Efficient Estimation of Distribution-free dynamics in the Bradley-Terry Model

Heejong Bong*, **Wanshan Li***, **Shamindra Shrotriya***
Department of Statistics and Data Science
Carnegie Mellon University
Pittsburgh, PA 15213
{hbong,wanshanl,sshrotri}@andrew.cmu.edu

Abstract

We propose a time-varying convex generalization of the original Bradley-Terry model. Our model directly captures the temporal dependence structure of the pairwise comparison data without explicit temporal distribution assumptions. This enables the modeling of discrete-time dynamic global rankings of N distinct objects. Different choices of the convex penalization term provide a control on the degree of smoothing in the derived time-varying global rankings of the N distinct objects. Furthermore this directly enables analysis on sparse time-varying pairwise comparison data. We also prove that a relatively weak condition is both necessary and sufficient to guarantee the existence and uniqueness of the solution of our model. We implement various convex optimization algorithms to efficiently estimate the model parameters under the ℓ_2^2, ℓ_2, and ℓ_1 convex penalization norms. We conclude by thoroughly testing the practical effectiveness of our model under both simulated and real world settings, including ranking 5 seasons of National Football League (NFL) team data. Our generalized time-varying Bradley-Terry model thus provides a useful minimalist benchmarking tool for other feature-rich dynamic ranking models since it only relies on the time-varying pairwise comparison data between the N distinct objects.

1 Introduction and Prior Work

1.1 Pairwise Comparison Data and the Bradley-Terry Model

Pairwise comparison data is very common in daily life especially in cases where the goal is to rank several objects. Rather than directly ranking all objects simultaneously it is usually much easier and more efficient to first obtain results of pairwise comparisons. The pairwise comparisons can then be used to derive a *global* ranking across all individuals in a principled manner. One such statistical model for deriving global rankings using pairwise comparisons was presented by statisticians R. A. Bradley and M. E. Terry in their classic 1952 paper [3], and thereafter commonly referred to as the Bradley-Terry model in statistical literature. A similar model was also studied by Zermelo dating back to 1929 [19]. The Bradley-Terry model is one of the most popular models to analyze paired comparison data due to its interpretable setup and computational efficiency in parameter estimation. The Bradley-Terry model along with its various generalizations has been studied and applied in various ranking applications across many broad domains. This includes the ranking of sports teams ([11],[5],[7]), scientific journals ([14],[15]), and the quality of several brands ([1],[13]).

*Equal contribution. This version is as submitted for CMSAC 2020 review with minor grammatical corrections and de-anonymization. All code to reproduce analysis in this paper can be accessed from: https://github.com/shamindras/bradley-terry-convexopt

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

In order to describe the original Bradley-Terry model [3], suppose that we have N distinct objects, each with a (positive) score or index $(s_i)_{i\in[N]}$ showing their power of competing with each other at a single point in time (static). This model assumes that the comparisons between different pairs are independent and the results of comparisons between a given pair, object i and object j, are independent and identically distributed as Bernoulli random variables, with success probability defined as

$$\mathbb{P}(i \text{ beats } j) = \frac{s_i}{s_i + s_j} \tag{1}$$

A common way to parameterize $(s_i)_{i\in[N]}$ is to assume that $s_i = \exp(\beta_i)$, for all $i \in [N]$. In this case, equation (1) is usually expressed as $\mathsf{logit}(\mathbb{P}(i \text{ beats } j)) = \beta_i - \beta_j$, where $\mathsf{logit}(x) := \log \frac{x}{1-x}$. An important assumption in the original Bradley-Terry model is that comparisons of different pairs are independent. However, in practice such strong data independence assumptions are unlikely to hold, as discussed further in [2],[6],[4],[15]. A typical way to deal with such data dependence in the pairwise comparison scores $(s_i)_{i\in[N]}$ is to use quasi-likelihood approaches [17],[15]. Fortunately under the setting of the original Bradley-Terry model [3] the log-quasi-likelihood and usual log-likelihood are the same, in terms of optimizing for the target parameters $\boldsymbol{\beta} = (\beta_1, \cdots, \beta_N)$ as discussed further in [17].

1.2 The Time-varying (dynamic) Bradley-Terry Model

In many applications it is very common to observe paired comparison data spanning over multiple (discrete) time periods. A natural question of interest is then to understand how the *global* rankings *change* over time. For example in sports analytics the performance of teams often changes from season to season and thus explicitly incorporating the time-varying dependence into the model is crucial. In particular the paper [7] considers a state-space generalization of the Bradley-Terry model to modelling the sports tournaments data. In a similar manner bayesian frameworks for the dynamic Bradley-Terry model are studied further in [9]. Such dynamic analysis of paired comparison data is becoming increasingly important because of the rapid growth of openly available time-dependent paired comparison data.

Our main focus in this paper is to tackle the problem of efficiently estimating the parameters in the time-varying Bradley-Terry model under a frequentist framework. Our frequentist approach allows us to assume a *distribution-free* approach in the *changes in model parameters* over time in order to perform a dynamic parameter estimation with minimal assumptions. Our proposed model relies only on the time-varying pairwise comparison data across the N distinct objects. This is in contrast to more assumption-heavy recent frequentist dynamic Bradley-Terry models including [5]. The remainder of this paper discusses our model in detail and is organized as follows. First, we formulate our overall time-varying Bradley-Terry estimation requirement as a convex optimization problem under various penalty norms. We then describe the relatively weak necessary and sufficient conditions to guarantee uniqueness of our model solution. We proceed to apply various well known convex optimization algorithms to derive efficient estimation of time-varying parameters with provable computational convergence guarantees, exploiting specific structure of the convex penalty terms as applicable. We then provide two strategies to optimizing our penalization parameter, λ, using both data-driven and heuristic approaches. Finally we conclude by applying our model on synthetic datasets and a real-world example including NFL data demonstrating the practical viability of the model as minimalist benchmarking tool for time-varying rankings.

