6.3 LinUCB Analysis

There are two main propositions that the upper bounds follow.

[Proposition 6.6] (Confidence)

Let $\delta > 0$. We have that $\Pr(\forall t, \mu^* \in \mathrm{BALL}_t) \geq 1-\delta$

[Proposition 6.7] (Sum of Squared Regret bounds)

Suppose that $\|x\| \leq B$ for $x \in D$. Suppose that β_t is $\uparrow$ and $\beta_t \geq 1$. For LinUCB, if $\mu^* \in \mathrm{BALL}_t$ for all t, then

$$\sum_{t=0}^{T-1} \mathrm{reget}_t^2 \leq 8 \beta_T d \log\left(1 + \frac{TB^2}{d\lambda}\right)$$

where $\mathrm{reget}_t = \mu^* \cdot x^* - \mu^* \cdot x_t$ (the instantaneous regret)

We will prove the two propositions later. First we use the two props. to prove Th. 6.3.

[Th 6.3] Suppose $|\mu^* \cdot x| \leq 1$ for all $x \in D$, $\|\mu^*\| \leq W$ and $\|x\| \leq B$ for all $x \in D$, and that η_t is σ^2 sub-Guassian. Set $\lambda = \sigma^2 / W^2$, $\beta_t := \sigma^2 \left(2 + 4 d \log\left(1 + \frac{tB^2W^2}{d}\right) + 8 \log(4/\delta)\right)$, we have that with prob. $\geq 1-\delta$, for all $T \geq 0$,

$$R_T \leq c \sigma \sqrt{T} \left(d \log\left(1 + \frac{TB^2W}{d\sigma^2}\right) + \log(4/\delta)\right)$$

where c is an absolute constant.

Pf: By prop 6.6 and prop 6.7, with Cauchy-Schwartz ineq., we have:

$$R_T = \sum_{t=0}^{T-1} \mathrm{reget}_t \leq \sqrt{T \sum_{t=0}^{T-1} \mathrm{reget}_t^2} \leq \sqrt{8 T \beta_T d \log\left(1 + \frac{TB^2}{d\lambda}\right)}$$

$$\leq \left[8 T d \log\left(1 + \frac{TB^2W^2}{d\sigma^2}\right) \sigma^2 \left(2 + 4 d \log\left(1 + \frac{TB^2W^2}{d}\right) + 8 \log(4/\delta)\right)\right]^{\frac{1}{2}}$$

$$= 2\sigma\sqrt{Td} \left(4 \log\left(1 + \frac{TB^2W^2}{d\sigma^2}\right)\right)^{\frac{1}{2}} \left(1 + 2 d \log\left(1 + \frac{TB^2W^2}{d}\right) + 4 \log(4/\delta)\right)^{\frac{1}{2}}$$

$$\leq \sigma\sqrt{Td} \left[4 \log\left(1 + \frac{TB^2W^2}{\cdots}\right) + 1 + 2 d \log\left(1 + \frac{TB^2W^2}{\cdots}\right) + 4 \log(4/\delta)\right]$$

$$\le \sigma\sqrt{Td}\left[4\log\left(1+\frac{TB^2W^2}{d\sigma^2}\right) + 1 + 2d\log\left(1+\frac{TB^2W^2}{d}\right) + 4\log(4/\delta)\right]$$

$$\le \sigma\sqrt{Td}\,(4+1+dm)\left[\log\left(1+\frac{TB^2W^2}{d\sigma^2}\right) + \log(4/\delta)\right]$$

Let $c = (5+dm)\sqrt{d}$, we complete the proof.

Addition: Since $4/\delta \geq 4$, we have $1 \le \log\left(1+\frac{TB^2W^2}{d\sigma^2}\right) + \log(4/\delta)$

Since $2d\log\left(1+\frac{TB^2W^2}{d}\right) = d\log\left(\left(1+\frac{TB^2W^2}{d}\right)^2\right) \le d\log\left(1+\frac{TB^2W^2}{d\sigma^2}\right) + d\log(1+\sigma^2)$

and $\log(1+\sigma^2) = \frac{\log(1+\sigma^2)}{\log(4/\delta)}\log(4/\delta)$, let $m = \max\left\{1, \frac{\log(1+\sigma^2)}{\log(4/\delta)}\right\}$

then we have $2d\log\left(1+\frac{TB^2W^2}{d}\right) \le dm\left(\log\left(1+\frac{TB^2W^2}{d\sigma^2}\right) + \log(4/\delta)\right)$

which implies c, m are constant independently with T. $\square$

Now we turn back to the proofs of these two props.

