- Rewards, $X_t = \langle A_t, \theta \rangle + \eta_t$
- Action $A_t \in \mathcal{A}_t$
- Action set $\mathcal{A}_t \subset \mathbb{R}^d$.
- η_t noise.

Regret,
$$R_n = \mathbb{E}\left[\sum_{t=1}^{T} \langle A_t^*, \theta \rangle - \sum_{t=1}^{T} \langle A_t, \theta \rangle\right]$$

$$A_t^* = \arg\max_{a \in \mathcal{A}_t} \langle a, \theta \rangle .$$

Least square estimation:

$A_1, \dots, A_t \in \mathbb{R}^d$ and $X_1 \dots X_t \in \mathbb{R}$, $\hat{\theta}_t = G_t^{-1} S_t$.

$$G_t = \lambda I + \sum_{s=1}^{t} A_s A_s^T \qquad S_t = \sum_{s=1}^{T} A_s X_s .$$

when $\lambda = 0$,

- Unbiased $\mathbb{E}\hat{\theta}_t = \theta$
- Variance $\mathbb{E}\left[\langle x, \hat{\theta}_t - \theta \rangle^2\right] = \|x\|^2_{G_t^{-1}}$

$$= x^T G_t^{-1} x .$$

Scanned with CamSca

$$A_t = \underset{a \in \mathcal{A}_t}{\text{argmax}} \; \langle \hat{\theta}_t, a \rangle + \underbrace{\|a\|_{G_t^{-1}} \beta_t}$$

$\beta_t \propto \sqrt{d \log t}$ → some confidence bound.

$$R_n = \mathbb{E}\left[\sum_{t=1}^{n} \langle A_t^*, \theta_* \rangle - \langle A_t, \theta_* \rangle\right]$$

$$r_t = \sum_{t=1}^{n} \langle A_t^*, \theta_* \rangle - \langle A_t, \theta_* \rangle.$$

On theoretical note, Reward with an action chosen is,

$$\langle A_t, \theta \rangle + 2\beta_t \|A_t\|_{G_{t-1}^{-1}} \geq \langle \hat{\theta}_t, A_t \rangle + \beta_t \|A_t\|_{G_{t-1}^{-1}}$$

$$\geq \langle \theta^*, A_t^* \rangle + \beta_t \|A_t\|_{G_{t-1}^{-1}}$$

$$\geq \langle \theta, A_t^* \rangle \rightarrow \text{reward with optimal action.}$$

so now, $r_t \leq 2\beta_t \|A_t\|_{G_{t-1}^{-1}}$.

*). Let's now say, we have some "normalizing" assumptions on 'A_t' and 'θ' as. $\|A_t\| \leq L$, $\|\theta^*\| \leq M$.

then, the regret in each round, r_t should be ≤ 1.

$$\Rightarrow r_t \leq 2\beta_t \|A_t\|_{G_{t-1}^{-1}} \wedge 1.$$

$$\sum_{t=1}^{n} r_t \leq 2\beta_{n} \sum_{t=1}^{n} \min\left(1, \|A_t\|_{G_{t-1}^{-1}}\right).$$

Consider this part $\displaystyle\sum_{t=1}^{n} \min\left(1, \|A_t\|_{G_{t-1}^{-1}}\right)$.

$$\leq \sqrt{n \sum_{t=1}^{n} \min\left(1, \|A_t\|^2_{G_{t-1}^{-1}}\right)}. \quad \text{— (I)}$$

(by Cauchy Swarz inequality).

We know that $\min(1, u) \leq 2\log(1+u)$, using this inequality,

$$\leq 2 \sum_{t=1}^{n} \log\left(1 + \|A_t\|_{G_{t-1}^{-1}}\right).$$

$$\leq 2 \prod_{t=1}^{n} \log \cancel{\det(G_n}\ \det\left(\frac{G_n}{G_0}\right).$$

by AM-GM inequality.

$$\det(G_t) = \prod_i \lambda_i \leq \left(\frac{1}{d} \sum \lambda_i\right)^d$$

-) $\lambda_1 \ldots \lambda_t$ be the eigen values of 'G'.
-) Sum of λ of G = Traces of G.

$$= \left(\frac{Tr(G_n)}{d}\right)^d \simeq \left(\frac{n}{d}\right)^d$$

$$\text{(I)} \Rightarrow \sum_{t=1}^{n} \min\left(1, \|A_t\|_{G_{t-1}^{-1}}\right) \leq 2\sqrt{n \cdot \log \det\left(\frac{G_n}{G_0}\right)}$$

$$\leq 2\sqrt{n \cdot \log\left(\frac{n}{d}\right)^d}$$

$$\leq 2\sqrt{nd \cdot \log\left(\frac{n}{d}\right)}.$$

We know, $\beta_t \simeq \sqrt{d \log t}$,

$$R_n \leq \sqrt{d \log t} \cdot 2\sqrt{nd \log\left(\frac{n}{d}\right)}.$$

$$\leq 2d\sqrt{n}\sqrt{\log\left(t + \frac{n}{d}\right)}.$$

$$R_n = \tilde{O}\left(2d\sqrt{n}\right)$$

More generally $\|\theta_*\| \leq M$, $\forall a \in A$, $\|a\| \leq L$.

Then,

$$\boxed{R_n \leq 2LM\sqrt{n}.}$$