Online Prediction with Bradley-Terry Models

Issei Matsumoto
Toyota Corporation
sc106049@gmail.com

Kohei Hatano
Department of Informatics
Kyushu University
hatano@inf.kyushu-u.ac.jp

Eiji Takimoto
Department of Informatics
Kyushu University
eiji@inf.kyushu-u.ac.jp

Abstract

We consider an online density estimation problem under the Bradley-Terry model which determines the probability of the order between any pair in the set of n teams. An annoying issue is that the loss function is not convex. A standard solution to the avoid the non-convexity is to change variables so that the new loss function w.r.t. new variables is convex. But, then the radius of the new domain might be huge or unknown in general, for which standard algorithms such as OGD and ONS have suboptimal regret bounds. We propose an algorithm with regret $O(\ln T)$ which works without the knowledge of the radius.

1 Introduction

Prediction problems of rankings over a set of items appear in many contexts such as information retrieval and recommendation tasks. Probabilistic models on rankings are useful for these tasks. Among probabilistic models over ranking, Bradley-Terry model [1, 2, 3] would be arguably the most fundamental one. Given a set $[n] = \{1, \ldots, n\}$ of n teams (or items), let $\boldsymbol{\Theta} = \{\boldsymbol{\theta} \in \mathbb{R}_+^n \mid \sum_{i=1}^n \theta_i = 1\}$ be the set of model parameters. In the Bradley-Terry model, given a pair of team i and j, the probability that the team i wins the team j, denoted as the ordered pair (i, j), under the parameter $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is defined as

$$P((i,j) \mid \boldsymbol{\theta}, \{i,j\}) = \frac{\theta_i}{\theta_i + \theta_j}.$$

Here, each weight θ_i can be interpreted as the strength of the player i. For simplicity, we do not consider ties.

In this paper, we consider an online density estimation problem for Bradley-Terry models with the logarithmic loss. The protocol is defined as follows: For each trial $t = 1, \ldots, T$, (i) Player guesses $\boldsymbol{\theta}_t \in \Theta$. (ii) Adversary chooses a pair of teams i_t and j_t and their game result (i_t, j_t). (iii) Player incurs the loss $f_t(\boldsymbol{\theta}_t) = -\ln P((i_t, j_t) \mid \boldsymbol{\theta}, \{i_t, j_t\}) = -\ln \frac{\theta_{i_t}}{\theta_{i_t} + \theta_{j_t}}$. The goal of Player is to minimize the regret: $Regret(T) = \sum_{t=1}^T f_t(\boldsymbol{\theta}_t) - \min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \sum_{t=1}^T f_t(\boldsymbol{\theta})$, where the second term corresponds to cumulative losses of the maximum likelihood estimator in hindsight.

There are many researches on online density estimation problems for the exponential family including Bernoulli and Gaussian (e.g., [4, 5, 6]). The exponential family has various nice properties which imply robust algorithms with $O(\ln T)$ regret bounds, say, Bayesian algorithms. But, Bradley-Terry model or logistic regression do not belong to the exponential family. So, previous works on the exponential family do not seem to be directly applicable to these models. Also, to the best of our knowledge, there is no previous result on online density estimation for Bradley-Terry model.

Further, an annoying issue is that the loss function $f_t(\boldsymbol{\theta})$ is *not* convex w.r.t. $\boldsymbol{\theta}$. Equivalent convex loss functions can be obtained by replacing the variable $\boldsymbol{\theta}$ with a new one $\boldsymbol{\gamma}$ using some bijection $\phi : \boldsymbol{\Gamma} \to \boldsymbol{\Theta}$ (see, e.g., [3]). That is, we can reduce the online density estimation problem for Bradley-Terry model with an online convex optimization problem with convex loss functions $g_t(\boldsymbol{\gamma}) = f_t(\phi(\boldsymbol{\gamma})) = f_t(\boldsymbol{\theta})$. In fact, the new loss function g_t can be viewed as a special case of the logistic function. So, by a reparametrization, our online prediction problem falls into an online logistic regression problem.

There is, however, a drawback of the reparameterization approach. The problem is that , the radius of the new domain $\boldsymbol{\Gamma}$ is unknown or infinitely large in general. Let $\boldsymbol{\gamma}^*$ be the offline minimum in hindsight over the new domain. and $K = \max_{i,j\in[n]} e^{\gamma_i^* - \gamma_j^*}$. Then, it can be shown that $\ln K \leq \|\boldsymbol{\gamma}\|_2 \leq \sqrt{2n} \ln K$. There are sequences of game results for which $K = \infty$, which implies $\|\boldsymbol{\gamma}^*\|_2 = \infty$.