6.3.1 Regret Analysis.

Goal: prove Prop. 6.7

Recape: $\mathrm{BALL}_t = \{\mu \mid (\hat{\mu}_t - \mu)'\Sigma_t(\hat{\mu}_t - \mu) \le \beta_t\}$

where $\Sigma_t = \lambda I + \sum_{\tau=0}^{t-1} x_\tau x_\tau'$ with $\Sigma_0 = \lambda I$.

> [Lemma 6.8] Let $x \in D$. If $\mu \in \mathrm{BALL}_t$, then
> $$|(\mu - \hat{\mu}_t)'x| \le \sqrt{\beta_t x'\Sigma_t^{-1}x}$$

Pf:
$$\begin{aligned}
|(\mu-\hat{\mu}_t)'x| &= |(\mu-\hat{\mu}_t)'\Sigma_t^{1/2}\Sigma_t^{-1/2}x| \\
&= |(\Sigma_t^{1/2}(\mu-\hat{\mu}_t))'\Sigma_t^{-\frac{1}{2}}x| \\
&\le \|\Sigma_t^{1/2}(\mu-\hat{\mu}_t)\|\,\|\Sigma_t^{-\frac{1}{2}}x\| \\
&= \|\Sigma_t^{1/2}(\mu-\hat{\mu}_t)\|\sqrt{x'\Sigma_t^{-1}x} \\
&\le \sqrt{\beta_t x'\Sigma^{-1}x} \qquad \text{since } \mu \in \mathrm{BALL}_t
\end{aligned}$$
$\square$

To show the upper bound for the instantaneous regret, we define

$$w_t := \sqrt{x_t'\Sigma_t^{-1}x_t}$$

which we interpret as the "normalized width" at time t.

> **[Lemma 6.9]** Fix $t \le T$. If $\mu^* \in BALL_t$ and $\beta_t \geq 1$, then
> $$\text{regret}_t \le 2\min(\sqrt{\beta_t w_t}, 1) \le 2\sqrt{\beta_t}\min(w_t, 1)$$

Pf: Let $\tilde{\mu} \in BALL_t$ s.t. $\hat{\mu}$ maximises $\tilde{\mu}' x_t$.

$$\tilde{\mu}' x_t = \max_{\mu \in BALL_t} \mu' x_t = \max_{x \in D} \max_{\mu \in BALL_t} \mu' x \underset{\Uparrow}{\geq} \mu^* \cdot x^*$$

where $x^* \in \arg\max_{x \in D} \mu^* \cdot x$. with the hypothesis $\mu^* \in BALL_t$.

Hence,

$$\begin{aligned}
\text{regret}_t &= \mu^* \cdot x^* - \mu^* \cdot x_t \\
&\le (\tilde{\mu} - \mu^*)' x_t \\
&= (\tilde{\mu} - \hat{\mu}_t)' x_t + (\hat{\mu}_t - \mu^*)' x_t \\
&\le 2\sqrt{\beta_t} w_t \quad \text{(by Lemma 6.8)}
\end{aligned}$$

Since $r_t \in [-1, 1]$, regret_t is always at most 2.

$$\begin{aligned}
\Rightarrow \text{regret}_t &\le 2\min\{\sqrt{\beta_t} w_t, 1\} \\
&\le 2\min\{\sqrt{\beta_t} w_t, \sqrt{\beta_t}\} \quad (\text{since } \beta_t \geq 1) \\
&= 2\sqrt{\beta_t}\min\{w_t, 1\}
\end{aligned}$$

□

The following two lemmas provide usefull tools to view the log determinant as a potential function, where can bound the sum of the width.

