ||A_t||^2_{V_{t-1}^{-1}}), \text{ since } \det(I + UV^T) = 1 + U^T V
\end{aligned} \tag{1}$$

Thus, $\det(V_t) = \det(V_{t-1})(1 + ||A_t||^2_{V_{t-1}^{-1}})$. Therefore, $\det(V_T) = \det(\lambda I) \prod_{t=1}^{T}(1 + ||A_t||^2_{V_{t-1}^{-1}})$. This means that

$$\log\left(\frac{\det(V_T)}{\lambda^d}\right) = \sum_{t=d+1}^{T} \log\left(1 + ||A_t||^2_{V_{t-1}^{-1}}\right)$$

We know,

$$\begin{aligned}
\det(V_T) = \prod_{i=1}^{d} eig_1(V_T) &\leq \left(\frac{1}{d}\sum_{i=1}^{d} eig_1(V_T)\right) \text{ by AM-GM inequality} \\
&= \frac{1}{d^d}\operatorname{trace}(V_T)^d = \frac{1}{d^d}\left(\operatorname{trace}(\lambda I) + \sum_{s=1}^{T}\operatorname{trace}(A_S A_S^T)\right)^d \\
= \frac{1}{d^d}\left(d\lambda + \sum_{s=1}^{T} A_s^T A_s\right)^d &\leq \frac{1}{d^d}\left(d\lambda + \sum_{s=1}^{T} L^2\right)^d = \frac{1}{d^d}\left(d\lambda + TL^2\right)^d
\end{aligned}$$

We will now use the following inequality: $x \leq 2\log(1+x),\ \forall x \in [0, 2\log(2)]$ and the above result, to get

$$\begin{aligned}
\sum_{t=d+1}^{T} \min(1, ||A_t||^2_{V_{t-1}}) &\leq 2\sum_{t=d+1}^{T} \log\left(1 + \min\left(1, ||A_t||^2_{V_{t-1}}\right)\right) \\
&\leq 2\sum_{t=d+1}^{T} \log\left(1 + ||A_t||^2_{V_{t-1}^{-1}}\right)
\end{aligned}$$

$$\leq 2\log(\frac{1}{\lambda^d}\frac{(d\lambda+TL^2)^d}{d^d})$$
$$= 2d\log(1+\frac{TL^2}{d\lambda})$$

We now prove the claim $\mathbb{P}(G^c) \leq 1/T$, where $G = \left\{ \left|\theta_*^T a - \hat{\theta}_t^T a\right| \leq \beta_t \|a\|_{V_t^{-1}}, \forall a \in \mathcal{A}, \forall t \in \{T\} \right\}$

Proof $\mathbb{P}(G^c) \leq \sum_{t=1}^{T} \mathbb{P}\left(\left|\theta_*^T a - \hat{\theta}_t^T a\right| > \beta_t \|a\|_{V_t^{-1}} \text{for any } a \in \mathcal{A} \right)$
Now, Consider any $a \in A$,

$$|\theta_\star^\top a - \hat{\theta}_t^\top a| = |(\theta_\star - \hat{\theta}_t)^\top a| = |((\theta_\star - \hat{\theta}_t)V_t^{1/2})^\top (V_t^{-1/2}a)|$$
$$\leq \|(\theta_\star - \hat{\theta}_t)V_t^{1/2}\|_2 \cdot \|V_t^{-1/2}a\|_2 \quad \text{by Cauchy-Schwarz}$$
$$= \|\theta_\star - \hat{\theta}_t\|_{V_t} \cdot \|a\|_{V_t^{-1}} \quad \text{as } \|xQ^{1/2}\|_2^2 = x^\top Q x = \|x\|_Q \text{ for symmetric Q.}$$

Therefore, it is sufficient to show $\|\theta_\star - \hat{\theta}_t\|_{V_t} \leq \beta_t$ with probability $\geq 1 - \frac{1}{T^2}$.

$$\hat{\theta}_t = V_t^{-1}\sum_{s=1}^{t} A_s X_s, \quad \text{where,} \quad V_t = \lambda I + \sum_{s=1}^{t} A_s A_s^\top, \quad X_s = \theta_\star^\top A_s + \epsilon_s.$$

Let $W_t = \sum_{s=1}^{t} A_s A_s^\top$ so that $V_t = \lambda I + W_t$.
Now write,

$$\hat{\theta}_t = V_t^{-1}\sum_{s=1}^{t} A_s(A_s^\top \theta_\star + \epsilon_s) = V_t^{-1}W_t\theta_\star + V_t^{-1}\sum_{s=1}^{t} A_s\epsilon_s \triangleq V_t^{-1}W_t\theta_\star + V_t^{-1}\xi_t.$$

