JMLR: Workshop and Conference Proceedings vol 40:1–17, 2015

Online Density Estimation of Bradley-Terry Models

Issei Matsumoto SC106049@GMAIL.COM
Toyota Motor Corporation

Kohei Hatano HATANO@INF.KYUSHU-U.AC.JP
Eiji Takimoto EIJI@INF.KYUSHU-U.AC.JP
Department of Informatics, Kyushu University

Abstract

We consider an online density estimation problem for the Bradley-Terry model, where each model parameter defines the probability of a match result between any pair in a set of n teams. The problem is hard because the loss function (i.e., the negative log-likelihood function in our problem setting) is not convex. To avoid the non-convexity, we can change parameters so that the loss function becomes convex with respect to the new parameter. But then the radius K of the reparameterized domain may be infinite, where K depends on the outcome sequence. So we put a mild assumption that guarantees that K is finite. We can thus employ standard online convex optimization algorithms, namely OGD and ONS, over the reparameterized domain, and get regret bounds $O(n^{\frac{1}{2}}(\ln K)\sqrt{T})$ and $O(n^{\frac{3}{2}}K \ln T)$, respectively, where T is the horizon of the game. The bounds roughly means that OGD is better when K is large while ONS is better when K is small. But how large can K be? We show that K can be as large as $\Theta(T^{n-1})$, which implies that the worst case regret bounds of OGD and ONS are $O(n^{\frac{3}{2}}\sqrt{T} \ln T)$ and $\tilde{O}(n^{\frac{3}{2}}(T)^{n-1})$, respectively.

We then propose a version of Follow the Regularized Leader, whose regret bound is close to the minimum of those of OGD and ONS. In other words, our algorithm is competitive with both for a wide range of values of K. In particular, our algorithm achieves the worst case regret bound $O(n^{\frac{5}{2}}T^{\frac{1}{3}} \ln T)$, which is slightly better than OGD with respect to T. In addition, our algorithm works without the knowledge K, which is a practical advantage.

Keywords: online density estimation, Bradley-Terry model, ranking

1. Introduction

Prediction problems of ranking over a set of items appear in many contexts, such as information retrieval and recommendation tasks. Probabilistic modeling of rankings is useful for such tasks. The Bradley-Terry model (Ford, Jr., 1957; Marden, 1995; Hunter, 2004) is arguably the most fundamental model. Given a set $[n] = \{1, \ldots, n\}$ of n teams (or items), let $\Theta = \{\boldsymbol{\theta} \in \mathbb{R}_+^n \mid \sum_{i=1}^n \theta_i = 1\}$ be the set of model parameters. In the Bradley-Terry model, given a pair of team i and j, the probability that team i beats team j, denoted as the ordered pair (i, j), under the parameter $\boldsymbol{\theta} \in \Theta$ is defined as follows:

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

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

Algorithm	Regret Upper Bounds		
	General case (K: arbitrary)	Easy case ($K, \widetilde{K}$: constant)	Hard case ($K, \widetilde{K} : O(T^{n-1})$)
OGD (Zinkevich, 2003)	$n^{\frac{1}{2}}(\ln K)\sqrt{T}$	$n^{\frac{1}{2}}\sqrt{T}$	$n^{\frac{3}{2}}\sqrt{T}\ln T$
ONS (Hazan et al., 2007)	$n^{\frac{3}{2}}K\ln T$	$n^{\frac{3}{2}}\ln T$	$n^{\frac{3}{2}}T^{n-1}\ln T$
a version of FTRL (this paper)	$\boldsymbol{n \min\left\{(\ln \widetilde{K})\sqrt{\frac{T}{\lambda}}, \widetilde{K}\ln T\right\} + \lambda n^2(\ln K)}$	$\boldsymbol{n \ln T}$	$\boldsymbol{n^{\frac{5}{2}}T^{\frac{1}{3}}\ln T}$

Table 1: Comparison of regret bounds of the proposed and previous algorithms for Bradley-Terry models. Better bounds w.r.t. T are shown in bold face. For OGD and ONS, the radius K needs to be known a priori.

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 $t = 1, \ldots, T$.

1. The player guesses a parameter $\boldsymbol{\theta}_t \in \Theta$.
2. The adversary chooses a pair of teams i_t and j_t and their game result (i_t, j_t), meaning that team i_t beats team j_t.
3. The player incurs the loss $f_t(\boldsymbol{\theta}_t) = -\ln P((i_t, j_t) \mid \boldsymbol{\theta}_t, \{i_t, j_t\}) = -\ln \frac{\theta_{i_t}}{\theta_{i_t}+\theta_{j_t}}$.

The goal of the player is to minimize the regret: $Regret(T) = \sum_{t=1}^{T} f_t(\boldsymbol{\theta}_t) - \min_{\boldsymbol{\theta}\in\Theta} \sum_{t=1}^{T} f_t(\boldsymbol{\theta})$, where the second term corresponds to the cumulative loss of the maximum likelihood estimator in hindsight.

Many studies have examined online density estimation problems for the exponential family including Bernoulli and Gaussian (e.g., Shtar'kov (1987); Freund (1996); Takimoto and Warmuth (2000); Azoury and Warmuth (2001); Kotlowski et al. (2010)). The exponential family has various nice properties that imply robust algorithms with $O(\ln T)$ regret bounds. However, previous work on the exponential family does not appear to be directly applicable to Bradley-Terry models and logistic regression. In addition, to the best of our knowledge, online density estimation for Bradley-Terry models has not been previously presented.

One issue is that the loss function $f_t(\boldsymbol{\theta})$ is *not* convex w.r.t. $\boldsymbol{\theta}$. An equivalent convex loss function can be obtained by replacing the variable $\boldsymbol{\theta}$ with a new variable $\boldsymbol{\gamma}$ using some bijection $\phi : \Gamma \to \Theta$ (see, e.g., Hunter (2004)). In other words, we can reduce the online density estimation problem for Bradley-Terry models to 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 loss function. Thus, by the reparametrization, our online prediction problem becomes an online logistic regression problem.

However, there is a drawback to the reparametrization approach, i.e., the radius of the new domain Γ is unknown and infinitely large in general. Let $\boldsymbol{\gamma}^*$ be the offline minimizer in hindsight over the new domain and let $K = \max_{i,j\in[n]} e^{\gamma_i^* - \gamma_j^*}$. Then, it can be shown that $\ln K \le 2\|\boldsymbol{\gamma}^*\|_2 = O(\sqrt{n}\ln K)$. Note that there are sequences of game results for which $K = \infty$, which implies $\|\boldsymbol{\gamma}^*\|_2 = \infty$.

When we have knowledge of K, standard algorithms are applicable, such as Online Gradient Descent (OGD, Zinkevich (2003)) and Online Newton Step (ONS). When the radius K is known, both algorithms can be shown to have regret bounds $\mathrm{O}(n^{\frac{1}{2}}(\ln K)\sqrt{T})$ and $O(n^{\frac{3}{2}}K\ln T)$, which seems suboptimal w.r.t. parameters K or T. In addition, these regret bounds are not meaningful when the radius K is infinitely large.

