Generalized Latent Bandits

Adobe Advisor(s): Branislav Kveton, Anup Rao *Intern: Subhojyoti Mukherjee*

Abstract

To be written.

1 Introduction

We have proposed and deliberated on a generalized latent bandit model.

2 Related Works

In Maillard and Mannor (2014) the authors propose the Latent Bandit model where there are two sets: 1) set of arms denoted by $\mathcal{A}$ and 2) set of types denoted by $\mathcal{B}$ which contains the latent information regarding the arms. The latent information for the arms are modeled such that the set $\mathcal{B}$ is assumed to be partitioned into $|C|$ clusters, indexed by $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_C \in \mathcal{C}$ such that the distribution $v_{a,b}, a \in \mathcal{A}, b \in \mathcal{B}_c$ across each cluster is same. Note, that the identity of the cluster is unknown to the learner. At every timestep t, nature selects a type $b_t \in \mathcal{B}_c$ and then the learner selects an arm $a_t \in \mathcal{A}$ and observes a reward $X_{a,t}$ from the distribution $v_{a,b}$.

Another way to look at this problem is to imagine a matrix of dimension $|A| \times |B|$ where again the rows in $\mathcal{B}$ can be partitioned into $|C|$ clusters, such that the distribution across each of this clusters are same. Now, at every timestep t one of this row is revealed to the learner and it chooses one column such that the $v_{a,b}$ is one of the $\{v_{a,c}\}_{c \in \mathcal{C}}$ and the reward for that arm and the user is revealed to the learner.

This is actually a much simpler approach because note that the distributions across each of the clusters $\{v_{a,c}\}_{c \in \mathcal{C}}$ are identical and estimating one cluster distribution will reveal all the information of the users in each cluster.

3 Contributions

To be written.

4 Notations and Assumptions

Assumption 1. *We assume that there exists d-column base factors, denoted by $V(J^*, :)$, such that all row of V can be written as a convex combination of $V(J^*, :)$ and the zero vector and $J^* = [d]$. We denote the column factors by $V^* = V(J*, :)$. Therefore, for any $i \in [L]$, it can be represented by*

$$V(i, :) = a_i V(J^*, :),$$

where $\exists a_i \in [0,1]^d$ *and* $||a_i||_1 \leq 1$.

5 Problem Formulation: Generalized Latent Bandit Model

Extending the discussion in the previous section, we propose a Generalized Latent Bandit model where each $v_{a,b}$ can be considered as a mixture of several $\{v_{a,c}\}_{c \in \mathcal{C}}$. This is a harder problem than Latent Bandit model because now the rows in each cluster are not exactly identical but the index of the optimal arm (column) is similar in each row.

To define more formally, let R be matrix of dimension $K \times L$ where $|A| = K$ is the number of rows and $L = |B|$ is the number of columns. Also, let us assume that this matrix R has a low rank structure of rank $d << \min\{L, K\}$. Let U and V denote the latent matrices which are not visible to the learner such that,

$$R = UV^\intercal \qquad \textbf{s.t.} \qquad U \in [0,1]^{K \times d}, V \in [0,1]^{L \times d}$$

Furthermore, we put a constraint on V such that, $\forall j \in [L]$, $||V(j,:)||_1 \leq 1$.

6 Proposed Algorithms

To be written.

7 Main Results

Lemma 1. *For any arbitrary row* $i \in [K]$*,*

$$\arg\max_{j \in [L]} U(i,:)V(j,:)^\intercal \leq \arg\max_{j \in [d]} U(i,:)V(j,:)^\intercal.$$

8 Proofs

Proof. Considering any arbitrary row $i \in [K]$, we can show that,

$$\begin{aligned}
\arg\max_{j \in [L]} U(i,:)V(j,:)^\intercal &= U(i,:)V(j^*(i),:)^\intercal \\
&\overset{(a)}{=} U(i,:)\left(a_{j^*(i)}V(J^*,:)\right)^\intercal \\
&= \sum_{k=1}^{d} a_{j^*(i)}(k)U(i,:)V(j^*(i),:)^\intercal \\
&\leq \arg\max a_{j^*(i)}(k)U(i,:)V(k,:)^\intercal \\
&\leq \arg\max_{k \in [d]} U(i,:)V(k,:)^\intercal,
\end{aligned}$$

where (a) is from Assumption 1. □

9 Experiments

To be written.

10 Conclusions and Future Direction

To be written.

References

Maillard, O.-A. and Mannor, S. (2014). Latent bandits.