High Dimensional Linear Contextual Bandit and Thompson Sampling using Langevin dynamics

Sunrit Chakraborty, Saptarshi Roy

University of Micigan, Ann Arbor

April 20, 2021

Introduction

- Linear contextual bandit is a particular version of MAB.
- In this problem we have some "contexts" x_i associated with each arm.
- The learner uses these feature vectors along with the feature vectors and rewards of the arms played by her in the past to make the choice of the arm to play in the current round.
- The goal is to minimize regret over time.

Existing works

Adversarial setting:

- [Agrawal and Goyal, 2013] has given a near optimal algorithm (based on TS) to solve this problem. They have shown that their algorithm achieves on an average of $O(\frac{d^2}{\epsilon}\sqrt{T^{1+\epsilon}})$ for some positive $\epsilon > 0$ which is close to the theoretical lower bound $\Omega(d\sqrt{T})$ for this problem.
- Following the work of [Agrawal and Goyal, 2013] on linear contextual bandits a further progress was made by [Abeille and Lazaric, 2017]. They showed that it is not necessary for TS to sample from the actual posterior distribution and still get a regret of order $O(d^{3/2}\sqrt{T})$.

Existing works

Stochastic setting:

- [Chu et al., 2011] has provided a near optimal algorithm in stochastic setting with a regret bound of $O(\sqrt{Td\log^3(KT\log(T/\delta))})$.
- They have also proved a $\Omega(\sqrt{dT})$ lower bound on the regret under this setting. This bound is improvement over the previously existing bound $\Omega(T^{3/4}K^{1/4})$ by [Abe et al., 2003].

Linear contextual bandit

We will work with Stochastic linear contextual bandit:

- Consider a collection K arms.
- We denote those K arms via $d-$dimensional vectors $\mathcal{X} = \{x_1, \ldots, x_K\}$.
- At each time point $t \in \{0, 1, \ldots, T\}$ the user selects an arm A_t and receives a reward $y_t \in \mathbb{R}$.
- We model the reward as

$$y_t = x_{A_t}^\top \beta^* + \eta_t.$$

Linear contextual bandit

Setup and Assumptions:

- We assume that $\|x_k\|_2 \leq 1$ and $\|\beta^*\|_2 \leq 1$.
- $\eta_t \sim N(0, \sigma^2)$, and they are all i.i.d across t.
- β^* is unknown weight vector.

Optimal arm: $x^* = \arg\max_{x \in \mathcal{X}} x^\top \beta^*$.
Goal: Our goal is to minimize the cumulative regret over T rounds,

$$R(T) = \sum_{t=1}^{T} \langle \beta^*, x^* \rangle - \langle \beta^*, x_{A_t} \rangle .$$

How do we tackle this problems?

Typically we have two basic approaches.

1. **Frequentist Approach**: Lin-UCB, OFUL. Both of these algorithms are based on constructing confidence ellipsoid for β^* and taking decision based on that.
2. **Bayesian Approach**: Thompson sampling. Typically we construct a prior on the space of β and compute the posterior. Then we generate one sample from the posterior and use it as a proxy for β. Then we try to choose an arm based on the procured sample and we continue.

Lin UCB

Algorithm 1 LinUCB with disjoint linear models.

Inputs: $\alpha \in \mathbb{R}_+$
for $t = 1, 2, 3, \ldots, T$ **do**
  Observe features of all arms $a \in \mathcal{A}_t$: $\mathbf{x}_{t,a} \in \mathbb{R}^d$
  for all $a \in \mathcal{A}_t$ **do**
    if a is new **then**
      $\mathbf{A}_a \leftarrow \mathbf{I}_d$ (d-dimensional identity matrix)
      $\mathbf{b}_a \leftarrow \mathbf{0}_{d \times 1}$ (d-dimensional zero vector)
    end if
    $\hat{\boldsymbol{\theta}}_a \leftarrow \mathbf{A}_a^{-1} \mathbf{b}_a$
    $p_{t,a} \leftarrow \hat{\boldsymbol{\theta}}_a^\top \mathbf{x}_{t,a} + \alpha \sqrt{\mathbf{x}_{t,a}^\top \mathbf{A}_a^{-1} \mathbf{x}_{t,a}}$
  end for
  Choose arm $a_t = \arg\max_{a \in \mathcal{A}_t} p_{t,a}$ with ties broken arbitrarily, and observe a real-valued payoff r_t
  $\mathbf{A}_{a_t} \leftarrow \mathbf{A}_{a_t} + \mathbf{x}_{t,a_t} \mathbf{x}_{t,a_t}^\top$
  $\mathbf{b}_{a_t} \leftarrow \mathbf{b}_{a_t} + r_t \mathbf{x}_{t,a_t}$
end for