A natural question would be how large is the radius K when K is finite? In this paper, we first give a complete answer to this question. We show that there exists a set of matches for which the radius K of the reparameterized space must be $\Omega((T/n)^{n-1})$. In addition, under a weak assumption that, for each pair of teams i and j, team i beats team j a constant number of times (and vice versa), the radius K is bounded above as $O(T/n)^{n-1}$. Thus, the bound is tight. We believe that the tight bound of the radius is valuable in its own right. Our analysis of the radius implies that the radius can grow polynomially in T! This is quite strange since the radius is assumed constant in the standard online prediction literature. As a result, when K is constant (we call this the "easy" case), the regret of ONS is $O(n^{\frac{3}{2}}\ln T)$ which is better than the bound $O(n^{\frac{1}{2}}\sqrt{T})$ of OGD in terms of T. But, when $K = O(T^{n-1})$ (the "hard" case), the regret bound of OGD becomes $O(n^{\frac{3}{2}}\sqrt{T}\ln T)$, while that of ONS increases up to $O(nT^{n-1}\ln T)$.

We then propose an algorithm for the online density estimation of Bradley-Terry models that performs well in both easy and hard cases. The proposed algorithm is a Follow the Regularized Leader with a natural regularizer, which we just call FTRL for simplicity. At each trial t, FTRL guesses the offline optimizer for the past $t-1$ trials with $\lambda n(n-1)$ "virtual" matches where team i beats team j for λ times (and vice versa) for any $i \neq j$. That is, our regularizer is log loss over such additional fictitious "even" matches. For FTRL with any parameter $\lambda > 0$, we show a regret bound $O(\min\{n^{\frac{3}{2}}(\ln \widetilde{K})\sqrt{T/\lambda}, n\widetilde{K}\ln T\} + \lambda n^2(\ln K))$, where $\widetilde{K}$ is the maximum of radius of the set of guesses produced by FTRL. Therefore, FTRL performs competitively in both easy and hard cases. Our first result also guarantees that $\widetilde{K}$ is bounded as $O(T^{n-1})$. Furthermore, by tuning the parameter λ appropriately, we obtain the regret bound $O(n^{\frac{5}{2}}T^{\frac{1}{3}}\ln T)$ of FTRL, which improves OGD in the worst case w.r.t. T. Regret bounds are summarized in Table 1. In addition, the proposed algorithm works without any prior knowledge of K or $\widetilde{K}$, which is a practical advantage.

2. Preliminaries

2.1. Reparametrization and Offline Algorithms

Recall that in our problem, each loss function $f_t(\boldsymbol{\theta})$ is not convex w.r.t. θ. Yet, by simple reparametrization (see, e.g., Hunter (2004)), an equivalent convex formulation can be obtained. Let $\gamma_i = \ln\theta_i - \ln\theta_1$ for $i = 1, \dots, n$. Then, $\theta_i = e^{\gamma_i}/\sum_{i=1}^n e^{\gamma_i}$. This mapping is a bijection between $\Theta = \{\boldsymbol{\theta} \in \mathbb{R}_+^n \mid \sum_i \theta_i = 1\}$ and $\Gamma = \{\boldsymbol{\gamma} \in \mathbb{R}^n \mid \gamma_1 = 0\}$. Then, observe that

$$f_t(\boldsymbol{\theta}) = -\ln\frac{\theta_{i_t}}{\theta_{i_t} + \theta_{j_t}} = -\ln\frac{e^{\gamma_{i_t}}}{e^{\gamma_{i_t}} + e^{\gamma_{j_t}}} = \ln(1 + e^{\boldsymbol{\gamma}\cdot\boldsymbol{x}_t}) \stackrel{\text{def}}{=} g_t(\boldsymbol{\gamma}),$$

where $\boldsymbol{x}_t \in \mathbb{R}^n$ is such that $x_{i_t} = -1$, $x_{j_t} = 1$, and $x_k = 0$ for $k \in [n]$ s.t. $k \neq i_t, j_t$. Here, g_t is convex w.r.t. the new variable $\boldsymbol{\gamma} \in \Gamma$. Note that, g_t can be viewed as a logistic loss function for sparse instances with the specific form. Therefore, by this reparametrization, online density estimation for Bradley-Terry models is reduced to an online logistic regression problem where each loss function is defined as g_t.

Unfortunately, there is a drawback with the reparametrization scheme. For the corresponding offline optimizer $\boldsymbol{\gamma}^* = \arg\min_{\boldsymbol{\gamma}\in\Gamma} \sum_{t=1}^{T} g_t(\boldsymbol{\gamma})$, let $\boldsymbol{\theta}^*$ be the corresponding optimizer in the original domain Θ and let

$$K = \max_{i,j\in[n]} \theta_i^*/\theta_j^* = \max_{i,j\in[n]} e^{\gamma_i^*-\gamma_j^*}.$$

Note that $\ln K = \max_{i,j\in[n]} |\gamma_i^* - \gamma_j^*| \le 2\|\boldsymbol{\gamma}^*\|_2$ and $\|\boldsymbol{\gamma}^*\|_2 = O(\sqrt{n}\ln K)$. Generally, K and $\|\boldsymbol{\gamma}^*\|_2$ are infinitely large. For example, consider the case where $n = 2$ and only the event $(1, 2)$ is observed. Then, the offline optimum $\boldsymbol{\theta}^* = (1, 0)$; however, the corresponding optimum $\boldsymbol{\gamma}^*$ in the new domain is $\boldsymbol{\gamma}^* = (0, -\infty)$. So, $K = \|\boldsymbol{\gamma}^*\|_2 = \infty$. In addition, for the case with T games where there exists a team who beats others and is not beaten by anyone, $K = \infty$. Infinitely large domain is not desirable especially for regret analyses, which typically requires knowledge of the diameter of the domain $\|\boldsymbol{\gamma}^*\|_2$.

Typical research of offline optimization of Bradley-Terry models assumes that there is no "too strong team" for which $K = \infty$. In particular, Hunter (2004) 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 team $i \in S_2$ such that i beats some team 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 Minorization-Maximization (MM) algorithm, which works on the original domain Θ. The MM algorithm iteratively approximates the non-convex part of the objective with a linear function and maximizes the approximated objective. The MM algorithm has been shown to converge to the offline optimum and often runs faster than the Newton-Raphson method (Hunter, 2004).

In the following, for simplicity, we sometimes neglect the constraint $\gamma_1 = 0$ for Γ and simply assume $\Gamma = \mathbb{R}^n$.

2.2. Application of Existing Algorithms

In this subsection, we assume that the parameter K is known. Under this assumption, we review OGD (Zinkevich, 2003) and ONS (Hazan et al., 2007), as well as their applications to the Bradley-Terry Model. Here we consider the bounded reparameterized space $\Gamma_K = \{\boldsymbol{\gamma} \in \mathbb{R}^n \mid |\gamma_i - \gamma_j| \le \ln K \text{ for any } i, j \in [n]\}$.

The standard algorithm OGD has the regret bound $O(GD\sqrt{T})$ when its learning rate $\eta = G/(D\sqrt{T})$. where $\boldsymbol{\Gamma} = \{\boldsymbol{\gamma} \in \mathbb{R}^n \mid \|\boldsymbol{\gamma}\|_2 \le D\}$ and $\|\nabla g_t(\boldsymbol{\gamma})\|_2 \le G$ for any $t = 1, \dots, T$ and $\boldsymbol{\gamma} \in \boldsymbol{\Gamma}$. For our problem, it can be shown that $G = \max_{t=1}^{T} \|\nabla g_t(\boldsymbol{\gamma})\| \le \sqrt{2}$ and $D = O(\sqrt{n}\ln K)$. Thus, when K is known, the regret of OGD for our problem is $O(n^{\frac{1}{2}}(\ln K)\sqrt{T})$. If the loss functions are m-strongly convex, i.e., $\nabla^2 g_t(\boldsymbol{\gamma}) \succ mI$ for some $m > 0$ and the identity matrix I, then, OGD with a different learning parameter setting has the regret bound $O((G/m)\ln T)$. However, for Bradley-Terry models or logistic regression models, loss functions g_t are not strongly convex and the improved $O(\ln T)$ bound is not applicable.

