Generalized Linear Bandits

julian.zimmert

July 2016

1 Generalized Linear Model

The arms are denoted as $x_i \in \mathbb{R}^d$, whose features are specific to each arm and known to the agent.
There exists a known link function $\mu : \mathbb{R} \to \mathbb{R}$. μ is monotonously increasing and satisfies for any $x < y$: $c_\mu(y - x) \leq \mu(y) - \mu(x) \leq \kappa_\mu(y - x)$. For some unknown parameter vector $\theta_* \in \mathbb{R}^d$ the payoffs received at time t are

$$R_t = \mu(x_t^T \theta_*) + \eta_t, \tag{1}$$

where η_t is conditionally R-subgaussian for a fixed $R > 0$. This implies that

$$\mathbb{E}[R_t | X_t = x_t] = \mu(x_t^T \theta_*). \tag{2}$$

2 c_μ independent lower bound

We are interested in the case where the link function is Lipschitz, but might have a very small slope. For any finite time horizon T and monotonously increasing function μ, we can construct an infinitesimally larger function μ', that is strictly increasing and indistinguishable from μ w.h.p. in the given horizon. Therefore for a c_μ independent lower bound, it is sufficient to show a lower bound for a μ that is only restricted by $\dot{\mu} \geq 0$.

Theorem 1. *For any finite time horizon $T > 4^5$ there exists a generalized linear bandit problem with finitely many arms, a 1-Lipschitz link function μ, such that the regret of any algorithm will be at least of the order $\Omega(T^{\frac{1}{2}+\frac{1}{10}})$.*

Corollary 1.1. *It is impossible to derive a c_μ independent upper bound for the generalized linear bandit of the order $\tilde{\mathcal{O}}(\sqrt{T})$.*

Proof. Define the GLB Problem $P(T)$ as follows:
We chose the link function

$$\mu(x) := \max\{0, x + \epsilon - 1\} \tag{3}$$

with $\epsilon = T^{-\frac{1}{2}+\frac{1}{10}}$. We have $T^{\frac{1}{5}}$ many arms that are distributed uniformly on the 2d unit ball.
θ_* is one arm chosen at random.

Lemma 2. *For the problem $P(T)$, $T > 4^5$ it holds: $\forall\, x_i \neq \theta_* : \mathbb{E}[R_i | X_i = x_i] = 0$.*

Proof. For any arm $x_i \neq \theta_*$, uniform distribution of the arms implies

$$x_i^T \theta_* \leq \cos(2\pi T^{-\frac{1}{5}})$$

For any $|x| \leq \frac{\pi}{2}$, simple analysis shows that $\cos(x) \leq 1 - (\frac{2x}{\pi})^2$. That implies that for $T \geq 4^5$, we have

$$1 - x_i^T \theta_* \geq (4T^{-\frac{1}{5}})^2 = 16T^{-\frac{2}{5}} > \epsilon$$

Therefore $x_i^T \theta_* + \epsilon - 1 < 0$ and $\mu(x_i^T \theta_*) = 0$. □

Assume an oracle tells the agent that the set of possible values for θ_* is restricted to the set of arms. Then this problem is equivalent to a k-armed bandit problem, where each arm has expected reward 0, but one good arm has expected reward ϵ. Given $\epsilon = \sqrt{\frac{k}{T}}$, the lower bound for k-armed bandits shows in this case that any algorithm will at least observe a regret of the order $\Omega(\sqrt{\frac{k}{T}})$. For any $T \geq 4^5$, we can construct a problem of this kind with $k = T^{\frac{1}{5}}$. As oracle information cannot make an algorithm perform worse, so we obtain a lower bound of $\Omega(T^{\frac{1}{2}+\frac{1}{10}})$ □

3 μ independent upper bound for Explore then Commit with least square estimator

3.1 Problem assumptions

The following assumptions are currently made,

- the arms consist of the surface of the unit ball $\mathcal{B}_1$ (scaling can be absorbed in κ).
- we are playing the arms for a total of T time-steps and this value is known beforehand.
- the norm of $||\theta_*||$ is fixed to 1. (scaling can be absorbed in κ). We can further use this value in our exploration budget. should be generalized to $||\theta_*|| \leq 1$ later.
- $0 \leq \dot{\mu} \leq \kappa$ RHS will be eventually replaced by $\mu(x) \geq \mu(||\theta_*||) - \kappa(1 - x)||\theta_*||$
- $\dot{\mu}(x) = \dot{\mu}(-x)$ shouldn't be required, but also doesn't hurt much and makes the analysis much easier
- the noise is σ-subgaussian

We are further defining the following functions

$$\begin{aligned} L(\theta) &= \mathbb{E}\left[(\mu(x^T\theta) - R)^2\right] \\ &= \mathbb{E}\left[(\mu(x^T\theta) - \mu(x^T\theta_*) - \eta)^2\right] = \mathbb{E}\left[(\mu(x^T\theta) - \mu(x^T\theta_*))^2\right] + \sigma^2 \end{aligned} \tag{4}$$

$$L_n(\theta) = \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\theta) - R_k)^2 = \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\theta) - \mu(x_k^T\theta_*) - \eta_k)^2 \tag{5}$$

Obviously θ_* is the minimum of L.
The least square estimator is

$$\hat{\theta} := \arg\min_{\theta} L_n(\theta) \tag{6}$$

Our algorithm is Explore-then-Commit:
Explore the arms uniformly until a stopping time $\mathcal{T}$ currently this is a fixed value because we know the length of $||\theta_*||$. After this commit to playing $\hat{\theta}$ for the remaining time-steps.