Therefore,
$$\hat{\theta}_t - \theta_\star = (V_t^{-1}W_t - I)\theta_\star + V_t^{-1}\xi_t.$$

Hence,

$$\|\hat{\theta}_t - \theta_\star\|_{V_t} \leq \|(V_t^{-1}W_t - I)\theta_\star\|_{V_t} + \|V_t^{-1}\xi_t\|_{V_t} \quad \text{(triangle inequality)}$$
$$= \|(V_t^{-1}W_t - I)\theta_\star\|_{V_t} + \|\xi_t\|_{V_t^{-1}} \quad \text{as } \|V_t^{-1}\xi_t\|_{V_t} = \xi_t^\top V_t^{-1}V_tV_t^{-1}\xi_t.$$

The following calculations show $\|(V_t^{-1}W_t - I)\theta_\star\|_{V_t} \leq \sqrt{\lambda}B$,

$$\|(V_t^{-1}W_t - I)\theta_\star\|_{V_t}^2 = \theta_\star^\top (V_t^{-1}W_t - I)V_t(V_t^{-1}W_t - I)\theta_\star$$
$$= \theta_\star^\top (V_t^{-1}W_t - I)(W_t - V_t)\theta_\star = \lambda\theta_\star^\top (I - V_t^{-1}W_t)\theta_\star \leq \lambda\theta_\star^\top\theta_\star \leq \lambda B^2.$$

Recall, we need to show, that the following holds with probability $\geq 1 - \frac{1}{T^2}$

$$\|\theta_\star - \hat{\theta}_t\|_{V_t} \leq \beta_t = \max\left(\frac{1}{2}, \sqrt{\lambda}B + \sigma\sqrt{d\log(dT^2)(2+\log(1+tL^2))}\right)$$

Therefore, it is sufficient to show that the following holds with probability $\geq 1 - \frac{1}{T^2}$

$$\|\xi_t\|_{V_t^{-1}} \leq \sigma\sqrt{d}\sqrt{\log(dT^2)(2+\log(1+tL^2))}$$

Let us write,

$$
\begin{aligned}
\|\xi_t\|^2_{V_t^{-1}} &= \xi_t^\top V_t^{-1}\xi_t = \xi_t^\top V_t^{-1/2} \cdot I \cdot V_t^{-1/2}\xi_t \\
&= \xi_t^\top V_t^{-1/2}\left(\sum_{i=1}^d e_i e_i^\top\right) V_t^{-1/2}\xi_t \\
&= \sum_{i=1}^d \xi_t^\top V_t^{-1/2} e_i e_i^\top V_t^{-1/2}\xi_t \\
&= \sum_{i=1}^d (\xi_t^\top V_t^{-1/2} e_i)^2.
\end{aligned}
$$

We need to show, the following holds with probability $\geq 1 - \frac{1}{T^2}$

$$
\|\xi_t\|^2_{V_{t-1}^{-1}} \leq \sigma^2 d \log(dT^2)\left(2 + \log(1 + tL^2)\right) \tag{*}
$$

By a union bound,

$$
\begin{aligned}
\mathbb{P}\left(\|\xi_t\|^2_{V_{t-1}^{-1}} > \sigma^2(*)\right) &\leq \sum_{i=1}^d \mathbb{P}\left(\left(\xi_t^T V_t^{-1/2} e_i\right)^2 > \sigma^2(*)\right) \\
&= \sum_{i=1}^d \mathbb{P}\left(\frac{\xi_t^T V_t^{-1/2} e_i}{\sigma} > \sqrt{(*)}\right)
\end{aligned}
$$

Therefore, sufficient to show, for $a \in \{V_t^{-1/2}e_1, \ldots, V_t^{-1/2}e_d\}$,

$$
\mathbb{P}\left(\frac{\xi_t^T a}{\sigma} > \sqrt{(*)}\right) \leq \frac{1}{T^2 d}.
$$

We can summarize our proof of Claim 1 so far as follows:

1. (Prediction to estimation error) Sufficient to show, the following holds with probability $\geq 1 - \frac{1}{T^2}$,

$$
|\theta_* - \hat{\theta}_t|_{V_t} \leq \beta_t = \max\left(\frac{1}{2}, \sqrt{\Lambda B} + \sigma\sqrt{d\log(dT^2)(2 + \log(1 + tL^2))}\right)
$$