> **[Lemma 6.10].** We have
> $$\det \Sigma_T = \det \Sigma_0 \prod_{t=0}^{T-1} (1 + w_t^2).$$

Pf: By the definition of Σ_{t+1}, we have that

$$\det \Sigma_{t+1} = \det(\Sigma_t + x_t x_t')$$
$$= \det\left(\Sigma_t^{\frac{1}{2}} (I + \Sigma_t^{-\frac{1}{2}} x_t x_t' \Sigma_t^{-\frac{1}{2}}) \Sigma_t^{\frac{1}{2}} \right)$$
$$= \det(\Sigma_t) \det\left(I + (\Sigma_t^{-\frac{1}{2}} x_t)(\Sigma_t^{-\frac{1}{2}} x_t)' \right)$$
$$=: \det(\Sigma_t) \det(V_t V_t')$$

Now observe that $V_t' V_t = W_t^2$

$$(I + V_t V_t') V_t = V_t + V_t (V_t' V_t) = (1 + W_t^2) V_t$$

Hence, $(1+W_t^2)$ is an eigenvalues of $I + V_t V_t'$. Since $V_t V_t'$ is a rank one matrix, all other eigenvalues of $I + V_t V_t'$ $= 1$. (which can be seen by $|\lambda I - V_t V_t'| = (\lambda - W_t^2)\lambda^{n-1}$)

$\Rightarrow \det(I + V_t V_t') = 1 + W_t^2$.

$\Rightarrow \det(\Sigma_{t+1}) = \Sigma_0 \prod_{k=0}^{t} (1 + W_k^2)$ $\square$

[Lemma 6.11] ("Potential Function" Bound)

For any sequence $x_0, \cdots, x_{T-1}$ s.t. for $t < T$, $\|x_t\|_2 \le B$, we have

$$\log(\det \Sigma_{T-1} / \det \Sigma_0) = \log\det\left(I + \frac{1}{\lambda} \sum_{t=0}^{T-1} x_t x_t' \right) \le d \log\left(1 + \frac{TB^2}{d\lambda}\right)$$

Pf: Denote the eigenvalues of $\sum_{t=0}^{T-1} x_t x_t'$ as $\sigma_1, \cdots, \sigma_d$.

$$\sum_{i=1}^{d} \sigma_i = \mathrm{Trace}\left(\sum_{t=0}^{T-1} x_t x_t' \right) = \sum_{t=0}^{T-1} \|x_t\|^2 \le T B^2$$

$$\Rightarrow \log\det\left(I + \frac{1}{\lambda} \sum_{t=0}^{T-1} x_t x_t'\right) = \log\left(\prod_{i=1}^{d} (1 + \sigma_i/\lambda) \right)$$
$$= d \log\left(\left[\prod_{i=1}^{d} (1 + \sigma_i/\lambda) \right]^{\frac{1}{d}} \right)$$
$$\le d \log\left(\frac{1}{d} \sum_{i=1}^{d} (1 + \sigma_i/\lambda) \right)$$
$$\le d \log\left(1 + \frac{TB^2}{d\lambda}\right) \quad \square$$

Finally we prove that if μ^* alway stays in BALL_t, then our regret is under control, which is prop. 6.7.

[Proposition 6.7] Suppose $\|x\| \le B$, $\forall x \in D$. Suppose β_t is increasing

[Proposition] Suppose $\|x\| \le B$ $\forall x \in D$. Suppose β_t is increasing and $\beta_t \geq 1$. For LinUCB, if $\mu^* \in \text{BALL}_t$ $\forall t$, then

$$\sum_{t=0}^{T-1} \text{regret}_t^2 \leq 8\beta_t d \log\left(1+\frac{TB^2}{d\lambda}\right).$$

Pf: with the conditions, we have

$$\begin{aligned}
\sum_{t=0}^{T-1} \text{regret}_t^2 &\leq \sum_{t=0}^{T-1} 4\beta_t \min(w_t^2, 1) \leq 4\beta_T \sum_{t=0}^{T-1} \min(w_t^2, 1) \\
&\lesssim 8\beta_t \sum_{t=0}^{T-1} \ln(1+w_t^2) \quad (\text{for } 0 \leq y \leq 1,\ y \leq 2\ln(1+y)) \\
&= 8\beta_t \log(\det \Sigma_{T-1} / \det \Sigma_0) \quad (\text{by Lemma 6.10}) \\
&\leq 8\beta_t d \log\left(1+\frac{TB^2}{d\lambda}\right) \quad (\text{by lemma 6.11}) \quad \square
\end{aligned}$$

We now prove Prop. 6.6.