3.2 Results

This is what we aim for, the proof still has gaps though

Theorem 3. *For any time T and any μ under the given constraints, with probability at least $1-\delta$, the regret of the given algorithm will be bounded by*

$$\mathrm{Reg}(T) \leq cT^\wedge\left(\frac{1}{2} + \frac{d-1}{2d+6}\right)\cdot\left(\kappa\sigma^2\left(\log(T)^2 + \log(\delta^{-1})\right)\right)^p \tag{7}$$

The proof will require the following lemmas:

Lemma 4. *With probability at least $1-\delta$, is holds that*

$$L(\hat{\theta}) - L(\theta_*) \leq \frac{8}{n}\sigma^2(c_1\log(2n) + c_2\log\delta^{-1}) + ? \tag{8}$$

Proof. As $\hat{\theta}$ is the minimizer of L_n, it holds that

$$\begin{aligned} \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2 &\leq \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2 + L_n(\theta_*) - L_n(\hat{\theta}) \\ &= \frac{2}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))\eta_k \\ &= \frac{2}{\sqrt{n}}\sqrt{\frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2}\sum_{k=1}^{n}\frac{(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))}{\sqrt{\sum_{l=1}^{n}(\mu(x_l^T\hat{\theta}) - \mu(x_l^T\theta_*))^2}}\eta_k \\ w.h.p. &\leq \frac{2}{\sqrt{n}}\sqrt{\frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2}\sigma(c_1\sqrt{\log(2n)} + c_2\sqrt{\log\delta^{-1}}) \\ \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2 &\leq \frac{8}{n}\sigma^2(c_1\log(2n) + c_2\log\delta^{-1}) \end{aligned}$$

TODO: bound $\mathbb{E}\left[(\mu(x^T\theta) - \mu(x^T\theta_*))^2\right] - \frac{1}{n}\sum_{k=1}^{n}(\mu(x_k^T\hat{\theta}) - \mu(x_k^T\theta_*))^2$ for bound on $L(\hat{\theta}) - L(\theta_*)$. Write down the constants c_1, c_2 explicitly.

□

Lemma 5. *Let $\tilde{\theta}$ be the vector $\hat{\theta}$ rescaled such that it matches the length of θ_*. Then $L(\hat{\theta}) - L(\theta_*) \geq \frac{1}{2}(L(\tilde{\theta}) - L(\theta_*))$*

Proof. Let θ_1 and θ_2 be the two orthogonal vectors such that $\theta_* = \theta_1 - \theta_2$ and $\tilde{\theta} = \theta_1 + \theta_2$.
Define the two sets

$$\mathcal{B}_1 = \left\{x | \operatorname{sign}(x^T\theta_1) = \operatorname{sign}(x^T\theta_2)\right\}$$
$$\mathcal{B}_2 = \left\{x | \operatorname{sign}(x^T\theta_1) = -\operatorname{sign}(x^T\theta_2)\right\}.$$

Due to Symmetry it holds that

$$\mathbb{E}\left[(\mu(x^T\tilde{\theta}) - \mu(x^T\theta_*))^2 | x \in \mathcal{B}_1\right] = \mathbb{E}\left[(\mu(x^T\tilde{\theta}) - \mu(x^T\theta_*))^2 | x \in \mathcal{B}_2\right]$$

If $||\hat{\theta}||_2 > ||\tilde{\theta}||_2$, then $(\mu(x^T\hat{\theta}) - \mu(x^T\theta_*))^2 \geq (\mu(x^T\tilde{\theta}) - \mu(x^T\theta_*))^2$ for all $x \in \mathcal{B}_1$.
If $||\hat{\theta}||_2 < ||\tilde{\theta}||_2$, then $(\mu(x^T\hat{\theta}) - \mu(x^T\theta_*))^2 \geq (\mu(x^T\tilde{\theta}) - \mu(x^T\theta_*))^2$ for all $x \in \mathcal{B}_2$.
Finally we get

$$\begin{aligned}
&L(\hat{\theta}) - L(\theta_*) = \mathbb{E}\left[(\mu(x^T\hat{\theta}) - \mu(x^T\theta_*))^2\right] \\
=&\frac{1}{2}\left(\mathbb{E}\left[(\mu(x^T\hat{\theta}) - \mu(x^T\theta_*))^2 | x \in \mathcal{B}_1\right] + \mathbb{E}\left[(\mu(x^T\hat{\theta}) - \mu(x^T\theta_*))^2\right] | x \in \mathcal{B}_2\right) \\
\geq&\frac{1}{2}\mathbb{E}\left[(\mu(x^T\tilde{\theta}) - \mu(x^T\theta_*))^2\right] = \frac{1}{2}(L(\tilde{\theta}) - L(\theta_*))
\end{aligned}$$

□

3.2.1 $d = 2$

This is the most important gap. It holds for $\mu = (x - 1 + \epsilon)_+$, but I haven't succeeded in a general proof.