ONS is another popular algorithm for online convex optimization tasks, designed especially for when each loss function g is α-exp concave, i.e., $\exp(-\alpha g(\boldsymbol{\gamma}))$ is concave for any $\boldsymbol{\gamma} \in \Gamma$. It is known that m-strong convexity implies α-exp concavity for $\alpha \leq m/G^2$ (see,e.g., Hazan et al. (2007)). Thus, α-exp concavity is a weaker assumption than strong convexity.

When each loss function g_t is α-exp concave, ONS is known to achieve $O((\frac{1}{\alpha} + GD)n \ln T)$ regret (Hazan et al., 2007). In particular, it can be shown that g_t or logistic loss is $1/K$-exp concave (see, e.g., McMahan and Streeter (2012)). Thus, the regret of ONS for Bradley-Terry models is $O((K + \sqrt{n} \ln K)n \ln T)$.

2.3. Relationship with Online Logistic Regression

Recently, Hazan et al. (2014) showed a lower bound $\Omega(\sqrt{DT})$ of the regret for online logistic regression. But, the proof of their lower bound does not hold for online density estimation of Bradley-Terry models, because it assumes any vectors in a bounded Euclidean ball as instances, while the instances transformed from Bradley-Terry models have a restricted form wherein only a pair of components takes values ± 1 and others are zeros. In addition, our analysis exploits the structures of instances and is not applicable to general online logistic regression problems. Therefore, our regret bounds do not contradict their results.

In particular, for a one-dimensional online logistic regression problem with instances $\boldsymbol{x} \in \{-1, 0, 1\}$, McMahan and Streeter (2012) showed that FTRL with a virtual match regularization achieves $O(\sqrt{D} + \ln T)$ regret bound. The algorithm discussed later is a multi-dimensional extension of their algorithm. Furthermore, online density estimation for Bradley-Terry models can be reduced to $(n-1)$-dimensional online logistic regression; thus their algorithm also obtains $O(\ln K + \ln T)$ for Bradley-Terry models with $n = 2$.

3. Bounds on the Radius K

In this section, we derive the upper and lower bounds of the radius K.

3.1. Lower Bound of the Radius K

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

Theorem 1 *There exists a set of T games for which the offline optimizer $\boldsymbol{\theta}^*$ of the Bradley-Terry model satisfies*

$$K = \max_{i,j \in [n]} \theta_i^* / \theta_j^* > \left(\frac{T-1}{n-1} - 1 \right)^{n-1}.$$

The proof is shown in the Appendix.

3.2. Upper Bound of the Radius K under a Mild Condition

Next, we derive the upper bound of the radius K under mild assumptions that satisfy Assumption 1. We begin with a general lemma that holds for arbitrary games.

Lemma 1 *Assume any set of T games and that the maximum likelihood parameter $\boldsymbol{\theta}^*$ of the Bradley-Terry model satisfies $\theta_1^* \geq \theta_2^* \geq \cdots \geq \theta_n^*$ without loss of generality. Then,*

$$K = \max_{i,j} \frac{\theta_i^*}{\theta_j^*} \leq \prod_{i=1}^{n-1} \frac{c(\{1,\ldots,i\},\{i+1,\ldots,n\})}{c(\{i+1,...,n\},\{1,\ldots,i\})},$$

where $c(U,V)$ denotes the total number of wins of teams in U against teams in V.

Proof For each $i,j \in [n]$, let T_{ij} be the number of games where team i beats team j. We denote T_i as $T_i = \sum_{j \neq i} T_{ij}$ and $T_{\{i,j\}} = T_{ij} + T_{ji}$, respectively. Furthermore, for $i = 1,\ldots,n-1$, let $T_{1:i,i+1:n}$ be the number of matches between teams in $\{1,\ldots,i\}$ and teams in $\{i+1,\ldots,n\}$. Let $L(\boldsymbol{\theta}) = \sum_{t=1}^{T} f_t(\boldsymbol{\theta})$. The maximum likelihood estimate $\boldsymbol{\theta}$ satisfies $\nabla L(\boldsymbol{\theta}) = 0$. We can then obtain

$$\sum_{k \neq j} T_{\{j,k\}} \frac{\theta_j}{\theta_j + \theta_k} = T_j, \tag{1}$$

for $j = 1,\ldots,n$. By summing up equation (1) for $j = 1,\ldots,i$,

$$\sum_{j=1}^{i} \sum_{k \in \{1,\ldots,i\} \setminus \{j\}} T_{\{j,k\}} \frac{\theta_j}{\theta_j + \theta_k} + \sum_{j=1}^{i} \sum_{k=i+1}^{n} T_{\{j,k\}} \frac{\theta_j}{\theta_j + \theta_k} = \sum_{j=1}^{i} T_j. \tag{2}$$

The first term of the left hand side of equation (2) is exactly $\sum_{j<k\leq i} T_{\{j,k\}}$. Since $f(x) = x/(x+a)$ for $x \geq 0$ and $a > 0$ monotonically increases, $\theta_j/(\theta_j + \theta_k) \geq \theta_i/(\theta_i + \theta_k)$ for $j < i$ and any k. This fact implies that the second term of the left hand side is lower bounded as

$$\begin{aligned}\sum_{j=1}^{i} \sum_{k=i+1}^{n} T_{\{j,k\}} \frac{\theta_j}{\theta_j + \theta_k} &\geq \sum_{j=1}^{i} \sum_{k=i+1}^{n} T_{\{j,k\}} \frac{\theta_i}{\theta_i + \theta_k} \geq \sum_{j=1}^{i} \sum_{k=i+1}^{n} T_{\{j,k\}} \frac{\theta_i}{\theta_i + \theta_{i+1}} \\ &= \frac{\theta_i}{\theta_i + \theta_{i+1}} \sum_{j=1}^{i} \sum_{k=i+1}^{n} T_{\{j,k\}} = \frac{\theta_i}{\theta_i + \theta_{i+1}} T_{\{1:i,i+1:n\}}, \end{aligned} \tag{3}$$

where the inequality holds because $\theta_i/(\theta_i + \theta_k) \geq \theta_i/(\theta_i + \theta_{i+1})$ for $k \geq i+1$ since $\theta_k \leq \theta_i$. By combining equation (2) and inequality (3), we obtain

$$\frac{\theta_i}{\theta_i + \theta_{i+1}} T_{\{1:i,i+1:n\}} \leq \sum_{j=1}^{i} T_j - \sum_{j<k\leq i} T_{\{j,k\}}.$$

Equivalently,

$$\frac{\theta_i}{\theta_{i+1}} \leq \frac{\sum_{j=1}^{i} T_j - \sum_{j<k\leq i} T_{\{j,k\}}}{T_{\{1:i,i+1:n\}} - \left(\sum_{j=1}^{i} T_j - \sum_{j<k\leq i} T_{\{j,k\}}\right)} = \frac{c(\{1,\ldots,i\},\{i+1,\ldots,n\})}{c(\{i+1,...,n\},\{1,\ldots,i\})}.$$

Finally, we obtain

$$\frac{\theta_1}{\theta_n} \leq \prod_{i=1}^{n-1} \frac{\theta_i}{\theta_{i+1}} \leq \prod_{i=1}^{n-1} \frac{c(\{1,\ldots,i\},\{i+1,\ldots,n\})}{c(\{i+1,...,n\},\{1,\ldots,i\})},$$

as claimed. ■

Now we consider two specific cases where Assumption 1 holds.