Source: https://arxiv.org/pdf/1003.0146.pdf

Thompson sampling: Basic idea

Let $\pi(\cdot)$ be a prior on β^*. For simplicity let

$$\beta^* \sim \mathcal{N}(0, \Lambda^{-1}).$$

Then we do the followings:

1. **Initialization**: Initialize $\hat{\beta}_1 = 0$ and $A = \Lambda_1 = \mathbb{I}_d$.
2. **Draw sample form posterior**: At time t draw a sample $\tilde{\beta}_t$ from $\mathcal{N}(\hat{\beta}_t, \Lambda_t^{-1})$.
3. **Play arm**: $A_t = \arg\max_{k \in [K]} x_k^\top \tilde{\beta}$.
4. **Update the parameters(i.e. the posterior parameters)**: $\Lambda_{t+1} = \Lambda_t + x_{A_t} x_{A_t}^\top$, $\hat{\beta}_{t+1} = \Lambda_{t+1}^{-1}[x_{A_1}, \ldots, x_{A_t}][y_1, \ldots, y_t]^\top$.
5. Return to step 2.

High dimensional linear contextual bandit

Notations:

1. β^* : Unknown $d-$ dimensional parameter.
2. T : Number of total iterations and $T \ll d$.
3. s : Sparsity of β^* and $s \ll d$.
4. $X_{t,i}$: Context covariate $X_{A_t} \in \mathbb{R}^d$ associated with each feasible action in $\mathcal{X}$ at time t.
5. y_t : Linear reward observed after taking an action $A_t \in [K]$.

Model: $y_t = x_{A_t}^\top \beta^* + \eta_t$.
Goal: Our goal is to minimize the cumulative regret over T rounds,

$$R(T) = \sum_{t=1}^{T} \langle \beta^*, x^* \rangle - \langle \beta^*, x_{A_t} \rangle .$$

Relevance Vector Machine

Relevance Vector Machine is bayesian frame work that is particularly useful in sparse regression and classification tasks([Tipping, 2001]). Consider the linear model as an example

$$y = X\beta^* + \eta.$$

- Prior on β^*: $\beta^* \sim \mathcal{N}(0, A^{-1})$.
- Structure of A: $A = \mathrm{diag}(\alpha_1, \ldots, \alpha_\mathrm{d})$. Thus $\beta_j^* \sim \mathcal{N}(0, \alpha_j^{-1})$.
- Hyperprior on A : $(\alpha_1, \ldots, \alpha_d) \sim \prod_{j=1}^{d} \mathrm{Gamma}(0, 0)$.
- Posterior of β^*: $\beta^*|y, \boldsymbol{\alpha} \sim \mathcal{N}(\hat{\beta}, \Sigma^{-1})$ where,

$$\Sigma = A + \sigma^{-2} X^\top X; \quad \hat{\beta} = \sigma^{-2} \Sigma^{-1} X^\top y.$$

Relevance Vector Machine

Since α is unknown and computing the full posterior is computationally intensive, [Tipping, 2001] proposes choosing α to maximize $p(y|\boldsymbol{\alpha})$ and describes an iterative approach to solving this optimization problem, with computational speedups detailed in [Tipping et al., 2003].

Algorithm using RVM

Now we will present the algorithm proposed in [Gilton and Willett, 2017].

1. Initialize $\hat{\beta}_1 = 0$ and $A = \Sigma_1 = I_d$.
2. Let $\delta \in (0, 1)$ and set $\nu = \sqrt{9d\sigma^2 \log(T/\delta)}$.
3. for $t \in [T]$ do the followings,
4. sample $\tilde{\beta}_t \sim \mathcal{N}(\hat{\beta}_t, \nu^2 \Sigma_t^{-1})$
5. Play arm $x_{A_t} = \arg\max_{x \in \mathcal{X}} x^\top \tilde{\beta}_t$
6. Update A using the procedure in [Tipping et al., 2003].
7. Set $\Sigma_{t+1} = \Sigma_t + x_{A_t} x_{A_t}^\top = A + \sum_{s=1}^{t} x_{A_s} x_{A_s}^\top$.
8. $\hat{\beta}_{t+1} = \Sigma_{t+1}^{-1} (\sum_{s=1}^{t} y_s x_{A_s})$.