When we have the knowledge on K, standard algorithms, Online Gradient Descent (OGD, [7]) and Online Newton Step (ONS) are applicable. It can be shown that, when the radius K is known in advance, Both algorithms have regret bounds O$(n^{\frac{1}{2}}(\ln K)\sqrt{T})$ and $O(nK \ln T)$, which seems suboptimal w.r.t. parameters K or T.

In this paper, we propose an algorithm for the online density estimation of Bradley-Terry models without any prior knowledge of K. Our algorithm is a Follow the Regularized Leader with a natural reguralizer, which we just call FTRL for simplicity. At each trial t, FTRL guesses the offline optimizer for the past $t-1$ trials with $n(n-1)/\eta$ "imaginary" trials where team i wins team j for $1/\eta$ times (and vice versa) for any $i \neq j$. That is, our regularizer is log loss over such additional fictitious "even" matches. We show that a regret bound $O(n^2(\ln K + \widetilde{K}) \ln T)$ for FTRL, where $\widetilde{K}$ is the maximum of radius over T past incremental offline optimizers assuming the additional $n(n-1)/\eta$ trials. Note that $\widetilde{K}$ is always bounded. Further, in our experiments, $\widetilde{K}$ is often much smaller than K. We summarize these regret bounds in Table 1.

Previous Algorithms		Our Algorithm
OGD[7]	ONS [8]	FTRL
$n^{\frac{1}{2}}(\ln K)\sqrt{T}$	$(nK + n^{\frac{3}{2}} \ln K) \ln T$	$n^2(\ln K + \widetilde{K}) \ln T$

Table 1: Comparison of regret bounds of previous algorithms and ours for Bradley-Terry model.

Unfortunately, we have not obtained any upper bound of K or $\widetilde{K}$ under mild assumptions. On the other hand, we show a finite lower bound $K = \Omega(T/n)^{n-1}$ by constructing a set of match results explicitly. Note that this lower bound implies that there is a situation where the regret of ONS becomes $O(T \ln T)$ for a fixed n, which is worse than the trivial regret bound of $O(T)$. On the other hand, the regret of FTRL is $O(\ln^2 T + \widetilde{K} \ln T)$ in that situation.

Finally, our preliminary experiments on real data sets demonstrate advantages of Bradley-Terry models over Bernoulli models for ordering tasks.

2 Preliminaries

2.1 Offline algorithms

Typical researches of the offline optimization for BT models assume that there is no "too strong player" for which $K = \infty$. In particular Hunter [3] considered the following assumption.

Assumption 1. *Let S_1 and S_2 be any partition of $[n]$, i.e., $S_1 \cup S_2 = [n]$ and $S_1 \cap S_2 = \emptyset$. Then, there exists a player $i \in S_2$ such that i beats some player in S_1 at least once.*

Under Assumption 1, K is always finite and $\|\boldsymbol{\gamma}^*\| = O(\sqrt{n} \ln K)$. In particular, Hunter proposed the algorithm called MM (Minorization-Maximization) algorithm which works on the original domain $\boldsymbol{\Theta}$. MM algorithm iteratively approximates the non-convex part of the objective with a linear function and maximizes the approximated objective. It is shown that MM algorithm converges to the offline optimum and often runs faster than Newton-Raphson method [3].

3 Algorithm without the knowledge of K

In this section, we consider the situation in which no prior knowledge on K is available. For this situation, we propose a version of Follow the Regularized Leader with some "virtual matches" as a regularizer. We simply call the second algorithm FTRL in this paper. FTRL, at each trial t, given the initial guess $\boldsymbol{\theta}_1 = \frac{1}{n}\mathbf{1}$, predicts

$$\boldsymbol{\theta}_t = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{\tau=1}^{t-1} f_t(\boldsymbol{\theta}) + \frac{1}{\eta}\sum_{i,j} f_{ij}(\boldsymbol{\theta}), \tag{1}$$

where $f_{ij}(\boldsymbol{\theta}) = -\ln \frac{\theta_i}{\theta_i+\theta_j}$. Intuitively speaking, FTRL just predicts the maximum likelihood estimator of all the past data with additional "virtual matches" in which for each team i and j, each beats the other for $1/\eta$ times.