2 Our proposed Time-varying Bradley-Terry Model

In our time-varying frequentist setup of the Bradley-Terry model, we generalize the approach taken in the original Bradley-Terry paper [3] in estimating parameters of interest via maximizing the likelihood (or mimimizing negative likelihood) over the discrete observed time points $\{1, 2, \ldots, T\} =: [T]$. The parameters of interest, $\boldsymbol{\beta}^{(t)} \in \mathbb{R}^N$, are now given for each time point $t \in [T]$, for each of the N distinct objects. We assume that the pairwise comparison between a given pair, object i and object j, at a given time point t is determined by the corresponding parameter $\boldsymbol{\beta}^{(t)}$ so that

$$\mathsf{logit}(\mathbb{P}(i \text{ beats } j \text{ at time } t)) = \boldsymbol{\beta}_i^{(t)} - \boldsymbol{\beta}_j^{(t)}. \tag{2}$$

Given a $\boldsymbol{\beta}^{(t)}$, by equation (2) we can derive the log-likelihood $\ell_t(\boldsymbol{\beta}^{(t)})$ at each time $t \in [T]$. As the size of data in each t is much smaller than the global data, we might want to *smooth* the parameter $\boldsymbol{\beta}^{(t)}$ across different time points and to leverage the dynamic structure in the global data. Our proposal is to include the penalization term $\lambda \sum_{t=1}^{T-1} h(\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)})$ for some convex penalty function h that can effectively smooth the estimation of parameter by penalizing large differences in subsequent time points. In summary, our proposed time-varying setup of the Bradley-Terry model can be framed as the following convex optimization problem:

$$\min_{\{\boldsymbol{\beta}^{(t)}\}_{t\in[T]}} -\sum_{t=1}^{T} \ell_t(\boldsymbol{\beta}^{(t)}) + \lambda \sum_{t=1}^{T-1} h(\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}), \text{ s.t. } \sum_{i=1}^{N} \beta_i^{(1)} = 0 \tag{3}$$

where $\lambda \geq 0$, $\boldsymbol{\beta}^{(t)} \in \mathbb{R}^N$, and $h : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}_{\geq 0}$ is the convex penalty function.

First, we observe that if $T = 1$ and $\lambda = 0$ then equation (3) reduces to the same objective as the original (static) Bradley-Terry model [3]. As such our model represents a generalization of the original Bradley-Terry model to the time-varying setting. Furthermore we include an additional constraint, namely $\sum_{i=1}^{N} \beta_i^{(1)} = 0$ which sets the sum of the estimated parameters to 0 at the initial time point $t = 1$. This artificial constraint is used to guarantee the existence and uniqueness of the global solution set in the convex formulation discussed further in section 3. Since we are only concerned with obtaining *relative* global object rankings this constraint still maintains our ultimate goal and further ensures that equation (3) is invariant to constant shifts in the model parameters. In our project we consider 3 cases for the convex penalty function h, namely $h = \|\cdot\|_1$, $h = \|\cdot\|_2$, and $h = \|\cdot\|_2^2$ to allow for user flexibility in the degree of smoothing in the output global rankings depending on the underlying sparsity of the available pairwise comparison data.

3 Existence and uniqueness of solution

The existence or the uniqueness of solutions for the model (3) is not guaranteed in general. This is an innate property of the original Bradley-Terry model [3]. As pointed out by Ford Jr. [8] the Bradley-Terry model requires a sufficient amount of pairwise comparisons so that there is enough information of relative performance between any pair of two entries for parameter estimation purposes. For example, if there is an object which has never been compared to the others, there is no information which the model can exploit to assign a score for the object, so its derived rank could be arbitrary. In addition if there are several entries which have never outperformed the others then Bradley-Terry model would assign negative infinity for the performance of these entries, and we would not be able to compare among them for global ranking purposes.

In 1957 Ford Jr. [8] discovered the necessary and sufficient conditions which guarantee the uniqueness of the solution in the original Bradley-Terry model. The two equivalent conditions are:

Condition (1)**.** In every possible partition of the objects into two nonempty subsets, some object in the second set has been preferred at least once to some object in the first set

Condition (2)**.** For each ordered pair (i, j), there exists a sequence of indices $i_0 = i, i_1, \ldots, i_n = j$ such that

$$x_{i_{k-1},i_k} > 0 \tag{4}$$

for $k = 1, \ldots, n$.

We prove that these conditions can also be adapted to guarantee the uniqueness of the solution in our time-varying Bradley-Terry model. This is summarized in the following theorem.

Theorem 3.1. *Given data $\{x_{i,j}^{(t)}\}_{i,j,t}$ satisfies the previous condition, let $x_{i,j} = \sum_{t=1}^{T} x_{i,j}^{(t)}$. If $\{x_{i,j}\}$ satisfies the condition above, then*

1. *If h is continuous and $h(x) \to \infty$ as $x \to \infty$, the solution $\boldsymbol{\beta}^\star$ for* (2) *is attainable in $\mathbb{R}^{N\times T}$; and*
2. *With squared-ℓ_2 penalty, $h = \|\cdot\|_2^2$, the uniqueness of solution classes for* (3)*.*

Hence, in our proposed time-varying Bradley-Terry model we do not require the strong conditions [8] to hold at each time point, but simply require the *aggregated conditions* in Theorem 3.1 to hold overall. This is a quite weak data requirement in the sense that it is satisfied even when each object does not have pairwise comparisons with all other objects at every single time point. For example, even if one player did not play any game in a season, as long as we have a game record in another season (with at least one win and one loss respectively), we can assign a rank to this player in the missed season. This minimal data requirement for our proposed model is in our view a key advantage to other frequentist time-varying approaches which could be alternatively considered here. A typical such alternative frequentist time-varying approach would be to fit a separate Bradley-Terry model at each discrete time point and then 'smooth' the derived global rankings post-hoc. This could be done using kernel smoothing techniques over the T time points. However, such an approach would require the stronger conditions in [8] to hold at *every time point* rather than the much weaker aggregated conditions required for our model to hold in order for a unique solution. In this sense our model requires not only minimal assumptions on the time-varying dependence but also on the pairwise data information requirements between the N distinct objects.