Algorithm using RVM

Suppose we have perfect knowledge of $S := \mathrm{Supp}(\beta^*)$. then if $\alpha_i = \infty$ for $i \in S^c$ and $\alpha_i = \lambda$ for $i \in S$, and for now let us forget about updating A. Then this RVM-LTS is basically same as LTS on the true support of β^*.

Regret Bound: fixed A (i.e. we do not update A)

Define $S_\alpha := \{i : 1/\alpha_i = 0\}$ and $s_\alpha = |S_\alpha|$.

Theorem

Assume $\mathrm{Supp}(\theta^*) \subseteq \mathrm{S}_\alpha$. *For* $\delta > 0$, *with probability at least* $1 - \delta$ *we have,*

$$R(T) \leq O\left(\min\{\sqrt{s_\alpha}, \sqrt{\log K}\}\left[\sigma\sqrt{s_\alpha \log\left(\frac{T^2\alpha_{AM} + T^3}{\delta.\alpha_{GM}}\right)} + \|\alpha_S\|_2\right]\right),$$

where $\alpha_{AM} = \frac{1}{s_\alpha}\sum_{i\in S_\alpha}\alpha_i$ *and* $\alpha_{GM} = (\prod_{i\in S_\alpha}\alpha_i)^{1/s_\alpha}$.

Proof sketch

- The proof mainly follows the steps of [Chu et al., 2011].
- Key element of the proof is to find concentration inequality for $x_{A_t}^\top \hat{\beta}_t$.
- This requires basically bounding $\left| x_{A_t}^\top (\hat{\beta}_t - \beta^*) \right|$.
- As x_k's are bounded it is enough to get a bound for $\hat{\beta} - \beta^* = \Sigma_{t-1}^{-1}(\zeta_{t-1} - A\beta^*)$, where $\zeta_t = \boldsymbol{X}_t^\top \boldsymbol{y}_t$ and $\Sigma_t = A + \boldsymbol{X}_t^\top \boldsymbol{X}_t$.
- The first thing to show is,

$$\|\zeta\|_{\Sigma_t^{-1}} \leq R\sqrt{d \log \frac{\alpha_{AM} + t}{\delta \alpha_{GM}}}.$$

Proof sketch

Using this one can further show that,

$$\|\zeta_{t-1} - A\beta^*\|_{\Sigma_{t-1}^{-1}} \leq R\sqrt{d \log \frac{\alpha_{AM} + t}{\delta \alpha_{GM}}} + \|\boldsymbol{\alpha s}\|_2 .$$

The remainder of the proof is detailed in [Chu et al., 2011].