2. Sufficient to show, the following holds with probability $\geq 1 - \frac{1}{T^2}$,

$$
\|\xi_t\|_{V_{t-1}^{-1}} \leq \sigma\sqrt{d}\sqrt{\log(dT^2)}(2 + \log(1 + tL^2)).
$$

3. Sufficient to show, for $a \in \{V_t^{-1/2}e_1, \ldots, V_t^{-1/2}e_d\}$,

$$
\mathbb{P}\left(\frac{\xi_t^T a}{\sigma} > \sqrt{\log(dT^2)(2 + \log(1 + tL^2))}\right) \leq \frac{1}{T^2 d}.
$$

We know that $\xi_t = \sum_{s=1}^t A_s \epsilon_s$. We will show, for any $a \in \{V_t^{-1/2}e_1, \ldots, V_t^{-1/2}e_d\}$,

$$
\mathbb{P}\left(\frac{|\xi_t^\top a|}{\sigma} > \sqrt{\log(dT^2)(2 + \log(1 + tL^2))}\right) \leq \frac{1}{Td}.
$$

Key challenge: The actions and observations are not independent in here. We will instead use the fact that $\xi_t^\top a$ is a *martingale*. We can use a variety of martingale concentration results to obtain the above result.

Definition (Martingale). A sequence of random variables $\{Z_t\}_{t\in\mathbb{N}}$ is a martingale w.r.t another sequence $\{Y_t\}_{t\in\mathbb{N}}$ if $\mathbb{E}[Z_t|Y_{t-1}] = Z_{t-1}$ and $\mathbb{E}[|Z_t|] < \infty$, for all $t \in \mathbb{N}$

This proof technique is from Rusmevichientong and Tsitsiklis, 2008.

Lemma (Corollary 2.2 from de La Pena et al 2004). If A, B are random variables such that $\mathbb{E}\left[e^{\mu A - \frac{\mu^2 B^2}{2}}\right] \leq 1$ for all $\mu \in \mathbb{R}$, then for all $\tau \geq \sqrt{2}$, and $y > 0$, we have

$$\mathbb{P}\left(|A| \geq \tau\sqrt{(B^2+y)}\left(1+\frac{1}{2}\log\left(1+\frac{B^2}{y}\right)\right)\right) \leq e^{-\tau^2/2}.$$

Cf. If B is a constant and not a random variable, then the condition says that A is B-sub-Gaussian, $\mathbb{E}[e^{\mu A}] \leq e^{\frac{\lambda^2 B^2}{2}}$. In which case, we know $\mathbb{P}(A > B\tau) \leq e^{-\tau^2/2}$. Note that $\sqrt{(B^2+y)}\left(1+\log\left(1+\frac{B^2}{y}\right)\right) \asymp B$.

Lemma (Corollary 2.2 from de La Pena et al 2004). If A, B are random variables such that

$$\mathbb{E}\left[e^{\mu A - \frac{\mu^2 B^2}{2}}\right] \leq 1$$

for all $\mu \in \mathbb{R}$, then for all $\tau \geq \sqrt{2}$, and $y > 0$, we have

$$\mathbb{P}\left(|A| \geq \tau\sqrt{(B^2+y)}\left(1+\frac{1}{2}\log\left(1+\frac{B^2}{y}\right)\right)\right) \leq e^{-\tau^2/2}.$$

We will apply the above result with $A = \frac{a^\top \xi_t}{\sigma}$ and $B = \|a\|_{W_t}$, $y = \|a\|_2^2$, and $\tau = \sqrt{2\log(T^2 d)}$.

$$A = \frac{a^\top \xi_t}{\sigma} = \frac{1}{\sigma}\sum_{s=1}^{t} a^\top A_s \epsilon_s, \quad B^2 = a^\top W_t a = \sum_{s=1}^{t} a^\top A_s A_s^\top a.$$

Let us first check that the condition $\mathbb{E}\left[\exp\left(\mu A - \frac{\mu^2 B^2}{2}\right)\right] \leq 1$ holds. We can write,

$$\mu A - \frac{\mu^2}{2}B^2 = \sum_{s=1}^{t}\left(\frac{\mu}{\sigma}a^\top A_s \epsilon_s - \frac{\mu^2}{2}\left(a^\top A_s\right)^2\right) \triangleq Q_s.$$