Proposition 6. *Let $d = 2$. Then for any valid $\mu \in M$, it holds that*

$$L(\tilde{\theta}) - L(\theta_*) \geq \kappa^{-\frac{1}{2}}(\mu(x_*^T\theta_*) - \mu(\hat{x}^T\theta_*))(\mu(1) - \mu(-1))^{\frac{3}{2}}$$

The proof follows from the following Lemma

Lemma 7. *For any $\mu \in M_1$ and $z \in [0,1]$, it holds that*

$$\int_{-1}^{1} \left(\mu\left(zx + \sqrt{1-z^2}\sqrt{1-x^2}\right) - \mu\left(zx - \sqrt{1-z^2}\sqrt{1-x^2}\right)\right)^2 \frac{1}{\sqrt{1-x^2}}\,dx$$
$$\geq c\left(\mu(1) - \mu(z^2)\right)\left(\mu(1) - \mu(-1)\right)^{\frac{3}{2}}$$
$$\Leftrightarrow$$
$$\int_{-1}^{1} (1-z^2)\left(\int_{-1}^{1} \dot{\mu}\left(zx + \sqrt{1-z^2}\sqrt{1-x^2}y\right)\,dy\right)^2 \sqrt{1-x^2}\,dx$$
$$\geq c\left(\int_{z^2}^{1} \dot{\mu}(x)\,dx\right)\left(\int_{-1}^{1} \dot{\mu}(x)\,dx\right)^{\frac{3}{2}}$$

Proof of Theorem 6.

$$\begin{aligned}
L(\tilde{\theta}) - L(\theta_*) =& \mathbb{E}\left[(\mu(x^T\theta_1 - x^T\theta_2) - \mu(x^T\theta_1 + x^T\theta_2))^2\right] \\
=& \int_0^{2\pi} \left(\mu(\cos(x)|\theta_1| - \sin(x)|\theta_2|) - \mu(\cos(x)|\theta_1| + \sin(x)|\theta_2|)\right)^2\,dx \\
=& 2\int_0^{\pi} \left(\int_{-\sin(x)|\theta_2|}^{\sin(x)|\theta_2|} \dot{\mu}(\cos(x)|\theta_1| + y)\,dy\right)^2\,dx \\
=& 2|\theta_2|^2 \int_0^{\pi} \left(\int_{-1}^{1} \dot{\mu}(\cos(x)|\theta_1| + \sin(x)|\theta_2|y)\,dy\right)^2 \sin(x)^2\,dx \\
=& 2|\theta_2|^2 \int_{-1}^{1} \left(\int_{-1}^{1} \dot{\mu}(x|\theta_1| + \sqrt{1-x^2}|\theta_2|y)\,dy\right)^2 \sqrt{1-x^2}\,dx \\
& \ldots
\end{aligned}$$

□

3.3 $d \geq 3$

We express the Expectation in terms of integrals

$$
\begin{aligned}
& L(\tilde{\theta}) - L(\theta_*) \\
=& \mathbb{E}\left[(\mu(x^T\theta_1 - x^T\theta_2) - \mu(x^T\theta_1 + x^T\theta_2))^2\right] \\
=& \frac{1}{|S_d|}\int_{S_d} (\mu(x^T\theta_1 - x^T\theta_2) - \mu(x^T\theta_1 + x^T\theta_2))^2 \, d\mathcal{L}_{d-1}(x) \\
=& \frac{2|S_{d-2}|}{|S_d|}\int_0^{\frac{\pi}{2}} \cos(\alpha)^{d-2} \\
& \int_0^{2\pi} (\mu(\sin(\alpha)(\cos(x)|\theta_1| - \sin(x)|\theta_2|)) - \mu(\sin(\alpha)(\cos(x)|\theta_1| + \sin(x)|\theta_2|)))^2 \, dx \, d\alpha \\
\geq& \frac{4|S_{d-2}|}{|S_d|}\int_0^{\frac{\pi}{2}} \cos(\alpha)^{d-2} J(\mu, \sin(\alpha)||\theta_*||) \, d\alpha \\
=& \frac{4|S_{d-2}||\theta_2|^2}{|S_d|}\int_0^1 \frac{\alpha^{d-2}}{\sqrt{1-\alpha^2}} J(\mu, \sqrt{1-\alpha^2})||\theta_*||) \, d\alpha
\end{aligned}
$$