Experimental results

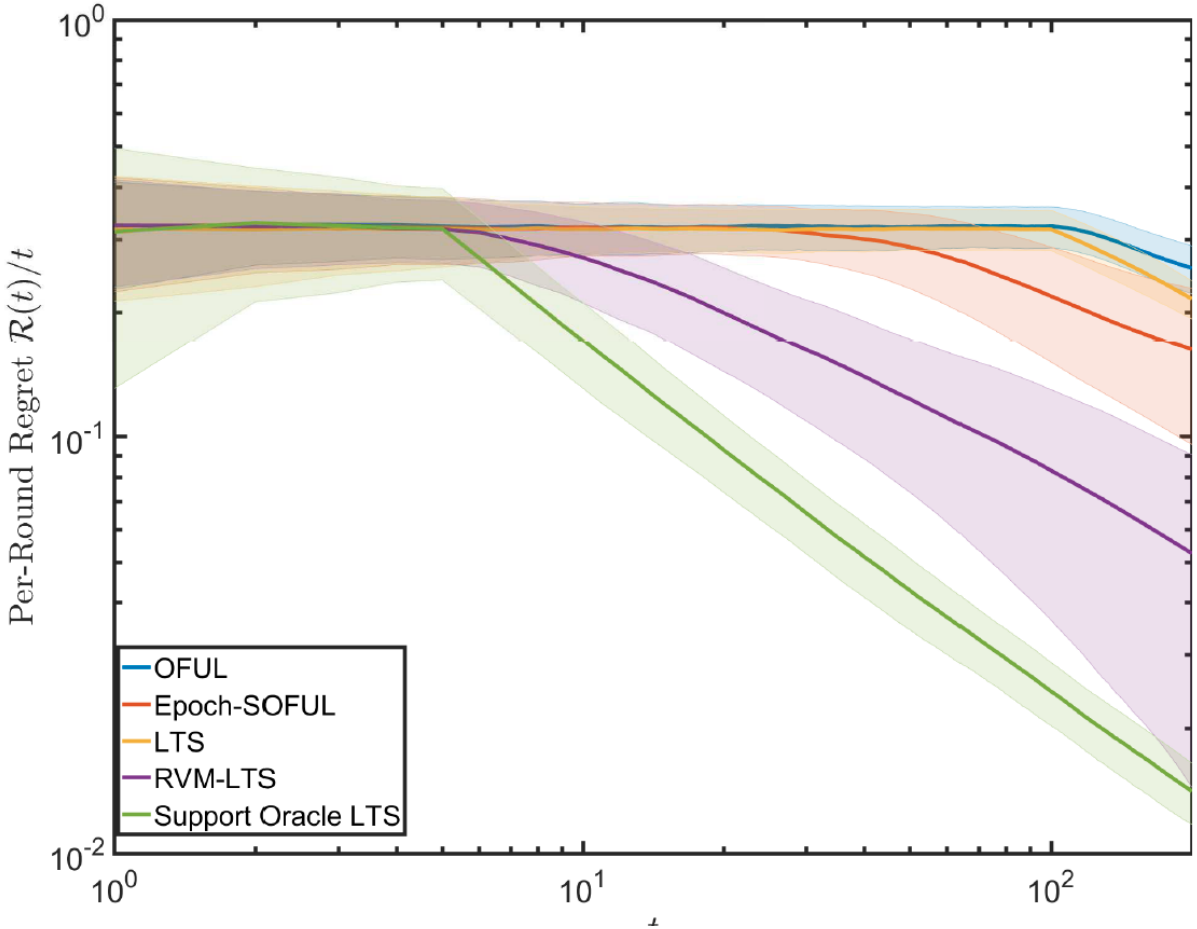


Figure: Per-round regret v/s time. $s = 5, d = 100, K = 1000$

Experimental results

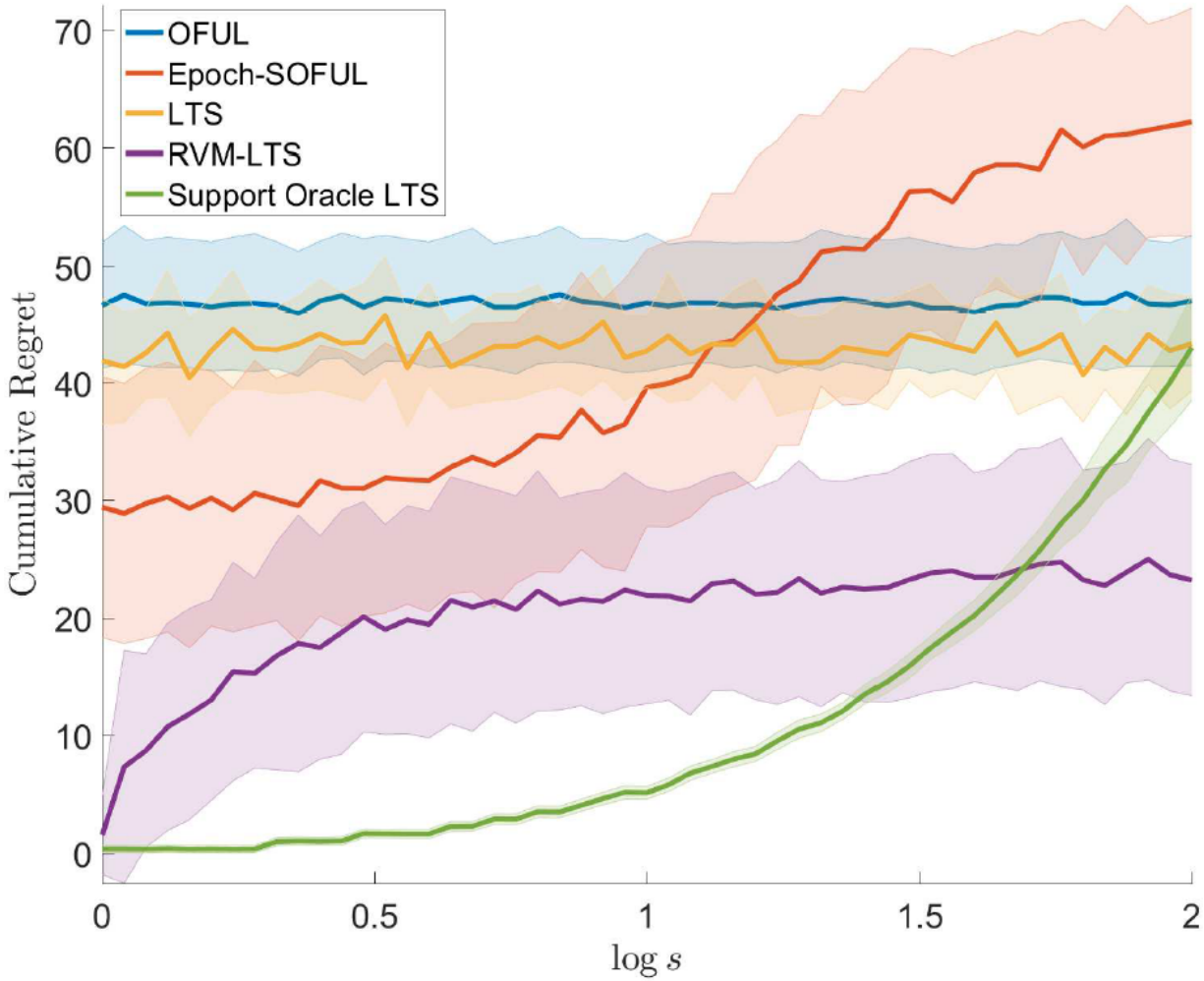


Figure: Total regret v/s sparsity. $d = 100, K = 1000$

Thompson sampling: Disadvantage

- The computation of posterior is easy if we consider a gaussian prior on β^*.
- In most of the cases if we work with complicated prior or hierarchical models then it is almost impossible to get a closed form of the posterior distribution.
- Even if we get a form of the posterior it is often the case we struggle to get a sample from the posterior.

Thompson sampling: Solution

- One well known methodology is Markov Chain Monte Carlo. For example: Gibb's Sampler, Metropolis Hastings, Hamiltonian Monte Carlo.
- Experiments shows that these methods are quite slow. It mainly occurs due to slow convergence of the Markov chain to its stationary distribution.
- Thus to curb time we need to generate some Markov chain which quickly converges to its stationary distribution.

Langevin dynamics: Langevin SDE

Langevin Process:

$$d\theta_t = -\nabla U(\theta_t)dt + \sqrt{\frac{2}{\gamma}}dB_t. \tag{1}$$

$(B_t, t \geq 0)$ is the standard Brownian motion, and the potential function $U : \mathbb{R}^d \to \mathbb{R}$ is assumed to satisfy the regularity condition:

