CS 6780 Research Proposal: Top N Recommender Systems

Bangrui Chen BC496@CORNELL.EDU
Saul Toscano Palmerin ST684@CORNELL.EDU
Zhengdi Shen ZS267@CORNELL.EDU

1. Motivation

We present a greedy Bayesian approach to solve the contextual bandit problem, and we study how well this method does. We will do numerical experiments and we might be able to prove some performance result.

2. Problem Formulation

In this problem, we are given a collection of items (such as movies, books, articles, etc), denote them as $I = \{I_1, I_2, \cdots, I_M\}$. For each item I_n, we use a k dimension vector $I_n = (i_{1,n}, \cdots, i_{k,n})$ to represent its score on k different features. For any time $t = 1, 2, \cdots, T$, we are asked to pick N items $S_t = \{I_{t,1}, I_{t,2}, \cdots, I_{t,N}\} \subset I$, order them and forward these N items to the user.

We assume there is a known probability vector $(p_1, \cdots, p_N)$ associated with the N different positions. Since the top positions are more likely to be seen by the user, we further assume $p_1 \geq p_2, \cdots \geq p_N$. The user has a user preference vector $\theta = \{\theta_1, \cdots, \theta_k\}$ which is unknown but remain static over the time. Every time the user sees N items, he/she will click the item $I_{t,i}$ with probability $(1 + \exp(-p_i \theta \cdot I_{t,i}))^{-1}$. If the user clicks the item, then we receive a reward 1 and 0 otherwise. Thus, the feedback from the user in this case is a N dimensional vector $(Y_{1,n}, \cdots, Y_{N,n})$ and $Y_{i,n} = 1$ if the users clicks the item on the i_{th} position and 0 otherwise. For each $Y_{i,n}$, we know

$$Y_{i,n} | \theta, I_{t,i}, p_i \sim Bernoulli\left(\frac{1}{1 + \exp(-p_i \theta \cdot I_{t,i})}\right). \quad (1)$$

Denote the total reward received at time n as $Y_n = \sum_{i=1}^{N} Y_{i,n}$.The goal of this problem is to find a strategy π to maximize the following expression

$$E^{\pi}\left[\sum_{t=1}^{T} Y_t\right] \quad (2)$$

The questions we are going to solve are the following:

(1) Suppose there are many historical users in our database and we know their user preference vectors. Then we can calculate the empirical distribution for those user preference vectors and treat it as the prior distribution of the current θ. Every time we observe the user's feedback $(Y_{1,n}, \cdots, Y_{N,n})$, we can use this feedback and maximum likelihood method to find a new estimate for θ, denote as $\hat{\theta}_n$. Then we make our next selection based on this $\hat{\theta}_n$. How good is this greedy method compared to other existing methods?

(2) Further, when we make the selection based on $\hat{\theta}_n$, what's the best trade off between exploitation and exploration?

(3) When k is large, this problem is computational intractable. How could we solve this problem efficiently? Can we prove any structural results for this problem?

3. Possible Applications

Recommender systems want to find the preference that a user would give to a subset of a finite set of items. They're widely applied to different problems. For example, they're used in Netflix where there are thousands of movies and TV episodes. The biggest challenge of these problems is that there are millions of objets and hundreds of million of users, and so it's necessary to find a model that performs well and be sufficiently fast.

4. Approach and Recourses

We will develop our method and test the performance on the dataset provided by companies like Yelp or Netflix.

5. Schedule

March 1: Test the performance of greedy method.
March 23: Trade off between expatiation and exploration.
April 20: Improve the efficiency for large data.
May 11: Finish the report.

References

1] X Zhao, P Frazier, *Exploration vs. Exploitation in the Information Filtering Problem*