4 Optimization Methods for Model Estimation

Since we have formulated our model as a convex optimization problem per equation (3), we discuss well known algorithms which can be used to efficiently estimate our time-varying model parameters $\boldsymbol{\beta}^{(t)}$. More specifically we note that the objective function in equation (3) can be expressed as follows:

$$\underbrace{f(\boldsymbol{\beta})}_{\text{objective in } \boldsymbol{\beta}} = \underbrace{g(\boldsymbol{\beta})}_{\text{negative log-likelihood}} + \underbrace{H(\boldsymbol{\beta})}_{\text{convex penalty}} \tag{5}$$

where

$$g(\boldsymbol{\beta}) = -\sum_{i=1}^{T} l_t(\boldsymbol{\beta}^{(t)}), \text{ and } H(\boldsymbol{\beta}) = \lambda \sum_{t=1}^{T-1} h(\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}^{(t+1)}), \lambda \geq 0 \tag{6}$$

Since g is convex differentiable and H is convex for the convex penalty function $h : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}_{\geq 0}$ then the proximal gradient descent method (PGD) is a natural optimization technique for this problem when using the ℓ_2^2-norm for the smoothing penalty. In this case we can derive the closed form of the proximal operator and run PGD at an inexpensive computational cost. However in the case of other smoothing penalties such as using the ℓ_1-norm or ℓ_2-norm, the proximal operator for h does not have a closed form thereby reducing the computational feasibility of the PGD technique. In such cases we resort to the Alternating Direction Method of Multipliers (ADMM) technique. In our optimization implementations we consider both fixed descent step sizes s_i and also determined by the backtracking line search with the initial step size s_{init}. Below we separately detail the efficient closed form for the proximal operator with the ℓ_2^2-penalty and discuss the computational feasibility under more general ℓ_1-norm and ℓ_2-norm smoothing penalties.

4.1 Efficient PGD for ℓ_2^2-penalty

The proximal operator for the step size $s > 0$ and the smoothing penalty H is defined as follows:

$$\text{prox}_{s,H}(\boldsymbol{\beta}') = \arg\max_{\boldsymbol{\beta} \in \mathbb{R}^{N \times T}} \left[\frac{1}{2s} \|\boldsymbol{\beta} - \boldsymbol{\beta}'\|_2^2 + \lambda \sum_{t=1}^{T-1} \|\boldsymbol{\beta}^{(t+1)} - \boldsymbol{\beta}^{(t)}\|_2^2 \right] \tag{7}$$

Hence, we can decompose the optimization in the proximal operator into the marginal optimizations for $i = 1 \ldots N$ so that if $\boldsymbol{\beta}^+ = \text{prox}_{s,H}(\boldsymbol{\beta}')$, then

$$\boldsymbol{\beta}_i^+ = \arg\max_{\boldsymbol{\beta}_i \in \mathbb{R}^T} \left[\frac{1}{2s} \|\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\beta}_i^{'(t)}\|_2^2 + \lambda \sum_{t=1}^{T-1} (\boldsymbol{\beta}_i^{(t+1)} - \boldsymbol{\beta}_i^{(t)})^2 \right]. \tag{8}$$

Hence, $\nabla_{\boldsymbol{\beta}_i} \mathcal{F}_i = 0$ if and only if

$$\begin{bmatrix} 1+2s\lambda & -2s\lambda & & \\ -2s\lambda & 1+4s\lambda & \ddots & \\ & \ddots & \ddots & -2s\lambda \\ & & -2s\lambda & 1+4s\lambda \end{bmatrix} \boldsymbol{\beta}_i^+ = \boldsymbol{\beta}_i'. \tag{9}$$

Hence by solving the linear system (9) for each $i = 1, \ldots, N$, we get the output of the proximal operator for the squared ℓ_2^2-norm penalty. By exploiting the specific sparse *tri-diagonal* structure of the linear system we can solve it in $O(n)$ computations which is much more efficient that just using matrix inversion.

Even though we can solve the proximal operator quite efficiently, we find that in practice the classical Newton's method can be even more efficient. However, when λ is close to 0, the numerical performance of Newton's method would become a little bit unstable, so we still need PGD in cases where λ is very small. For moderate λ we recommend Newton's method.

4.2 ADMM for ℓ_1-penalty and ℓ_2-penalty

On the other hand, getting the proximal operator for general smoothing penalty functions is nontrivial. In particular the l_1-norm and l_2-norm penalty terms do not have a closed-form proximal solution and require nontrivial amount of computation to approximate the optimum. In these two cases, we can use ADMM (Alternating Direction Method of Multipliers), as detailed below for our setting.

Let $\boldsymbol{\beta} := \mathrm{vec}(\boldsymbol{\beta}^{(1)}, \cdots, \boldsymbol{\beta}^{(T)}) \in \mathbb{R}^{TN}$ be the vector of scores of all individuals at all time points. Define $\boldsymbol{\theta} \in \mathbb{R}^{(T-1)N}$ by $\theta_{t\cdot N+i} = \mu_i^{t+2} - \mu_i^{t+1}$ for $i \in [N],\ 0 \leq t \leq T-2$. By introducing a matrix $A \in \mathbb{R}^{[(T-1)N]\times TN}$ we can express $\boldsymbol{\theta}$ as $\boldsymbol{\theta} = A\boldsymbol{\mu}$ and rewrite the optimization (3) as

$$\begin{aligned} &\text{minimize} \quad -\boldsymbol{\ell}_T(\boldsymbol{\beta}) + \lambda\tilde{h}(\boldsymbol{\theta}), \\ &\text{subject to } A\boldsymbol{\beta} = \boldsymbol{\theta}, \end{aligned} \tag{10}$$

where $\boldsymbol{\ell}_T(\boldsymbol{\beta}) = \sum_{t=1}^{T} \ell_t(\boldsymbol{\beta}^{(t)})$ and $\tilde{h} = \sum_{t=1}^{T-1} h(\boldsymbol{\theta}^{I_t})$ with the index set $I_t = \{i : N(t-1)+1 \leq i \leq Nt\}$. The ADMM scheme can be written as

$$\begin{aligned} \boldsymbol{\beta}^{k+1} &= \arg\min_{\boldsymbol{\beta}} -\boldsymbol{\ell}_T(\boldsymbol{\beta}) + (A\boldsymbol{\beta})^\top \boldsymbol{\mu}^k + \frac{\eta}{2}\|A\boldsymbol{\beta} - \boldsymbol{\theta}^k\|^2, \\ \boldsymbol{\theta}^{k+1} &= \arg\min_{\boldsymbol{\theta}} \lambda\tilde{h}(\boldsymbol{\theta}) - \boldsymbol{\theta}^\top \boldsymbol{\mu}^k + \frac{\eta}{2}\|A\boldsymbol{\beta}^{k+1} - \boldsymbol{\theta}\|^2, \\ \boldsymbol{\mu}^{k+1} &= \boldsymbol{\mu}^k + \eta(A\boldsymbol{\beta}^{k+1} - \boldsymbol{\theta}^{k+1}). \end{aligned} \tag{11}$$