- ∇U is locally lipschitz.
- $\langle \nabla U(\theta), \theta \rangle \geq c_1 \|\theta\|_2 - c_2$ for any $\theta \in \mathbb{R}^d$,

where c_1, c_2 are positive constants.

Langevin dynamics: Langevin SDE

Theorem

Under the stated regularity condition, the solution to the Langevin SDE (1) exists and is unique. Further the density of θ_t converges in L_2 to the stationary distribution with density proportional to $e^{-\gamma U}$.

Langevin dynamics: Connection with Bayesian Inference

- Consider a parametric family family of distribution $\{P_\theta \,|\theta \in \Theta\}$.
- $X_1^n := (X_1, \ldots, X_n)$ be i.i.d sequence from P_{θ^*}.
- log-likelihood: $F_n(\theta) := \frac{1}{n}\sum_{i=1}^n \log p_\theta(X_i)$.
- Posterior: $\Pi(\theta|X_1^n) \propto e^{nF_n(\theta)}\pi(\theta)$, where $\pi(\cdot)$ is prior on θ.

Langevin dynamics: Connection with Bayesian Inference

Now consider the following SDE:

$$d\theta_t = \underbrace{\frac{1}{2}\nabla F_n(\theta_t)dt + \frac{1}{2n}\nabla \log \pi(\theta_t)dt}_{:=-\nabla U_n(\theta_t)dt} + \frac{1}{\sqrt{n}}dB_t.$$

Thus if we show the aforementioned regularity conditions for $\nabla U_n(\theta)$ then density of θ_t converges to $\Pi(\cdot|X_1^n)$ in L_2.

Langevin dynamics: Posterior contraction

Let $F(\theta) = \mathbb{E}(\log p_\theta(X))$.