Recall that the optimization problem (1) is not convex. However, the additional virtual matches ensures Assumption 1. Then, the solution $\boldsymbol{\theta}_t^*$is unique and we can solve the problem (1) in the original domain $\boldsymbol{\Theta}$ by, e.g., MM algorithm [3], without changing the variables. As shown in [3], MM runs faster than Newton method.

Now we explain the outline of the analysis of our second algorithm. The first idea is to analyze FTRL in the reparameterized domain $\boldsymbol{\Gamma}$, without the knowledge of K.Then the analysis, again, falls into an online logistic regression problem. The FTRL framework is well understood when each loss function is linear or its second-order approximation is available (see, e.g., [9]). But, we take an alternative approach. We will exploit the fact that in the underlying online logistic regression problem, each instance $\boldsymbol{x}_t$ is sparse, i.e., only two component has values ± 1 and other components are 0s. Based on this fact, our analysis reduces the problem into $O(n^2)$ different problems of one-dimensional online logistic regression. For each one-dimensional problem, a tight analysis can be obtained (say, as shown by McMahan and Streeter [10]).

Let $\widetilde{K} = \max_{t=1}^{T} \max_{i,j\in[n]} \theta_{t,i}/\theta_{t,j}$.It can be shown that $\widetilde{K}$ is always finite even if K is infinitely large.

We are ready to show our main result.

Theorem 1. *For Bradley-Terry model. the regret of FTRL is $O(n^2(\ln K + \widetilde{K})\ln T)$ for $\eta = 1/\ln T$.*

Implementation of FTRL As mentioned earlier, we can solve the underlying non-convex problems exactly by taking advantage of the fact that the corresponding offline problem has a unique minimum and there are algorithms to solve it under a natural assumption (see, e.g., [1, 3]). As is often the case, such algorithms like MM algorithm solve the non-convex problems faster than solving the corresponding convex problems in the reparameterized domain. In addition, we suggest to use the parameter $\eta = c/\ln T$ for some constant. In fact, the larger the value of $1/\eta$, the smaller the value of $\widetilde{K}$ is. In our experiments, we observe that $\widetilde{K}$ is almost as small as $O(\ln K)$ (not shown) with this choice of η.

Discussion on Bayesian Approach For online density estimation problems, Bayesian approaches are shown to be quite effective to get $O(\ln T)$ regret bounds (e.g., [4, 5, 6]). The typical Bayesian approach assumes a prior distribution over parameters and predicts the average of parameters w.r.t. the posterior distribution. Unfortunately, unlike the exponential family, it is not straightforward to get a regret bound for Bradly-Terry and Logistic models using this approach. Note that FTRL has a natural Bayesian interpretation that, FTRL predicts the maximum a posterior estimate of parameters w.r.t. the posterior distribution, where the prior distribution is defined as the likelihood of the virtual $2/\eta$ matches between each two players. It is an interesting open question whether a Bayesian approach achieves $O(\ln T)$ regret bound for these models.

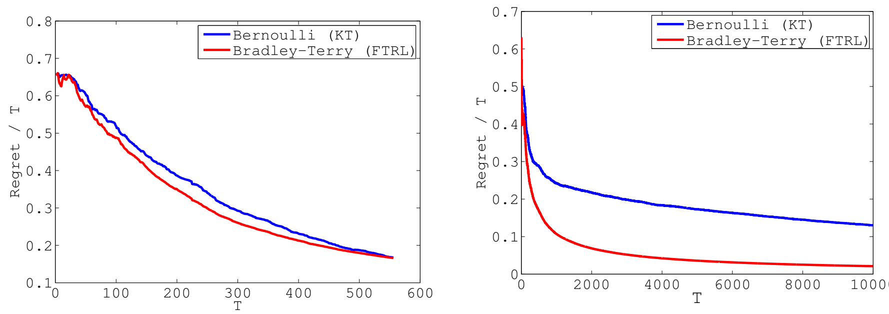

Figure 1: Regret of each algorithm for EURO2012 (left) and Movie Lense (right). The regret curve of each algorithm is obtained with its best parameter.

4 Lower Bound of the Radius K

Even when K is finite, we show a lower bound on K by explicitly constructing a set of games. The set of games is the following: 1 beats 2 for a times, 2 beats 3 for a times, ..., $n-1$ beats n for a times and n beats 1 at once. So, $T = a(n-1)+1$ Then, we can show the lower bound for this set.