The update of $\boldsymbol{\theta}$ is simply the proximal operator of $\tilde{h}$. The update of $\boldsymbol{\beta}$ is a convex optimization problem of a well-behaved function with no constraints and can be solved by some basic methods, like Newton's method.

5 Tuning λ (smoothing penalty parameter) in practical settings

In our convex formulation of the time-varying Bradley Terry model (3) we note that the penalty coefficient $\lambda \in \mathbb{R}_{\geq 0}$ is effectively a global smoothing parameter for the fitted $\boldsymbol{\beta}^{(t)}$ values between subsequent time periods under the various penalty norms. Increasing λ , all else held constant, thus increases the penalty on the difference between subsequent $\boldsymbol{\beta}^{(t)}$ values over a single time period. This leads to $\boldsymbol{\beta}^{(t)}$ values (and hence the derived global rankings) becoming 'smoothed' together across time.

Naturally the question remains on how to *tune* λ in practical applications in a principled manner. This is a fundamentally challenging question since our proposed model is unsupervised i.e. the true $\boldsymbol{\beta}^{(t)}$ or score $(s_i)_{i\in[N]}$ values are not directly observed. We note that in practice the end user of the time-varying Bradley Terry model typically seeks the global time-varying team rankings, rather than the individual fitted $\boldsymbol{\beta}^{(t)}$ coefficients used to derive pairwise comparisons. As such the usual approach of changing λ to fit the $\boldsymbol{\beta}^{(t)}$ values and then assessing the associated impact on the time-varying global rankings changes is a rather *indirect approach* to controlling the degree of smoothing in the sought after time-varying global rankings. We propose both a direct heuristic approch and a data-driven approach in tuning λ in real world settings. We detail the approaches and the conditions under which they would be useful below.

5.1 Heuristic Approach

We propose a simple heuristic whereby the user controls the degree of smoothing in global rankings *directly* by specifying a maximal ranking change parameter $\alpha \in [N]$ over all N teams and over all T time periods. By specifying this global (integer) parameter α we can search over a suitable finite grid of $\lambda \in \mathbb{R}^+$ values to meet this user specified global maximum team rank change requirement. In this heuristic we note that the user simply specifies a positive integer α indicating the maximum increase/decrease in global ranks for any team over all time periods. We claim that the user-specified α parameter is much more intuitive for the end user to directly control the global ranking changes. Here α is seen as a global smoothing parameter since controlling the maximum global rank change over all T periods is effectively controlling smoothing across all local consecutive time periods as well.

Furthermore, since α here is integer-valued it is naturally capped at the total number of teams N, since a team can't globally change rankings by more than the total number of teams across any time period. However in practice it will be much lower than this and easier to prescribe by the end user based on reasonable domain knowledge of expected time-varying ranking movements. We acknowledge that this heuristic trades the subjectivity of choosing λ for α but that it controls for the degree of smoothing in global rankings directly from the users point of view. The limited choice of α for the end user and the efficiency in grid-fitting λ makes it effective to apply in practical situations where the user has suitable domain knowledge on maximal global ranking changes.

5.2 Data Driven Approach - Sample Splitting and LOOCV (recommended)

Alternatively we propose a data-driven approach to tuning λ when sufficient domain knowledge is not available for the end-user. First, we note that in ranking objects from pairwise comparisons, we want the time-varying Bradley-Terry model to sort objects by means of win rates. That is, a higher globally-ranked object should be more likely to win against a lower globally-ranked object. Hence, we choose λ giving the best prediction on pair-wise *win rates*. In prediction, leave-one-out cross-validation (LOOCV) has been a provably successful model parameter tuning mechanism without the aid of human heuristics [10]. Given that we don't observe the ranking values we briefly describe how we adapt LOOCV to our unsupervised time-varying Bradley-Terry model. We then propose several techniques to reduce computational cost for our LOOCV λ-tuning approach.

In general settings where we have independent and identically distributed (i.i.d) samples, LOOCV assesses the performance of a predictive model by holding out one of the i.i.d samples. In our case, each pairwise comparison can be considered an i.i.d. sample if we take the compared objects and the time point on which they are compared as covariates. Let (t_k, i_k, j_k) denote k-th pairwise comparison where object i_k won against object j_k at time point t_k for $k = 1, \ldots, N$. Then, for a given smoothing penalty parameter λ, LOOCV is adapted to our model as follows:

1. For $k = 1, \ldots, N$, given λ:
   (a) Solve (3) with λ on the dataset where the k-th comparison is held-out.
   (b) Calculate the negative log-likelihood (nll) of the previous solution to (t_k, i_k, j_k).
2. Take the average of the negative log-likelihoods to obtain nll_λ as a loss in the predictive performance of time-varying Bradley-Terry model for given λ on our dataset
3. Choose the λ^* with the smallest nll_λ value overall

Solving (3) with every candidate λ, over a suitable finite grid of candidate values and on every hold-one-out dataset incurs heavy computational cost. To reduce this cost, we can approximate the exhaustive LOOCV with a stochastic estimate. The exhaustive LOOCV fits the model to every possible held-one-out data and takes their average. However, if we have a large number of pairwise comparisons, an average of much fewer losses makes tiny error. Hence, we can uniformly randomly sample a smaller number of matches to be held-out and obtain $\hat{\text{nll}}_\lambda$ efficiently. Also, we can quicken fitting on held-one-out data by giving the original estimate as an initial value. We assume that the parameter fitted on the entire dataset differs little to the fitted on held-one-out datasets. Hence, starting from the original estimate, optimizing functions could reach convergence much faster. These two speed-up techniques make LOOCV a practical and recommended option for time-varying Bradley-Terry models with a large number of parameter values, and where there is limited domain knowledge available.