(A1) $\|\nabla F(\theta_1) - \nabla F(\theta_2)\|_2 \leq \|\theta_1 - \theta_2\|_2$

(A2) $\|\nabla \log \pi(\theta_1) - \nabla \log \pi(\theta_2)\|_2 \leq L_2 \|\theta_1 - \theta_2\|_2$.

(A3) $\sup_{\theta \in \mathbb{R}^d} \langle \nabla \log \pi(\theta), \theta - \theta^* \rangle \leq B$.

(A4) $\langle \nabla F(\theta), \theta^* - \theta \rangle \geq \mu \|\theta - \theta^*\|_2^2$.

(A5) $\sup_{\theta \in B(\theta^*, r)} \|\nabla F_n(\theta) - \nabla F(\theta)\| \leq \varepsilon_1(n, \delta) + \varepsilon_2(n, \delta)$ with probability at least $1 - \delta$.

Langevin dynamics: Posterior contraction

Theorem (Informal version, [Mou et al., 2019])

If (A1)-(A5) is satisfied then for large enough n such that $\varepsilon_1(n,\delta) \leq \mu/6$, we have

$$\Pi\left(\|\theta - \theta^*\|_2 \geq C\sqrt{\frac{d + \log(1/\delta) + B}{n\mu} + \frac{\varepsilon_2(n,\delta)}{\mu}} \,\middle|\, X_1^n\right) \leq \delta,$$

with probability at least $1 - \delta$.

This establishes $(d/n)^{1/2}$ rate of posterior contraction.

Langevin dynamics: Stochastic optimization

- The basic structure of the Langevin based algorithms are hinged on *stochastic optimization* of some cost function.
- An ordinary stochastic gradient update step is:

$$\theta_{t+1} = \theta_t + \frac{h_t}{2}\left(\nabla \log \pi(\theta) + \frac{n}{k}\sum_{i=1}^{k} \nabla \log p(x_{ti}|\theta_t)\right)$$

  $k =$ number of sub-sample, $n =$ number of samples, x_{ti} are sub-samples.
- But this update step collapses to MAP estimator. But we need a sample from the posterior.

Langevin dynamics: Stochastic optimization

- But the Langevin update does the following:

$$\theta_{t+1} = \theta_t + \frac{h_t}{2}\left(\nabla \log \pi(\theta) + \frac{n}{k}\sum_{i=1}^{k} \nabla \log p(x_{ti}|\theta_t)\right) + \sqrt{2h_t}w_t$$

$w_t \sim \mathcal{N}_d(0,1)$. Typically we have,

$$\sum_t h_t = \infty, \quad \sum_t h_t^2 < \infty.$$

It turns out that adding this gaussian noise eventually generates a sample from the posterior.

K-armed stochastic MAB

Setup:

- $\mathcal{A} := \{1, \ldots, K\}$.
- At time t learner chooses an arm A_t and receives reward X_{A_t} from a distribution p_{A_t}.
- We assume that the reward distribution of arm a can be characterized by $\theta_a \in \mathbb{R}^{d_a}$, i.e, reward distribution is $p_a(X) = p_a(X; \theta_a^*)$.
- $\mathbb{E}_{x \sim p_a(x|\theta_a)}[X] = \alpha_a^T \theta_a$.

Goal: Minimize

$$R(T) = \mathbb{E}\left[\sum_{t=1}^{T} \bar{r}_{a^*} - \bar{r}_{A_t}\right].$$

$\bar{r}_a$ = mean reward of arm a, $\quad a^* = \arg\max_{a \in \mathcal{A}} \bar{r}_a$.

K-armed stochastic MAB: Langevin SDE

We consider the following SDE:

$$d\theta_t = \frac{1}{2}\nabla_\theta F_{n,a}(\theta_t)dt + \frac{1}{2n}\nabla_\theta \log \pi_a(\theta_t)dt + \frac{1}{\sqrt{n\gamma_a}}dB_t.$$

Thus we expect,

$$\lim_{t\to\infty} P_t(\theta|X_1^n) \propto \exp(-\gamma_a(nF_{n,a}(\theta) + \log \pi_a(\theta)))$$

Sampling from scaled posterior $\mu_a^{(n)}[\gamma_a]$ is needed for technical reasons.