[Proposition 6.6] Let $\delta > 0$. We have that

$$\Pr(\forall t, \mu^* \in \text{BALL}_t) \geq 1-\delta$$

Pf: Since $r_\tau = x_\tau \cdot \mu^* + \eta_\tau$, we have

$$\begin{aligned}
\hat{\mu}_t - \mu^* &= \Sigma_t^{-1} \sum_{\tau=0}^{t-1} r_\tau x_\tau - \mu^* \quad (\text{by def of } \hat{\mu}_t) \\
&= \Sigma_t^{-1} \sum_{\tau=0}^{t-1} x_\tau (x_\tau \cdot \mu^* + \eta_\tau) - \mu^* \\
&= \Sigma_t^{-1} \left(\sum_{\tau=0}^{t-1} x_\tau x_\tau'\right) \mu^* - \mu^* + \Sigma_t^{-1} \sum_{\tau=0}^{t-1} \eta_\tau x_\tau \\
&= \Sigma_t^{-1} (\Sigma_t - \lambda I)\mu^* - \mu^* + \Sigma_t^{-1} \sum_{\tau=0}^{t-1} \eta_\tau x_\tau \\
&= -\lambda \Sigma_t^{-1} \mu^* + \Sigma_t^{-1} \sum_{\tau=0}^{t-1} \eta_\tau x_\tau
\end{aligned}$$

For any $0 < \delta_t < 1$,

$$\begin{aligned}
\sqrt{(\hat{\mu}_t - \mu^*)' \Sigma_t (\hat{\mu}_t - \mu^*)} &= \| \Sigma_t^{1/2} (\hat{\mu}_t - \mu^*) \| \\
&\leq \| \lambda \Sigma_t^{-\frac{1}{2}} \mu^* \| + \| \Sigma_t^{-\frac{1}{2}} \sum_{\tau=0}^{t-1} \eta_\tau x_\tau \|
\end{aligned}$$

With prob. $\geq 1-\delta_t$, the RHS satisfies

$$\leq \sqrt{\lambda} \|\mu^*\| + \underbrace{\sqrt{2\sigma^2 \log(\det(\Sigma_t)\det(\Sigma_0)^{-1}/\delta_t)}}_{(\text{by lemma 6.10})} \quad \cdots (*)$$

where we used $\|\Sigma_t^{-1}\| \leq 1/\lambda$ and the following lemma:

[Lemma A.9] for $t \geq 1$, with prob $> 1-\delta$ $$\Big\|\sum_{i=1}^{t} X_i \varepsilon_i\Big\|^2_{\Sigma_t^{-1}} \leq \sigma^2 \log\left(\frac{\det(\Sigma_t)\det(\Sigma_0)^{-1}}{\delta^2}\right)$$ where $\{\varepsilon_i\}$: martingale difference sequence and conditionally σ sub-Guassian $\{X_i\}$: stochastic process. $\Sigma_t = \Sigma_0 + \sum_{i=1}^{t} X_i X_i'$

Note that at $t=0$. by our choice of λ, we have $BALL_0$ contains M, $\Pr(\mu^* \in BALL_0) = 1$. For $t \geq 1$, let us assign failure prob. $\delta_t = \frac{3\delta}{\pi^2 t^2}$ for t-th event.

$$\begin{aligned}
1 - \Pr(\forall t, \mu^* \in BALL_t) &= \Pr(\exists t, \mu^* \notin BALL_t) \\
&\leq \sum_{t=1}^{\infty} \Pr(\mu^* \notin BALL_t) \\
&< \delta \cdot \frac{3}{\pi^2} \cdot \sum_{t=1}^{\infty} \frac{1}{t^2} \\
&= \frac{\delta}{2}
\end{aligned}$$

By $(*)$ and Lemma 6.11. with prob. $> 1-\delta$

$$\begin{aligned}
\sqrt{(\hat{\mu}_t - \mu^*)' \Sigma_t (\hat{\mu}_t - \mu^*)} &\leq \sqrt{\lambda}\|\mu^*\| + \sqrt{2\sigma^2 \log\left((\det\Sigma_t / \det\Sigma_0)/\delta_t\right)} \\
&\leq \sqrt{\lambda}\|\mu^*\| + \sqrt{2\sigma^2\left(d\log\left(1+\frac{tB^2}{d\lambda}\right) + \log(1/\delta_t)\right)} \\
&\leq \sigma^2 + \sqrt{2\sigma^2\left(d\log\left(1+\frac{tB^2W^2}{d\sigma^2}\right) + \log(1/\delta_t)\right)} \\
&\leq \sqrt{\beta_t}
\end{aligned}$$

(Remain confusion about $1/\delta_t$) □