We need to show,

$$\mathbb{E}\left[\exp\left(\mu A - \frac{\mu^2}{2}B^2\right)\right] = \mathbb{E}\left[\exp\sum_{s=1}^{t} Q_s\right] \leq 1.$$

Denote $\mathcal{F}_s \triangleq \{A_1, \epsilon_1, \ldots, A_{s-1}, \epsilon_{s-1}\}$. We will first bound, $\mathbb{E}[Q_s|\mathcal{F}_s]$,

$$\mathbb{E}[Q_s|\mathcal{F}_s] = \exp\left(\frac{\mu^2}{2}(a^\top A_s)^2\right)\mathbb{E}\left[\exp\left(\frac{\mu}{\sigma}(a^\top A_s)\epsilon_s\right)\Big|\mathcal{F}_s\right]$$

$$\leq \exp\left(\frac{\mu^2}{2}(a^\top A_s)^2\right)\exp\left(-\frac{\sigma^2}{2}\times\left(\frac{\mu}{\sigma}(a^\top A_s)\right)^2\right) \leq 1$$

As, given $\mathcal{F}_{s-1} = \{A_1, \epsilon_1, \ldots, A_{s-1}, \epsilon_{s-1}\}$, A_s is fixed and ϵ_s is σ-sub-Gaussian.

Therefore,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\mu A - \frac{\mu^2}{2}B^2\right)\right] &= \mathbb{E}\left[\exp\sum_{s=1}^{t} Q_s\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{s=1}^{t} Q_s \middle| \mathcal{F}_{t-1}\right]\right] \\
&= \mathbb{E}\left[\sum_{s=1}^{t-1} Q_s \mathbb{E}\left[e^{Q_t}\middle|\mathcal{F}_{t-1}\right]\right] \quad \text{as we have fixed } A_1, \epsilon_1, \ldots, A_{t-1}, \epsilon_{t-1} \\
&\leq \mathbb{E}\left[\exp\sum_{s=1}^{t-1} Q_s\right] \leq \cdots \leq 1.
\end{aligned}$$

This verifies the condition for the lemma.

Therefore, we have the following with probability at least $1 - \frac{1}{T^2 d}$,

$$\begin{aligned}
\frac{a^\top \xi_t}{\sigma} &\leq \sqrt{2\log(T^2 d)}\sqrt{(a^\top W_t a + a^\top a)\left(1 + \frac{1}{2}\log\left(1 + \frac{a^\top W_t a}{\|a\|_2^2}\right)\right)} \\
&\leq \sqrt{\log(T^2 d)(2 + \log(1 + \mathrm{eig}_1(W_t)))} \cdot \|a\|_{V_t}.
\end{aligned}$$

As $\mathrm{eig}_1(A) = \max_x \frac{x^\top A x}{x^\top x}$ and $a^\top W_t a + a^\top a = a^\top (W_t + I)a = a^\top V_t a$.

We showed that, for any $a \in \mathbb{R}^d$, with probability at least $1 - \frac{1}{T^2 d}$,

$$\frac{a^\top \xi_t}{\sigma} \leq \sqrt{\log(T^2 d)(2 + \log(1 + \mathrm{eig}_1(W_t)))} \cdot \|a\|_{V_t}.$$

The proof is completed by the following observations:

- When $a = V_t^{-1/2} e_i$, $\|a\|_{V_t}^2 = a^\top V_t a = e_i^\top V_t^{-1/2} V_t V_t^{-1/2} e_i = e_i^\top e_i = 1$.
- $\mathrm{eig}_1(W_t) \leq \mathrm{trace}(W_t) = \sum_{s=1}^{t}(A_s^\top A_s) \leq tL^2$, as $\max_{a\in\mathcal{A}} a^\top a \leq L^2$.

LinUCB: Proof summary

This general strategy works in other settings as well (GP bandits, generalized linear bandits).

1. First, consider the pseudo-regret, $\overline{R}_T = \sum_{t=1}^{T}(\theta_*^\top a_* - \theta_*^\top A_t)$.
2. Define a good event G, where the confidence intervals trap the true means. Then,

$$R_T = \mathbb{E}[\overline{R}_T | G]\mathbb{P}(G) + \mathbb{E}[\overline{R}_T | G^c]\mathbb{P}(G^c).$$

3. Use martingale concentration to bound $\mathbb{P}(G^c)$.
4. Under G, we can bound the instantaneous pseudo-regret

$$\theta_*^\top a_* - \theta_*^\top A_t \leq 2 \times \text{“conf-width of } A_t \text{ at round } t-1\text{”}.$$

Then bound the summation.