With λ-cyclic matches: T arbitrary matches with additional $n\lambda$ cyclic matches where team 1 beats team 2 λ times, team 2 beats team 3 λ times, $\cdots$, team $n-1$ beats team n λ times, and team n beats team 1 λ times, respectively.

With λ-complete matches : T arbitrary matches with additional $n(n-1)\lambda$ matches where for each team i and j, team i beats team j λ times (and vise versa).

Theorem 2 *(i) For any T matches with λ-cyclic matches,*

$$K \leq \left(\frac{\frac{T}{n-1}+\lambda}{\lambda}\right)^{n-1} = \left(\frac{T}{(n-1)\lambda}+1\right)^{n-1}$$

(ii) For any T matches with λ-complete matches,

$$K \leq \prod_{i=1}^{n-1}\left(\frac{\frac{T}{n-1}+i(n-i)\lambda}{i(n-i)\lambda}\right) \leq \left(\frac{T}{(n-1)^2\lambda}+1\right)^{n-1}.$$

Proof The right hand side of Lemma 1 is maximized when T matches are distributed uniformly over $n-1$ factors. For any set of i teams, (i) there are λ wins against the other set and (ii) $\lambda i(n-i)$ wins against the other set, respectively. These facts imply the result. ■

Note that the bound of Theorem 2 is tight in general, since the upper bound of K for T games with 1-cyclic games matches the lower bound of Theorem 1. Theorem 2 states that adding λ-complete matches are more effective by roughly $1/(n-1)^{n-1}$ times than the cyclic matches. In addition, $\lambda/(n-1)$ complete matches induce the same upper bound of the radius obtained by λ-cyclic matches.

4. FTRL with "Virtual Match" Regularization

Here, we consider the situation wherein prior knowledge of K is unsavailable. For this situation, we propose a variant of FTRL, whose regularizer is defined using some "virtual even matches." FTRL, at each trial t, given the initial guess $\boldsymbol{\theta}_1 = \frac{1}{n}\mathbf{1}$ and a parameter $\lambda > 0$, predicts

$$\boldsymbol{\theta}_t = \arg\min_{\boldsymbol{\theta}\in\Theta} \sum_{\tau=1}^{t-1} f_\tau(\boldsymbol{\theta}) + \lambda \sum_{i\neq j} f_{ij}(\boldsymbol{\theta}), \tag{4}$$

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

Recall that the optimization problem (4) is not convex. However, the additional virtual matches ensure Assumption 1. Then, the solution $\boldsymbol{\theta}_t^*$ is unique and we can use previously proposed algorithms, e.g., the MM algorithm (Hunter, 2004) to solve the problem (4) in the original domain

Θ without changing the variables. Problem (4) can also be solved in the reparameterized convex domain using standard approaches.

Let
$$\widetilde{K} = \max_{t=1}^{T} \max_{i,j\in[n]} \theta_{t,i}/\theta_{t,j} = \max_{t=1}^{T} \max_{i,j\in[n]} e^{\gamma_{t,i}-\gamma_{t,j}}.$$

As shown in Theorem 2, $\widetilde{K}$ is always finite, even if K is infinitely large.

An outline of the analysis of FTRL is presented in the following. We analyze FTRL in the reparameterized domain Γ. Then the analysis, again, becomes an online logistic regression problem. Our analysis is close to the approach adopted by Azoury and Warmuth (2001), who investigated FTRL for online density estimation with the exponential family of distributions. However, our analysis is more specialized for Bradley-Terry models.

Here, we use the notion of Bregman divergence. Given a convex function $g : \Gamma \to \mathbb{R}$, the Bregman divergence $\Delta_g(\boldsymbol{p}, \boldsymbol{q})$ is defined as $\Delta_g(\boldsymbol{p}, \boldsymbol{q}) = g(\boldsymbol{p}) - g(\boldsymbol{q}) - \nabla g(\boldsymbol{q}) \cdot (\boldsymbol{p} - \boldsymbol{q})$. Let $\Phi_{t-1}(\boldsymbol{\gamma}) = \sum_{\tau=1}^{t-1} g_\tau(\boldsymbol{\gamma}) + \lambda \sum_{ij} g_{ij}(\boldsymbol{\gamma})$, where $g_{ij}(\boldsymbol{\gamma}) = \ln(1 + e^{\gamma_j - \gamma_i})$, and let $\boldsymbol{\gamma}_t = \inf_{\boldsymbol{\gamma}\in\Gamma} \Phi_{t-1}(\boldsymbol{\gamma})$. The following lemma, which is similar to the one proved by Azoury and Warmuth (2001), is a key tool for the analysis of our algorithm.

Lemma 2 *For each $t = 1, \dots, T$,*

$$\Phi_t(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t) \geq g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t).$$

Proof Using the fact that $\Phi_t(\boldsymbol{\gamma}_{t+1}) = g_t(\boldsymbol{\gamma}_{t+1}) + \Phi_{t-1}(\boldsymbol{\gamma}_{t+1})$, we have

$$\begin{aligned}\Phi_t(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t) &\geq g_t(\boldsymbol{\gamma}_{t+1}) + \Phi_{t-1}(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t) - \nabla\Phi_{t-1}(\boldsymbol{\gamma}_t) \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) \\ &= g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t),\end{aligned}$$

where the inequality holds because $\boldsymbol{\gamma}_t$ is the minimizer of Φ_{t-1}; thus, $\nabla\Phi_{t-1}(\boldsymbol{\gamma}_t) \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) \geq 0$ (see, e.g., Boyd and Vandenberghe (2004), p. 139) and the second equation holds by definition of Bregman divergence. ■

Next, we prove a regret bound of FTRL.

Theorem 3 *For Bradley-Terry models, the regret of FTRL is $O(n\widetilde{K} \ln T + \lambda n^2(\ln K))$.*

The proof is shown in the Appendix.

Unfortunately, the regret bound of Theorem 3 is not tight enough w.r.t. the radius $\widehat{K}$. Therefore, we need an alternative approach to analyze FTRL. We exploit the particular property of the convex function called generalized concordance proposed by Bach (Bach, 2010, 2014). We say that the function g is *R-generalized concordant*[1] (w.r.t. the infinity norm) if for any $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^n$ and any $t \in \mathbb{R}$, $q(t) = g(\boldsymbol{w} + t\boldsymbol{v})$ satisfies $|q'''(t)| \leq R\|\boldsymbol{v}\|_\infty q''(t)$. Note that the generalized concordant property is slightly different from the standard self-concordant property, where g is self-concordant if $|q'''(t)| \leq Rq''(t)^{\frac{3}{2}}$. In fact, for Bradley-Terry models, the loss function g in the reparameterized domain Γ is R-generalized concordant for some constant R.

Proposition 1 *For each $t = 1, \dots, T$, $g_t(\boldsymbol{\gamma})$ is 2-generalized concordant.*

1. We slightly modified Bach's original definition. The original definition is given in terms of the 2-norm.

The proof is given in the Appendix.