K-armed stochastic MAB: Assumptions

Assumption 1- Uniform

$$
\begin{aligned}
&-\log p_a(x|\theta_a') - \nabla_\theta \log p_a(x|\theta_a')^T(\theta_a - \theta_a') + \frac{m_a}{2}\left\|\theta_a - \theta_a'\right\|^2 \\
&\leq -\log p_a(x|\theta_a) \\
&\leq -\log p_a(x|\theta_a') - \nabla_\theta \log p_a(x|\theta_a')^T(\theta_a - \theta_a') + \frac{L_a}{2}\left\|\theta_a - \theta_a'\right\|^2 .
\end{aligned}
$$

Assumption 2: Strong log-concavity and lipschitz assumption on $p_a(x|\theta)$ in x.

K-armed stochastic MAB: Algorithm

Algorithm 2 (Stochastic Gradient) Langevin Algorithm for Arm a

Input : Data $\{x_{a,1}, \cdots, x_{a,n}\}$;
MCMC sample $\theta_{a,Nh^{(n-1)}}$ from last round

Set $\theta_0 = \theta_{a,t-1}$ for $a \in \mathcal{A}$

for $i = 0, 1, \cdots N$ **do**

Uniformly subsample $\mathcal{S} \subseteq \{x_{a,1}, \cdots, x_{a,n}\}$.

Compute $\nabla \widehat{U}(\theta_{ih^{(n)}}) = -\frac{n}{|\mathcal{S}|} \sum_{x_k \in \mathcal{S}} \nabla \log p_a(x_k|\theta_{ih^{(n)}}) - \nabla \log \pi_a(\theta_{ih^{(n)}})$.

Sample $\theta_{(i+1)h^{(n)}} \sim \mathcal{N}\left(\theta_{ih^{(n)}} - h^{(n)} \nabla \widehat{U}(\theta_{ih^{(n)}}), 2h^{(n)}\mathrm{I}\right)$.

Output : $\theta_{a,Nh^{(n)}} = \theta_{Nh^{(n)}}$ and $\theta_{a,t} \sim \mathcal{N}\left(\theta_{Nh^{(n)}}, \frac{1}{nL_a\gamma_a} I\right)$

Main regret result

Theorem (Informal, [Mazumdar et al., 2020])

When the likelihood and the true reward distributions satisfy aforementioned regularity conditions: The algorithm after T rounds satisfies the following,

$$\mathbb{E}[R(T)] \leq O\left(\frac{\log T}{\Delta_{min}}\right),$$

$\Delta_{min} =$ *minimum sub-optimality gap across the arms.*

Proof sketch

- Between the n-th and the $(n+1)$-th pull to arm a, samples $\theta_{a,t}$ approximately follows the posterior $\mu_a^{(n)}$:

$$W_p(\hat{\mu}_a^{(n)}, \mu_a^{(n)}) \leq O\left(\sqrt{\frac{d_a + C_1 p + C_2}{n}}\right).$$

  $\hat{\mu}_a^n$ is the probability measure associated with any of the samples(s) $\theta_{a,Nh_a^{(n)}}$.

- Then we show that $\theta_{a,t}$ concentrates to θ_a^*,i.e., $\|\theta_{a,t} - \theta_a^*\|_2 \leq \sqrt{\frac{d_a}{n}}$, with high probability.

Proof sketch

- Next we try to bound $\mathbb{E}(T_a(T))$, where $T_a(T)$= number of times the sub-optimal arm a is pulled up.
- $E_a(t) = \{r_{a,t}(T_a(t)) \geq \bar{r}_1 - \epsilon\}$. (1 is optimal arm).

$$\mathbb{E}(T_a(T)) = \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} 1(A_t = a, E_a^c(t))\right]}_{\text{A}} + \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} 1(A_t = a, E_a(t))\right]}_{\text{B}}$$

$p_{a,s} = \mathbb{P}(r_{a,t}(s) > \bar{r}_1 - \epsilon | \mathcal{F}_{t-1})$ for some $\epsilon > 0$.