Theorem 2. *There exists a set of T games for which the offline optimizer $\boldsymbol{\gamma}^*$ of Bradley-Terry model over reparameterized space $\boldsymbol{\Gamma}$ satisfies $K = \max_{i,j\in[n]} e^{\gamma_i - \gamma_j} > \left(\frac{T-1}{n-1} - 1\right)^{n-1}$.*

5 Experimental Results

In this section, we show some experimental results on real data to illustrate advantages of Bradley-Terry model to the naive Bernoulli model. Algorithms we compare are KT-estimator[11] for Bernoulli distributions for each pair and our FTRL for Bradley-Terry models. Note that the Bernoulli model is the strongest model in the sense that its offline MLE attains the least cumulative loss among any reasonable probabilistic models. In the experiments with other algorithms (OGD and ONS) for Bradley-Terry models, FTRL performs comparably or better than these algorithms (omitted).

The first real data set is the soccer game results at EURO 2012 [1] which contains 284 games among 53 teams. To deal with ties, we regard a game result that team i wins team j as "team i wins team j twice", and a tie between i and j as "team i and team j wins each other once", respectively. We randomly permute each data and give them to algorithms sequentially. The parameter η of FTRL is fixed as $\eta = p/\ln T$ with the choices of $p = 2, \ldots, 2^{10}$.

The second real data set is MovieLense data set which consists of rates of movies by users. In this experiment, we use a subset of the entire set, containing $n = 100$ movies and ratings by 100 users. To generate a sequence of matches between movies, first we pick up two movies randomly among n movies and choose a user randomly. The movie i wins the movie j if the rates of i by the user is higher than that of j. We construct $T = 10^4$ matches by this method. For both data sets, their diameter K are infty.

The results of algorithms are shown in Figure 1. Here, both regrets are computed w.r.t. the best Bradley-Terry model in hindsight. In both experiments, results are averaged over 10 repeats of these procedures. For each algorithm, its average regret with the best parameter is shown. In both experiments, FTRL performs better than KT estimator. These results make sense. KT estimator predicts game results on teams i and j without considering games between other players. So, KT estimator predicts $\Pr\{i \text{ beats } j\} = 1/2$ when there is no game between i and j. On the other hand, Bradley-Terry model takes into account other game results. So, even when game results are sparse, it can infer the strength of teams i and j from other games. In particular, FTRL for Bradley-Terry models performs much better than the other, since MovieLense data has a much sparse structure.

[1] We collect the game results from the website `http://archive.sportsnavi.yahoo.co.jp/soccer/euro/12/index.html`.

References

[1] L.R. Ford. Solution of a ranking problem from binary comparisons. *American Mathematical Montnly*, 64(8):28—-33, 1957.

[2] John I. Marden. *Analyzing and Modeling Rank Data*. Chapman & Hall, 1995.

[3] D.R Hunter. MM algorithms for generalized Bradley-Terry models. *Annals of Statistics*, 32(1):384–406, 2004.

[4] Yoav Freund. Predicting a binary sequence almost as well as the optimal biased coin. In *Proceedings of the Ninth Annual Conference on Computational Learning Theory (COLT '96)*, pages 89–98, 1996.

[5] K S Azoury and Manred K. Warmuth. Relative Loss Bounds for On-Line Density Estimation with the Exponential Family of Distributions. *Machine Learning*, 43:211–246, 2001.

[6] Wojciech Kotlowski, Peter Grünwald, and Steven de Rooij. Following the Flattened Leader. In *Proceedings of 23th Annual Conference on Learning Theory (COLT 2010)*, pages 106–118, 2010.

[7] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML 2003)*, pages 928–936, 2003.

[8] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3):169–192, 2007.

[9] Elad Hazan. The convex optimization approach to regret minimization. In Suvrit Sra, Sebastian Nowozin, and Stephen J. Wright, editors, *Optimization for Machine Learning*, chapter 10, pages 287–304. MIT Press, 2011.

[10] H. Brendan McMahan and Matthew J. Streeter. Open Problem: Better Bounds for Online Logistic Regression. In *Proceeding of the 25th Annual Conference on Learning Theory (COLT 2012)*, pages 44.1–44.3, 2012.

[11] R. E. Krichevsky and V. K. Trofimov. The performance of universal encoding. *IEEE Transactions on Information Theory*, 27(2):199–207, 1981.