The generalized self-concordant property ensures that the loss function has a tighter second order approximation than typical loss functions.

Proposition 2 *If $g(\boldsymbol{\gamma}) : \mathbb{R}^n \to \mathbb{R}$ is R-generalized concordant, then for any $\boldsymbol{\gamma}, \boldsymbol{\gamma}' \in \mathbb{R}^n$,*

$$g(\boldsymbol{\gamma}) \geq g(\boldsymbol{\gamma}') + \nabla g(\boldsymbol{\gamma}')(\boldsymbol{\gamma} - \boldsymbol{\gamma}') + \frac{1}{2\max\{e^2, 4R\|\boldsymbol{\gamma} - \boldsymbol{\gamma}'\|_\infty\}}(\boldsymbol{\gamma} - \boldsymbol{\gamma}')^\top \nabla^2 g(\boldsymbol{\gamma}')(\boldsymbol{\gamma} - \boldsymbol{\gamma}').$$

The proof is given in the Appendix.

Lemma 3 *For each $t = 1, \dots, T$,*

$$g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t) \geq g_t(\boldsymbol{\gamma}_t) - \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}}\lambda} \tag{5}$$

where $t_{\{i_t,j_t\}}$ is the number of matches between team i_t and team j_t.

Proof Let $h(x) = \ln(1 + e^x)$. Then, $g_t(\boldsymbol{\gamma}_t) = h(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_t)$. Note that $\nabla g_t(\boldsymbol{\gamma}) = h'(\boldsymbol{\gamma} \cdot \boldsymbol{x}_t)\boldsymbol{x}_t$ and

$$\nabla^2 \Phi_t(\boldsymbol{\gamma}) = \sum_{\tau=1}^{t} h''(\boldsymbol{\gamma} \cdot \boldsymbol{x}_\tau)\boldsymbol{x}_\tau \boldsymbol{x}_\tau^\top + \lambda \sum_{i,j} h''(\boldsymbol{\gamma} \cdot \boldsymbol{x}_{(ij)})\boldsymbol{x}_{(ij)}\boldsymbol{x}_{(ij)}^\top,$$

where $g_{ij}(\boldsymbol{\gamma}) = \ln(1 + e^{\boldsymbol{\gamma} \cdot \boldsymbol{x}_{(ij)}})$, $\boldsymbol{x}_{(ij)}$ is such that $x_{(ij),i} = -1$, $x_{(ij),j} = 1$ and $x_{(ij),k} = 0$ for $k \neq i, j$, respectively. Note that $\Phi_t(\boldsymbol{\gamma})$ is 2-generalized concordant by Proposition 1 and the fact that the concordant property is closed under summation. Then, using Proposition 2 around $\boldsymbol{\gamma}_t$, the l.h.s. of inequality (5) is given as

$$\begin{aligned} & g_t(\boldsymbol{\gamma}_{t+1}) + \Phi_{t-1}(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t) - \nabla \Phi_{t-1}(\boldsymbol{\gamma}_t) \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) \\ & \geq g_t(\boldsymbol{\gamma}_t) + \nabla g_t(\boldsymbol{\gamma}_t) \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) + \frac{1}{2\max\{e^2, 8\ln \widetilde{K}\}}(\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t)^\top \nabla^2 \Phi_t(\boldsymbol{\gamma}_t)(\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t), \end{aligned} \tag{6}$$

Then, since $h'(x) = e^x/(1 + e^x)$ and $h''(x) = e^x/(1 + e^x)^2$, $h''(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau) = h'(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau)^2 e^{-\boldsymbol{z}_t \cdot \boldsymbol{x}_\tau}$. Therefore,

$$\nabla^2 g_\tau(\boldsymbol{\gamma}_t) = h''(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau)\boldsymbol{x}_\tau \boldsymbol{x}_\tau^\top = e^{-\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau}(h'(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau))^2 \boldsymbol{x}_\tau \boldsymbol{x}_\tau^\top = e^{-\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau} \nabla g_\tau(\boldsymbol{\gamma}_t) \nabla g_\tau(\boldsymbol{\gamma}_t)^\top.$$

Now the second term of the r.h.s. of inequality (6) is

$$\frac{1}{2\max\{e^2, 8\ln \widetilde{K}\}}(\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t)^\top B_t(\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t),$$

where $B_t = \sum_{\tau=1}^{t} e^{-\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_\tau} \nabla g_\tau(\boldsymbol{\gamma}_t) \nabla g_\tau(\boldsymbol{\gamma}_t)^\top + \lambda \sum_{i,j} e^{-\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_{(ij)}} \nabla g_{ij}(\boldsymbol{\gamma}_t) \nabla g_{ij}(\boldsymbol{\gamma}_t)^\top$. Let t_{ij} be the number of team i' wins against team j in t trials (including wins in virtual matches). Then, by using the fact that $\boldsymbol{x}_{(ij)} = -\boldsymbol{x}_{(ji)}$ for each $i, j \in [n]$, B_t can be written as

$$B_t = \sum_{i<j} (t_{ij} e^{-\gamma_t \cdot \boldsymbol{x}_{(ij)}} + t_{ji} e^{\gamma_t \cdot \boldsymbol{x}_{(ij)}}) \nabla g_{ij}(\boldsymbol{\gamma}_t) \nabla g_{ij}(\boldsymbol{\gamma}_t)^\top.$$

Recall that $\nabla g_{ij}(\boldsymbol{\gamma}_t) = h'(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_{(ij)})\boldsymbol{x}_{(ij)}$. Then, by letting $\eta_{t,ij} = \boldsymbol{\gamma}_t \cdot \boldsymbol{x}_{(ij)}$, the second and the third terms of r.h.s. of inequality (6) are given as follows.

$$h'(\eta_{t,i_tj_t})(\eta_{t+1,i_tj_t} - \eta_{t,i_tj_t}) + \frac{\sum_{i<j}(t_{ij}e^{-\eta_{t,ij}} + t_{ji}e^{\eta_{t,ij}})(h'(\eta_{t,ij}))^2(\eta_{t+1,ij} - \eta_{t,ij})^2}{2\max\{e^2, 8\ln \widetilde{K}\}}.$$

By replacing $\boldsymbol{\eta}_{t+1}$ with a variable $\boldsymbol{\eta}$, the above terms are further lower bounded by

$$\min_{\boldsymbol{\eta}} \left(h'(\eta_{i_t,j_t})(\eta_{i_t,j_t} - \eta_{t,i_t,j_t}) + \frac{\sum_{i,j}(t_{ij}e^{-\eta_{t,ij}} + t_{ji}e^{\eta_{t,ij}})(h'(\eta_{t,ij}))^2(\eta_{ij} - \eta_{t,ij})^2}{2\max\{e^2, 8\ln \widetilde{K}\}} \right).$$

This is equivalent to

$$-\frac{\max\{e^2, 8\ln \widetilde{K}\}}{t_{ij}e^{-\eta_{t,i_tj_t}} + t_{ji}e^{\eta_{t,i_tj_t}}}.$$

Note that this term is minimized if the denominator is minimized as $2\sqrt{t_{ij}t_{ji}}$, which is achieved when $\eta_{t,i_tj_t} = (1/2)\ln(t_{ij}/t_{ji})$. In addition, $\sqrt{t_{ij}t_{ji}}$ is minimized if t_{ij} and t_{ji} are biased, i.e., $t_{ij} = t_{\{i,j\}} + \lambda$ and $t_{ji} = \lambda$ (or vice versa). This implies that the l.h.s. of inequality (5) is lower bounded as

$$g_t(\boldsymbol{\gamma}_t) - \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}}\lambda}.$$

■

We can now prove our second main result.