Proof sketch

Using standard techniques of [Agrawal and Goyal, 2012] one can show,

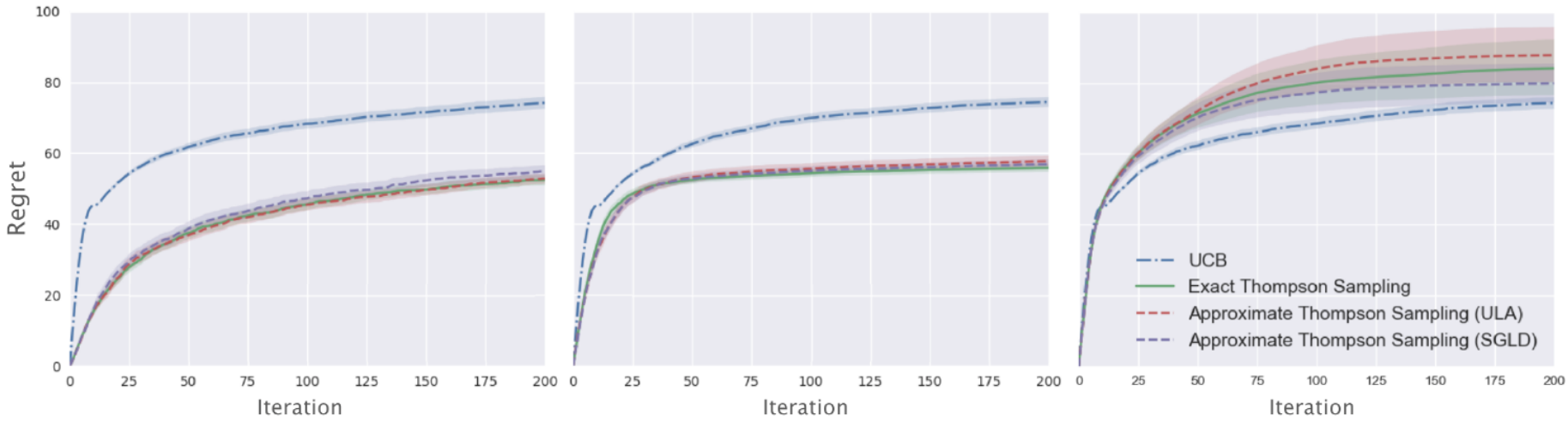
$$\text{(A)} \leq \mathbb{E}\left[\sum_{s=l}^{T-1} \frac{1}{p_{1,s}} - 1\right],$$

$$\text{(B)} \leq 1 + \mathbb{E}\left[\sum_{s=1}^{T} 1(p_{a,s} > \frac{1}{T})\right],$$

then one can further get an upper bound on the right hand terms.

- Then using standard regret decomposition for stochastic settings, one can prove the final bound to be $O(\log T/\Delta)$.

Numerical results



a. Good Priors b. Uniform Priors c. Adversarial Priors

Figure 1: Performance of exact and approximate Thompson sampling vs UCB on Gaussian bandits with (a) "good priors" (priors reflecting the correct ordering of the arms' means), (b) the same priors on all the arms' means, and (c) "bad priors" (priors reflecting the exact opposite ordering of the arms' means). The shaded regions represent the 95% confidence interval around the mean regret across 100 runs of the algorithm.


Abe, N., Biermann, A. W., and Long, P. M. (2003).
Reinforcement learning with immediate rewards and linear hypotheses.
Algorithmica, 37(4):263–293.


Abeille, M. and Lazaric, A. (2017).
Linear Thompson Sampling Revisited.
In Singh, A. and Zhu, J., editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pages 176–184, Fort Lauderdale, FL, USA. PMLR.


Agrawal, S. and Goyal, N. (2012).
Analysis of thompson sampling for the multi-armed bandit problem.
In *Conference on learning theory*, pages 39–1. JMLR Workshop and Conference Proceedings.


Agrawal, S. and Goyal, N. (2013).
Thompson sampling for contextual bandits with linear payoffs.

In *International Conference on Machine Learning*, pages 127–135. PMLR.


Chu, W., Li, L., Reyzin, L., and Schapire, R. (2011).
Contextual bandits with linear payoff functions.
In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pages 208–214. JMLR Workshop and Conference Proceedings.


Gilton, D. and Willett, R. (2017).
Sparse linear contextual bandits via relevance vector machines.
In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 518–522. IEEE.


Mazumdar, E., Pacchiano, A., Ma, Y.-a., Bartlett, P. L., and Jordan, M. I. (2020).
On thompson sampling with langevin algorithms.
arXiv preprint arXiv:2002.10002.


Mou, W., Ho, N., Wainwright, M. J., Bartlett, P., and Jordan, M. I. (2019).
A diffusion process perspective on posterior contraction rates for parameters.
arXiv preprint arXiv:1909.00966.


Tipping, M. E. (2001).
Sparse bayesian learning and the relevance vector machine.
Journal of machine learning research, 1(Jun):211–244.


Tipping, M. E., Faul, A. C., et al. (2003).
Fast marginal likelihood maximisation for sparse bayesian models.
In *AISTATS*.