6 Experiments

6.1 Synthetic Data Generation

We can conduct simulation experiments in which we know the underlying global ranks of all entries by generating data from a synthetic process. Given the number of entries N and the number of time points T, the overall synthetic data generation process is as follows:

1. For each $i \in [N]$, simulate $\boldsymbol{\beta}_i^* \in \mathbb{R}^T$ from a distribution of time-series.
2. For each $i \neq j$ and $t \in [T]$, generate $n_{ij}^{(t)}$ from some distribution and simulate $x_{ij}^{(t)}$ by $x_{ij}^{(t)} \sim \text{Binom}\Big(n_{ij}^{(t)}, 1/\{1 + \exp[\beta_j^{(t)} - \beta_i^{(t)}]\}\Big)$, $x_{ji}^{(t)} = n_{ij}^{(t)} - x_{ij}^{(t)}$.

For the choice of distribution of $\boldsymbol{\beta}_i^*$'s in step 1, we use Gaussian process to generate $\boldsymbol{\beta}^*$ since they are widely used in applications and can generate $\boldsymbol{\beta}^*$ that are *smoothed* path across time. For each $i \in [N]$, The generation of $\boldsymbol{\beta}_i^* \in \mathbb{R}^T$ from a Gaussian process $\text{GP}(\mu_i(\cdot), \sigma_i(\cdot, \cdot))$ includes three steps.

1. Set the values of the mean process $\mu_i(\cdot)$ at $t \in [T]$ and get mean vector $\boldsymbol{\mu}_i \in \mathbb{R}^T$
2. Set the values of the variance process $\sigma_i(\cdot, \cdot)$ at $(s, t) \in [T]^2$, to derive $\Sigma_i \in \mathbb{R}^{T \times T}$.
3. Generate a sample $\boldsymbol{\beta}_i^*$ from $\text{Normal}(\boldsymbol{\mu}_i, \Sigma_i)$.

6.2 Synthetic Simulation Results

First we generate the parameter $\boldsymbol{\beta}^*$ via a Gaussian process using the method described in 6.1 for $N = 10$ and $T = 10$ (see appendix for more details). We can then compare the true $\boldsymbol{\beta}^*$ and $\hat{\boldsymbol{\beta}}$ by different methods to see how effectively our formulation of time-varying Bradley-Terry Model fits the synthetic ground truth. The results are shown in Figure 1.

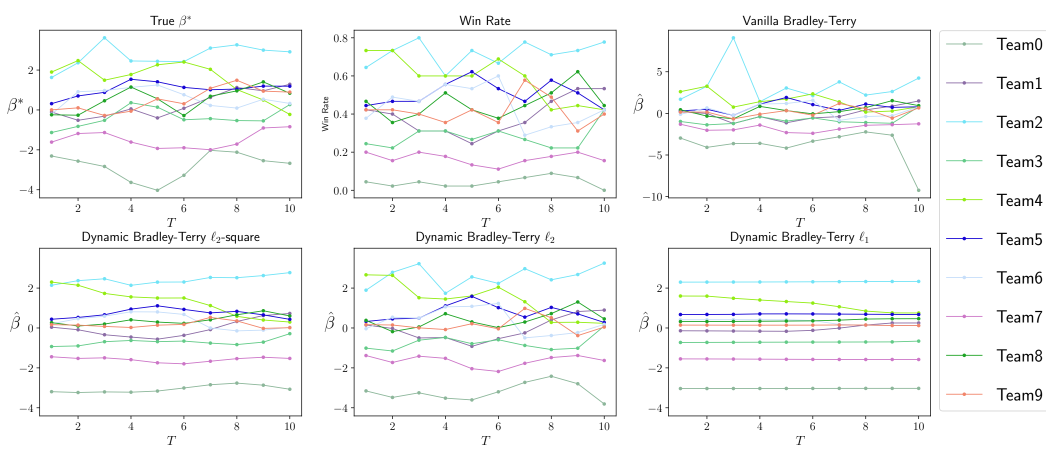

Figure 1: Comparisons of $\boldsymbol{\beta}^*$ and different solutions. From left to right: (first row) $\boldsymbol{\beta}^*$, win rate, vanilla BT; (second row) $\hat{\boldsymbol{\beta}}$ by $\|\cdot\|_2^2$ with LOOCV, $\|\cdot\|_2$, $\|\cdot\|_1$. x-axis: time points

In Figure 1, we also include the paths of the simulated "win rate", given by the proportion of games that each team wins at each time point, and the path of $\hat{\boldsymbol{\beta}}$ estimated by vanilla Bradley-Terry model at each time point. By comparing these curves, we note that our models produce estimates $\hat{\boldsymbol{\beta}}$ that recover the comprehensive global ranking of each team better than "win rate" and vanilla Bradley-Terry model, and meanwhile give more stable paths over time. In figures of $\boldsymbol{\beta}^*$ and $\hat{\boldsymbol{\beta}}$ for $\|\cdot\|_2^2$, $\|\cdot\|_2$, $\|\cdot\|_1$, we can verify that team 2 is almost always dominating and team 0 is maintaining the lowest position in the majority of the time points. However, we can also find some differences of $\hat{\boldsymbol{\beta}}$ from $\boldsymbol{\beta}^*$ which comes from the smoothing property of our model. For instance, the estimated parameters of team 2 and team 0 by our model tend to have larger gaps with other teams, compared with the true $\boldsymbol{\beta}^*$.

Another point illustrated by Figure 1 is that, different choices of the penalty function lead to different shapes of the solution. Specifically, the ℓ_2^2 norm produces the smoothest paths of $\boldsymbol{\beta}$, while the ℓ_1 norm imposes a piecewise constant structure on the paths of $\boldsymbol{\beta}$. In applications, the path of $\boldsymbol{\beta}$ could have various shapes. Therefore, our model is adaptive to different user requirements on the smoothed shape of $\boldsymbol{\beta}$ in different applications.