Theorem 4 *For Bradley-Terry models, the regret of FTRL is*

$$O(n(\ln \widetilde{K})\sqrt{T/\lambda} + \lambda n^2 \ln K).$$

Furthermore, the regret bound becomes $O(n^{\frac{3}{2}}(\ln \widetilde{K} + \ln K)T^{1/3})$ *if* $\lambda = T^{\frac{1}{3}}/\sqrt{n}$.

Proof By Lemma 2 and Lemma 3, summing up the inequality for $t = 1, \ldots, T$, we have

$$\Phi_T(\boldsymbol{\gamma}_{T+1}) - \Phi_0(\boldsymbol{\gamma}_1) = \sum_{t=1}^{T} (\Phi_t(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t)) \geq \sum_{t=1}^{T} g_t(\boldsymbol{\gamma}_t) - \sum_{t=1}^{T} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}}\lambda}$$

By rearranging and letting $\gamma^* = \arg\min_{\boldsymbol{\gamma}} \sum_{t=1}^{T} g_t(\boldsymbol{\gamma})$,

$$\begin{aligned}\sum_{t=1}^{T} g_t(\boldsymbol{\gamma}_t) &\leq \Phi_T(\boldsymbol{\gamma}_{T+1}) - \Phi_0(\boldsymbol{\gamma}_1) + \sum_{t=1}^{T} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}}\lambda} \\ &\leq \sum_{t=1}^{T} g_t(\boldsymbol{\gamma}^*) + \lambda \sum_{i,j} g_{ij}(\boldsymbol{\gamma}^*) + \sum_{t=1}^{T} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}}\lambda}\end{aligned}$$

where the last inequality holds by the definition of $\boldsymbol{\gamma}_{T+1}$.

The second term is $O(\lambda n^2(\ln K))$. The third term is bounded as follows: Let $T_{\{ij\}}$ be the number of matches between team i and j over T trials. Then, we obtain

$$\sum_{t=1}^{T} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{t_{\{i_t,j_t\}}\lambda}} \leq \sum_{i<j} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{\lambda}} \int_0^{T_{\{ij\}}} \frac{1}{\sqrt{t}} = \sum_{i<j} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{\lambda}} \sqrt{T_{\{i,j\}}}$$
$$\leq \binom{n}{2} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{\lambda}} \sqrt{\frac{T}{\binom{n}{2}}} = \sqrt{\binom{n}{2}} \frac{\max\{e^2, 8\ln \widetilde{K}\}}{2\sqrt{\lambda}} \sqrt{T},$$

where the last inequality follows since the sum of the square roots is maximized when all components are equal. ■

We obtain our final result by Theorem 3 and 4 .

Corollary 1 *The regret of FTRL is*

$$O\left(\min\left\{n\widetilde{K}\ln T, n(\ln \widetilde{K})\sqrt{T/\lambda}\right\} + \lambda n^2(\ln K)\right).$$

5. Discussion

In this section, we consider some issues related to our analyses.

FTRL vs FTL and FTAL It is natual to ask if the Follow the Leader (FTL, e.g., Shalev-Shwartz (2011)) has a similar regret bound for Bradley-Terry models. FTL simply predicts the parameter $\gamma_{t+1} = \arg\min_\gamma \sum_{\tau=1}^{t} g_t(\gamma)$ at each trial t. In other words, the question is whether the virtual match regularizer is really necessary or not. In fact, we can show that the regret bound of FTL is $O(n\hat{K}\ln T)$, where $\hat{K} = \max_{t=1}^{T} \max_{i,j} e^{\gamma_{t,i}-\gamma_{t,j}}$, by following the analysis in Theorem 3. However, in early trials, Assumption 1 does not hold; thus, $\hat{K} = \infty$ in general. Therefore the regret bound is not meaningful. A similar argument is true for Follow The Approximate Leader (FTAL, Hazan et al. (2007)). At each trial t, FTAL predicts the parameter that optimizes the cumulative second order approximations of loss functions. The reget bound of FTAL can be shown to be $O(nK\ln T))$ by following the analysis of Theorem 3. However, similar to ONS, FTAL requires the prior knowledge of K.

Virtual Match vs the 2-norm Regularization It is possible to employ the standard square norm regularization, i.e., a version of FTRL that predicts $\gamma_t = \arg\min_{\gamma\in\Gamma} \sum_{\tau=1}^{t-1} g_t(\gamma) + \lambda\|\gamma\|_2^2$ can be employed to obtain the same regret bound $O(n^{\frac{1}{2}}(\ln K)\sqrt{T})$ of OGD. However, the prior knowledge of K is required to obtain this bound. Since the objective function of this FTRL is λ-strongly convex, it can be shown that $\hat{K} = e^{O(T/\lambda)}$ using the analysis appeared in, e.g., Boyd and Vandenberghe (2004). The bound of the radius $\hat{K}$ is exponentially worse than the virtual match regularization.

Lower Bound of the Regret To date, we have not obtained any lower bounds of the regret for Bradley-Terry models. For Bradley-Terry models with $n = 2$, lower bounds are no greater than $O(\ln T)$, since the $O(\ln T)$ regret bound is shown by McMahan and Streeter (2012) as discussed in Section 2.3. In general, few $\Omega(\ln T)$ lower bound has been proven for online density estimation problems except the Bernoulli model (Shtar'kov, 1987) and the Gaussian model (Takimoto and Warmuth, 2000) via minimax regret analyses.

Bayesian Approach For online density estimation problems, Bayesian approaches have been quite effective for obtaining $O(\ln T)$ regret bounds (e.g., Freund (1996); Azoury and Warmuth (2001); Kotlowski et al. (2010)). 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 obtain 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 2λ matches between each two players. It is an interesting open question whether a Bayesian approach achieves $O(\ln T)$ regret bound for these models.

6. Conclusion

We considered the online density estimation problem for Bradley-Terry models. We derived matching upper and lower bounds of the radius, analyzed FTRL with virtual match regularization, and showed better regret bounds than standard algorithms.

There are some interesting open questions to explore. An obvious open problem is to obtain better upper/lower regret bounds for Bradley-Terry models. It might be possible to design Bayesian-based algorithms for Bradley-Terry models as well. In addition, minimax analyses for Bradley-Terry models would be an interesting approach.

Acknowledgments

We thank anonymous reviewers and Hiroe Inoue for many helpful comments. Hatano is grateful to the supports from JSPS KAKENHI Grant Number 25330261 and CORE Project Grant of Microsoft Research Asia. Takimoto is grateful to the supports from JSPS KAKENHI Grant Number 15H02667. In addition, both authors acknowledge the support from MEXT KAKENHI Grant Number 24106010 (the ELC project).

References

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

Francis Bach. Self-concordant analysis for logistic regression. *Electric Journal of Statisitics*, 4: 384–414, 2010.

Francis Bach. Adaptivity of averaged stochastic gradient descent to local strong convexity for logistic regression. *Journal of Machine Learning Research*, 15:595627, 2014.

Stephan Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Lester R. Ford, Jr. Solution of a ranking problem from binary comparisons. *American Mathematical Monthly*, 64(8):28—-33, 1957.

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.

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.

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

Elad Hazan, Tomer Koren, and Kfir Levy. Logistic Regression: Tight Bounds for Stochastic and Online Optimization. In *Proceedings of the 27th Conference on Learning Theory (COLT'14)*, pages 197–209, 2014.

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

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.

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

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.

Shai Shalev-Shwartz. Online Learning and Online Convex Optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.

Yu. M. Shtar'kov. Universal Sequential Coding of Single Messages. *Problems of Information Transmission*, 23(3):175–186, 1987.

Eiji Takimoto and Manfred K. Warmuth. The Minimax Strategy for Gaussian Density Estimation. In *Proceedings of the Thirteenth Annual Conference on Computational Learning Theory (COLT 2000)*, pages 100–106, 2000.

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.

Appendix

Proof of Theorem 1

Proof The log likelihood function $L(\boldsymbol{\theta})$ is given as

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n-1} a \ln(1 + \theta_{i+1}/\theta_i) + \ln(1 + \theta_1/\theta_n).$$

Now, we let $\tilde{\theta}_1 = \theta_1/\theta_1 = 1$, $\tilde{\theta}_i = \theta_i/\theta_{i-1}$ for $i = 2, \ldots, n$. Then,

$$L(\tilde{\boldsymbol{\theta}}) = \sum_{i=2}^{n} a \ln(1 + \tilde{\theta}_i) + \ln\left(1 + 1/\prod_{i=2}^{n} \tilde{\theta}_i\right).$$

The offline optimizer $\boldsymbol{\theta}^*$ must satisfy, for $i = 2, \ldots, n$

$$\frac{\partial(L(\tilde{\boldsymbol{\theta}}^*))}{\partial\tilde{\theta}_i} = \frac{a}{1+\tilde{\theta}_i^*} + \frac{\prod_{j=2}^n \tilde{\theta}_j^*}{1+\prod_{j=2}^n \tilde{\theta}_j^*}\frac{1}{\tilde{\theta}_i^*} - \frac{1}{\tilde{\theta}_i^*} = 0.$$

Equivalently,

$$\frac{\tilde{\theta}_i^*}{1+\tilde{\theta}_i^*} = \frac{1}{a\left(1+\prod_{i=2}^n \tilde{\theta}_i\right)},$$

implying that $\tilde{\theta}_2^* = \tilde{\theta}_3^* = \cdots = \tilde{\theta}_n^*$. Now we let $x = \tilde{\theta}_2^*$. Then, it holds that

$$\frac{x}{1+x} = \frac{1}{a(1+x^{n-1})}.$$

By rearranging, $ax^n + (a-1)x - 1 = 0$. Since $(a-1)x - 1 = -ax^n < 0$, $x < 1/(a-1)$. Now,

$$K = \max_{i,j\in[n]} \theta_i^*/\theta_j^* \geq \theta_1^*/\theta_n^* = \frac{1}{\prod_{i=2}^n \tilde{\theta}_i} = \left(\frac{1}{x}\right)^{n-1} > (a-1)^{n-1} = \left(\frac{T-1}{n-1} - 1\right)^{n-1}.$$

■

Proof of Proposition 1

Proof Let $q(t) = g(\boldsymbol{w} + t\boldsymbol{v}) = \ln(1 + e^{w_j - w_i} + t(v_j - v_i))$. Then the first, second and third order derivatives are given as follows:

$$q'(t) = \frac{(v_j - v_i)e^{w_j - w_i + t(v_j - v_i)}}{1 + e^{w_j - w_i + t(v_j - v_i)}}.$$

$$\begin{aligned} q''(t) &= \frac{(v_j - v_i)^2 e^{w_j - w_i + t(v_j - v_i)}(1 + e^{w_j - w_i + t(v_j - v_i)})}{(1 + e^{w_j - w_i + t(v_j - v_i)})^2} \\ &\quad - \frac{(v_j - v_i)^2 e^{w_j - w_i + t(v_j - v_i)} e^{w_j - w_i + t(v_j - v_i)}}{(1 + e^{w_j - w_i + t(v_j - v_i)})^2} \\ &= \frac{(v_j - v_i)^2 e^{w_j - w_i + t(v_j - v_i)}}{(1 + e^{w_j - w_i + t(v_j - v_i)})^2} \end{aligned}$$

$$\begin{aligned} q'''(t) &= \frac{(v_j - v_i)^3 e^{w_j - w_i + t(v_j - v_i)}(1 + e^{w_j - w_i + t(v_j - v_i)})^2}{(1 + e^{w_j - w_i + t(v_j - v_i)})^4} \\ &\quad - \frac{2(v_j - v_i)^3 e^{w_j - w_i + t(v_j - v_i)}(1 + e^{w_j - w_i + t(v_j - v_i)})}{(1 + e^{w_j - w_i + t(v_j - v_i)})^2} \\ &= \frac{(v_j - v_i)^3 e^{w_j - w_i + t(v_j - v_i)}(e^{w_j - w_i + t(v_j - v_i)})}{(1 + e^{w_j - w_i + t(v_j - v_i)})^3}. \end{aligned}$$

So, we have

$$
\begin{aligned}
|q'''(t)| &= |v_j - v_i| \left| \frac{e^{w_j - w_i + t(v_j - v_i)} - 1}{1 + e^{w_j - w_i + t(v_j - v_i)}} \right| |q''(t)| \\
&\le |v_j - v_i||q''(t)| \quad (\text{since } |(x-1)/(x+1)| \le 1 \text{ for } x \ge 0) \\
&\le 2\|\boldsymbol{v}\|_\infty |q''(t)|.
\end{aligned}
$$

■

Proof of Proposition 2

Proof For simplicity, we omit the subscript t of g_t, i.e., $g_t = g$. We use the following result of Bach (2010):

Proposition 3 (Bach (2010)) *Let $q : \mathbb{R} \to \mathbb{R}$ be any three times differentiable function such that $|q'''(t)| \le Sq''(t)$ for some $S > 0$. Then, for any $t \ge 0$,*

$$
q(t) \ge q(0) + q'(0)t + \frac{q''(0)}{S^2}(e^{-St} + St - 1).
$$

In addition, we use the following technical proposition.

Proposition 4 *For $x \ge 0$, $\frac{x^2}{e^{-x}+x-1} \le \max\{2e^2, 8x\}$.*

The proof is given in the next subsection. Using Proposition 4 and Proposition 3, we obtain that

$$
q(t) \ge q(0) + q'(0)t + \frac{q''(0)}{\max\{2e^2, 8St\}} t^2 \tag{7}
$$

By applying inequality (7) for $q(t) = g(\boldsymbol{\gamma}' + t(\boldsymbol{\gamma} - \boldsymbol{\gamma}'))$, $t = 1$ and $S = R\|\boldsymbol{\gamma} - \boldsymbol{\gamma}'\|_\infty$, we obtain

$$
g(\boldsymbol{\gamma}) \ge g(\boldsymbol{\gamma}') + \nabla g(\boldsymbol{\gamma}') \cdot (\boldsymbol{\gamma} - \boldsymbol{\gamma}') + \frac{(\boldsymbol{\gamma} - \boldsymbol{\gamma}')^\top \nabla^2 g(\boldsymbol{\gamma}')(\boldsymbol{\gamma} - \boldsymbol{\gamma}')}{\max\{2e^2, 8R\|\boldsymbol{\gamma} - \boldsymbol{\gamma}'\|_\infty\}},
$$

as claimed.