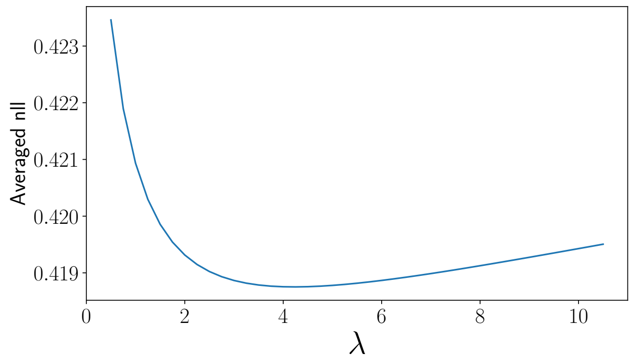

Figure 2: LOOCV curve of Dynamic Bradley-Terry model with ℓ_2-square penalty. x-axis: coefficient of the penalty, y-axis: the averaged negative log-likelihood. The best λ^* is 4.25.

Figure 2 shows the curve of LOOCV of our Dynamic Bradley-Terry model with ℓ_2-square penalty. The curve shows a typical shape of CV curve for tuning parameter. The λ with smallest nll_λ is $\lambda^* = 4.25$. For ℓ_2 and ℓ_1 penalty, since the LOOCV procedure takes much time, we currently just use $\lambda = 4.25$.

7 Application - NFL Data

In order to test our model in practical settings we consider ranking National Football League (NFL) teams over multiple seasons. Specifically we source 5 seasons of openly available NFL data from 2011-2015 (inclusive) using the `nflWAR` package [18]. Each NFL season is comprised of $N = 32$ teams playing $T = 16$ games each over the season i.e. $t \in [16]$ in this case. This means that at each point in time t the pairwise matrix of scores across all 32 teams is sparsely populated with only 16 entries. We fit our time-varying Bradley Terry model over all 16 rounds in the season using the ℓ_2^2 convex penalty and tuning λ using the LOOCV approach described in section 5.2. In order to gauge whether the rankings produced by our model are reasonable we compare our season ending rankings (fit over all games played in that season) with the relevant openly available NFL ELO ratings [12]. The top 10 season-ending rankings from each method across NFL seasons 2011-2015 are summarized in Table 1.

Based on table 1 we observe that we roughly match between 6 to 9 of the top 10 ELO teams consistently over all 5 seasons. However we can see that there are often misalignment with specific ranking values across both ranking methods. For example in the 2014 season we can see that our rankings are reasonably well aligned and notably a match with Seattle being the number one ranked team by both methods. The 2012 season had slightly more misalignment comparatively across both methods. This is captured in the average ranking difference between ELO and our time varying Bradley-Terry model being 3.2 which is slightly higher than the 2014 season value of 1.9. We observe that the average differences across all seasons between ELO and the Bradley Terry model are uniformly positive indicating that ELO ranks the same teams higher than the Bradley-Terry model on average across all seasons.

We note that it is difficult to interpret the differences in great further detail given that the underlying ranking methodologies are fundamentally different. In particular the NFL ELO ranking methodology uses both the pairwise scores between teams (similar to our time-varying Bradley Terry model) but also uses the location information of each game in the modeling process. In this sense we view the comparable top 10 ranking results as an encouraging indication of our model viability in this real world application. We thus view our time-varying Bradley Terry model as a useful *benchmarking*

rank	2011		2012		2013		2014		2015	
	ELO	**BT**	**ELO**	**BT**	**ELO**	**BT**	**ELO**	**BT**	**ELO**	**BT**
1	GB	GB	NE	DEN	SEA	SF	SEA	SEA	SEA	CAR
2	NE	NO	DEN	NE	SF	CAR	NE	DEN	CAR	ARI
3	NO	NE	GB	SEA	NE	SEA	DEN	GB	ARI	KC
4	PIT	SF	SF	MIN	DEN	ARI	GB	NE	KC	SEA
5	BAL	PIT	ATL	SF	CAR	NE	DAL	DAL	DEN	MIN
6	SF	BAL	SEA	GB	CIN	DEN	PIT	PIT	NE	DEN
7	ATL	DET	NYG	IND	NO	NO	BAL	IND	PIT	CIN
8	PHI	ATL	CIN	HOU	ARI	CIN	IND	ARI	CIN	PIT
9	SD	PHI	BAL	WAS	IND	IND	ARI	BUF	GB	GB
10	HOU	SD	HOU	CHI	SD	SD	CIN	DET	MIN	DET
Av. Diff.	2.6		3.2		2.6		1.9		2.8	

Table 1: Bradley-Terry vs. ELO NFL top 10 rankings. Blue: perfect match, yellow: top 10 match

tool for other feature-rich time-varying ranking models since (such as ELO) since our model simply relies on the minimalist time-varying score information for modeling.

8 Conclusion

In this paper we propose a time-varying (dynamic) convex generalization of the original (static) Bradley-Terry model [3]. The underlying goal of the model is to derive the *global* ranking of N distinct objects over T discrete time periods in a principled statistical manner using observed pairwise comparison data at each time point $t \in [T]$. Our proposed model directly captures the temporal dependence structure of the pairwise comparison data in the form of time-varying model parameters. These time-varying model parameters are modeled in a distribution-free frequentist manner in the form of a convex penalty term. The specific choice of convex penalty norm also acts as a 'smoothing' mechanism for the derived time-varying global rankings. In particular the convex penalty term also enables analysis on sparse time-varying pairwise comparison data, which is common in many practical settings including sports analytics.