■

6.1. Proof of Proposition 4

Proof Let

$$
f(x) = \frac{x^2}{e^{-x} + x - 1}.
$$

Let $a > 1$ be a positive real number. We consider two cases. (i) Suppose that $x \le a$. Then, by second order Taylor approximation around $x = 0$, $e^{-x} + x - 1 \ge e^{-a}x^2/2$,

$$
\frac{x^2}{e^{-x} + x - 1} \le \frac{x^2}{e^{-a}x^2/2} = 2e^a.
$$

(ii) Otherwise, assume that $x > a > 1$. Then by using Taylor expansion of f around $x = a$, $f(x) = f(a) + f'(z)(x-a)$ for some z s.t. $a < z < x$. Note that

$$f'(z) = \frac{-z^2(-e^{-z}+1) + 2z(e^{-z}+z-1)}{(e^{-z}+z-1)^2} = \frac{z^2e^{-z}+z^2+2ze^{-z}-2z}{(e^{-z}+z-1)^2} < \frac{2z^2}{(z-1)^2}.$$

Furthermore, the last term is maximized when $z = a$. Thus,

$$f(x) = f(a) + f'(z)(x-z) < \frac{a^2}{e^{-a}+a-1} + \frac{2^a}{(a-1)^2}x - \frac{a^3}{(a-1)^2} < \frac{2^a}{(a-1)^2}x.$$

Therefore, in both cases, $f(x)$ is bounded by $\max\{2e^a, 2a^2/(a-1)^2x\}$. Finally, by plugging $a = 2$ into the upper bound, we complete the proof. ■

6.2. Proof of Theorem 3

Proof By using second order Taylor expansion of $g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t)$ around $\boldsymbol{\gamma}_t$,

$$g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t) = g_t(\boldsymbol{\gamma}_t) + \nabla g_t(\boldsymbol{\gamma}_t) \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) + (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t)^\top \nabla^2 \Phi_t(\boldsymbol{z}_t)(\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t),$$

where $\boldsymbol{z}_t$ is a convex combination of $\boldsymbol{\gamma}_t$ and $\boldsymbol{\gamma}_{t+1}$. Then, since $h''(x) = e^x/(1+e^x)^2 = 1/(e^x+1+e^{-x})$, $h''(\boldsymbol{z}_t \cdot \boldsymbol{x}_\tau) \geq \frac{1}{2\tilde{K}+1}$. Similarly, $h'(x) = 1/(1+e^x)$. So, $h'(\boldsymbol{\gamma}_t \cdot \boldsymbol{x}_t) \geq 1/(1+\tilde{K}) \geq 1/(2\tilde{K}+1)$. Therefore,

$$g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t) \geq g_t(\boldsymbol{\gamma}_t) + \frac{1}{2\tilde{K}+1}(\boldsymbol{x}_t \cdot (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) + (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t)^\top C_t (\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t) - \varepsilon \|\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t\|^2),$$

where $C_t = \sum_{\tau=1}^t \boldsymbol{x}_\tau \boldsymbol{x}_\tau^\top + \lambda \sum_{i,j} \boldsymbol{x}_{ij} \boldsymbol{x}_{ij}^\top + \varepsilon I_n$ The second term of the r.h.s. is further lower bounded by

$$\frac{1}{2\tilde{K}+1}\left(\boldsymbol{x}_t \cdot (\boldsymbol{\gamma} - \boldsymbol{\gamma}_t) + (\boldsymbol{\gamma} - \boldsymbol{\gamma}_t)^\top C_t (\boldsymbol{\gamma} - \boldsymbol{\gamma}_t)\right),$$

which is minimized as

$$-(4\tilde{K}+2)\boldsymbol{x}_t^\top C_t^{-1} \boldsymbol{x}_t.$$

In addition, $-\varepsilon \|\boldsymbol{\gamma}_{t+1} - \boldsymbol{\gamma}_t\|^2$ can be lower bounded by $-2\varepsilon n(\ln \widetilde{K})^2$. Therefore we have

$$g_t(\boldsymbol{\gamma}_{t+1}) + \Delta_{\Phi_{t-1}}(\boldsymbol{\gamma}_{t+1}, \boldsymbol{\gamma}_t) \geq g_t(\boldsymbol{\gamma}_t) - (4\tilde{K}+2)\boldsymbol{x}_t^\top C_t^{-1} \boldsymbol{x}_t - \frac{2\varepsilon n(\ln \widetilde{K})^2}{2\widetilde{K}+1}. \tag{8}$$

Then, we use the following technical lemma.

Lemma 4 (Hazan et al. (2007)) *Let $\boldsymbol{u}_t \in \mathbb{R}^n$ for $t = 1, \ldots, T$, be a sequence of vectors s.t. for some $r > 0$, $\|\boldsymbol{u}_t\| \leq r$ and let $V_t = \sum_{\tau=1}^t \boldsymbol{u}_\tau \boldsymbol{u}_\tau^\top + \varepsilon I_n$. Then,*

$$\sum_{t=1}^T \boldsymbol{u}_t^\top V_t^{-1} \boldsymbol{u}_t \leq n \ln\left(r^2 T/\varepsilon + 1\right).$$

By Lemma 2 inequality (8), summing the inequality for $t = 1, \dots, T$, we have

$$\Phi_T(\boldsymbol{\gamma}_{T+1}) - \Phi_0(\boldsymbol{\gamma}_1) = \sum_{t=1}^{T} (\Phi_t(\boldsymbol{\gamma}_{t+1}) - \Phi_{t-1}(\boldsymbol{\gamma}_t))$$
$$\geq \sum_{t=1}^{T} g_t(\boldsymbol{\gamma}_t) - (4\tilde{K} + 2) \sum_{t=1}^{T} \boldsymbol{x}_t^\top C_t^{-1} \boldsymbol{x}_t - 2\frac{\varepsilon T n (\ln \widehat{K})^2}{2\widetilde{K} + 1}.$$

By rearranging and letting $\gamma^* = \arg\min_{\boldsymbol{\gamma}} \sum_{t=1}^{T} g_t(\boldsymbol{\gamma})$,

$$\sum_{t=1}^{T} g_t(\boldsymbol{\gamma}_t) \leq \Phi_T(\boldsymbol{\gamma}_{T+1}) - \Phi_0(\boldsymbol{\gamma}_1) + (4\widetilde{K} + 2) \sum_{t=1}^{T} \boldsymbol{x}_t^\top C_t^{-1} \boldsymbol{x}_t + \frac{2\varepsilon T n (\ln \widetilde{K})^2}{2\widetilde{K} + 1}$$
$$\leq \sum_{t=1}^{T} g_t(\boldsymbol{\gamma}^*) + \lambda \sum_{i,j} g_{ij}(\boldsymbol{\gamma}^*) + (4\widetilde{K} + 2) n \ln \left(\frac{4T}{\varepsilon} + 1 \right) + \frac{2\varepsilon T n (\ln \widetilde{K})^2}{2\widetilde{K} + 1},$$

where the last inequality holds by definition of $\boldsymbol{\gamma}_{T+1}$. Note that the second term of the r.h.s. of the above inequality is $O(\lambda n^2 (\ln K))$. Finally, by letting $\varepsilon = 1/T$, we complete the proof. ■