From a theoretical perspective we prove that a relatively weak condition is necessary and sufficient to guarantee the existence and uniqueness of the solution of our proposed time-varying Bradley-Terry model. This ensures that we fit using limited temporal distributional assumptions and also means that we require much less time-varying pairwise comparison data to derive our global rankings than other non-sparse time-varying frequentist ranking approaches. In order to tune the model penalty term λ we propose both a heuristic and data-driven approach. This heuristic is driven by user preferences in directly setting the degree of smoothing in the final global rankings. Although subjective, this is computational efficient and viable when such user-domain knowledge applies. The data-driven approach uses a sample-splitting approach via leave-one-out cross validation (LOOCV) to derive optimal lambda across a held out set. Although computationally exhaustive, we discuss methods to optimize this further using random sampling and better initial seeding. From an algorithmic perspective we implement various well studied convex optimization algorithms to solve the model efficiently under the ℓ_2^2, ℓ_2 and ℓ_1 penalization norms. This is done exploiting any specific structure in our proposed model under the chosen convex penalty norm for additional computational efficiency gain.

We finally test the practical effectiveness of our model under both simulated and real world settings. In the latter case we separately rank 5 consecutive seasons of open National Football League (NFL) team data [18] from 2011-2015. Our NFL ranking results compare favourably to the well-accepted and feature rich NFL ELO model rankings [12]. We thus view our distribution-free time-varying Bradley-Terry model as a useful *benchmarking tool* for other feature-rich time-varying ranking models since it simply relies on the minimalist time-varying score information for modeling.

References

[1] Alan Agresti. *Categorical Data Analysis*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2002. xvi+710.

[2] U. Böckenholt and W. R. Dillon. Modeling within-subject dependencies in ordinal paired comparison data. *Psychometrika*, 62:411–434, 1997.

[3] Ralph Allan Bradley and Milton E Terry. Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3/4):324–345, 1952.

[4] Manuela Cattelan. Models for paired comparison data: A review with emphasis on dependent data. *Statistical Science*, 27(3):412–433, 2012.

[5] Manuela Cattelan, Cristiano Varin, and David Firth. Dynamic bradley–terry modelling of sports tournaments. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62(1):135–150, 2012.

[6] R. Dittrich, R. Hatzinger, and W. Katzenbeisser. Modelling dependencies in paired comparison data: A log-linear approach. *Comput. Statist. Data Anal.*, 40:39–57, 2002.

[7] Ludwig Fahrmeir and Gerhard Tutz. Dynamic stochastic models for time-dependent ordered paired comparison systems. *Journal of the American Statistical Association*, 89(428):1438–1449, 1994.

[8] Lester R Ford Jr. Solution of a ranking problem from binary comparisons. *The American Mathematical Monthly*, 64(8P2):28–33, 1957.

[9] Mark E. Glickman. Paired comparison models with time varying parameters. *Doctoral thesis*, 2012.

[10] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An introduction to statistical learning*, volume 112. Springer, 2013.

[11] Guido Masarotto and Cristiano Varin. The ranking lasso and its application to sport tournaments. *Ann. Appl. Stat.*, 6(4):1949–1970, 12 2012.

[12] Neil Paine. NFL Elo Ratings Are Back! `https://fivethirtyeight.com/features/nfl-elo-ratings-are-back/`. Accessed: 2018-11-30.

[13] Filip Radlinski and Thorsten Joachims. Active exploration for learning rankings from click-through data. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '07, pages 570–579, New York, NY, USA, 2007. ACM.

[14] S. M. Stigler. Citation patterns in the journals of statistics and probability. *Statist. Sci.*, 9:94–108., 1994.

[15] C. Varin, M. Cattelan, and D. Firth. Statistical modelling of citation exchange between statistics journals. *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, 179(1):1–63, 2016.

[16] Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, Dec 2007.

[17] R. W. M. Wedderburn. Quasi-likelihood functions, generalized linear models, and the gauss-newton method. *Biometrika*, 61(3):439–447, 1974.

[18] Ronald Yurko, Samuel Ventura, and Maksim Horowitz. nflwar: A reproducible method for offensive player evaluation in football. *arXiv preprint arXiv:1802.00998*, 2018.

[19] E. Zermelo. Die berechnung der turnier-ergebnisse als ein maximumproblem der wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 29(1):436–460, Dec 1929.

9 Appendix

9.1 Proof of Theorem

Throughout this section, let f denote the target loss function:

$$f(\boldsymbol{\beta}) = -\sum_{t=1}^{T} \ell_t(\boldsymbol{\beta}^{(t)}) + \lambda \sum_{t=1}^{T-1} h(\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}^{(t+1)}) \tag{12}$$

9.1.1 Uniqueness of the solution with squared-ℓ_2 penalty

We can decompose the loss function with squared-ℓ_2 penalty into two parts:

$$f = \sum_{t=1}^{T} L_t + \sum_{i=1}^{N} R_i \tag{13}$$

where $L_t = -\ell_t(\boldsymbol{\beta}^{(t)})$ and $R_i = \lambda \sum_{t=2}^{T}(\boldsymbol{\beta}_i^{(t)} - \boldsymbol{\beta}_i^{(t-1)})$.

Elements of L_t's Hessian with respect to $\boldsymbol{\beta}^{(t)}$ are:

$$\nabla^2_{\boldsymbol{\beta}_i^{(t)}} L_t = \sum_{i \neq j} (x_{ij}^{(t)} + x_{ji}^{(t)}) \frac{\exp \boldsymbol{\beta}_i^{(t)} \exp \boldsymbol{\beta}_j^{(t)}}{(\exp \beta_i^{(t)} + \exp \beta_j^{(t)})^2} \tag{14}$$

$$\nabla_{\boldsymbol{\beta}_j^{(t)}} \nabla_{\boldsymbol{\beta}_i^{(t)}} L_t = -(x_{ij}^{(t)} + x_{ji}^{(t)}) \frac{\exp \boldsymbol{\beta}_i^{(t)} \exp \boldsymbol{\beta}_j^{(t)}}{(\exp \beta_i^{(t)} + \exp \beta_j^{(t)})^2} \tag{15}$$

Also, R_i's Hessian with respect to $\boldsymbol{\beta}_i := (\boldsymbol{\beta}_i^{(1)}, \boldsymbol{\beta}_i^{(2)}, \ldots, \boldsymbol{\beta}_i^{(T)})$ is:

$$\begin{bmatrix} 2 & -2 & 0 & \cdots & 0 & 0 \\ -2 & 4 & -2 & \cdots & 0 & 0 \\ 0 & -2 & 4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -2 \\ 0 & 0 & 0 & \cdots & -2 & 2 \end{bmatrix} \tag{16}$$

Hence, both kinds of Hessians have zero column sums and so does the sum of them, i.e., the Hessian of the loss function. Let H denote the Hessian of f and $H(\boldsymbol{\beta}_i^{(t)}, \boldsymbol{\beta}_j^{(s)})$ denote each element of H. Then, H has a positive diagonal, and

$$H(\boldsymbol{\beta}_i^{(t)}, \boldsymbol{\beta}_j^{(s)}) = \begin{cases} -(x_{ij}^{(t)} + x_{ji}^{(t)}) \frac{\exp \boldsymbol{\beta}_i^{(t)} \exp \boldsymbol{\beta}_j^{(t)}}{(\exp \beta_i^{(t)} + \exp \beta_j^{(t)})^2} & \text{if } t = s \\ -2 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

As $H(\boldsymbol{\beta}_i^{(t)}, \boldsymbol{\beta}_j^{(t)}) < 0$ if $x_{ij}^{(t)} > 0$ or $x_{ji}^{(t)}$, Condition (1) implies that H can be regarded as a graph Laplacian for a connected graph. Following the classical proof of the property of graph Laplacian [16],

$$v^T H v = \sum_{i<j} |X_{ij}|(v_i - v_j)^2 \geq 0,$$

and Condition (1) guarantees that "=" is achieved if and only if $v = c\mathbf{1}$. This proves the uniqueness up to constant shifts.

9.1.2 Existence of solution

Because of its continuity, h attains its minimum in $\mathbb{R}^T$. Since we still get an equivalent optimization after constant shifting h, we can assume h has minimum value 0 without loss of generality. Also, note that $-\ell_t(\boldsymbol{\beta}_t)$ is non-negative:

$$-\ell_t(\boldsymbol{\beta}_t) = -\sum_{i \neq j} x_{ij}^{(t)} \left(\log \left(\frac{\exp \boldsymbol{\beta}_i^{(t)}}{\exp \boldsymbol{\beta}_i^{(t)} + \exp \boldsymbol{\beta}_j^{(t)}} \right) \right) \geq 0 \tag{18}$$

Plugging in $\boldsymbol{\beta} = \mathbf{0}$, we get an upperbound for the minimum loss function $f^\star$:

$$f^\star \leq (\log 2) \sum_{t=1}^{T} \sum_{i \neq j} x_{ij}^{(t)}. \tag{19}$$

As f is continuous, it suffices to show that the level set with $\sum_{i=1}^{N} \boldsymbol{\beta}_i^{(1)} = 0$ is bounded so that it is compact.

We get an upper-bound on the extent to which $\boldsymbol{\beta}^{(t)}$'s are dispersed in the level set:

$$\|\boldsymbol{\beta}^{(t)} - \boldsymbol{\beta}^{(t+1)}\|_\infty \leq \sqrt{\frac{1}{\lambda} (\log 2) \sum_{i \neq j} x_{ij}^{(t)}} =: B \tag{20}$$

and

$$\|\boldsymbol{\beta}^{(t)} - \overline{\boldsymbol{\beta}}\|_\infty \leq BT \tag{21}$$

where $\overline{\boldsymbol{\beta}} = \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\beta}^{(t)}$.

Then,

$$-\ell_t(\boldsymbol{\beta}^{(t)}) \geq -\sum_{i \neq j} x_{ij}^{(t)} \log \left(\frac{\exp(\overline{\boldsymbol{\beta}}_i - B)}{\exp(\overline{\boldsymbol{\beta}}_i + B) + \exp(\overline{\boldsymbol{\beta}}_j + B)} \right) \tag{22}$$

$$\geq -\sum_{i \neq j} x_{ij}^{(t)} \log \left(\frac{\exp \overline{\boldsymbol{\beta}}_i}{\exp \overline{\boldsymbol{\beta}}_i + \exp \overline{\boldsymbol{\beta}}_j} \right) - 2BT \sum_{i \neq j} x_{ij}^{(t)} \tag{23}$$

$$\geq \sum_{i \neq j} x_i j^{(t)} \log(1 + \exp(\overline{\boldsymbol{\beta}}_j - \overline{\boldsymbol{\beta}}_i)) - 2BT \sum_{i \neq j} x_{ij}^{(t)} \tag{24}$$

Hence, under the level set,

$$(\log 2) \sum_{t=1}^{T} \sum_{i \neq j} x_{ij}^{(t)} \geq f(\boldsymbol{\beta}) \tag{25}$$

$$\geq \sum_{i \neq j} (\sum_{t=1}^{T} x_{ij}^{(t)}) \log(1 + \exp(\overline{\boldsymbol{\beta}}_j - \overline{\boldsymbol{\beta}}_i)) - 2BT \sum_{t=1}^{T} \sum_{i \neq j} x_{ij}^{(t)} \tag{26}$$

and if $\sum_{t=1}^{T} x_{ij}^{(t)} \neq 0$ then $\overline{\boldsymbol{\beta}}_j - \overline{\boldsymbol{\beta}}_i$ is upperbounded.

By Condition (2) and the constraint $\sum_{i=1}^{N} \boldsymbol{\beta}_i^{(1)} = 0$, now every elements of $\boldsymbol{\beta}$ in the level set are bounded. This proves the existence part of the theorem.

9.2 Synthetic Simulation Results

First we generate the parameter $\boldsymbol{\beta}^*$ via a Gaussian process. For each $1 \leq i \leq N$, the value of the mean process is set to be $c_i \mathbf{1} \in \mathbb{R}^T$ with c_i a random sample from standard Normal, and the covariance matrix is set to be a Toeplitz matrix G whose entry is given by $G_{s,t} = 1 - T^{-1/2}|s-t|^{1/2}$. Then we generated the pairwise comparison data from $\boldsymbol{\beta}^*$ with $n_{ij}^{(t)} = 5$ for every $i \neq j$ and $t \in [T]$ and solved (2) by different methods.

9.3 Statistical analyses

9.3.1 Oracle inequality

Suppose that an oracle gives the probability $p_{ij}^{(t)}$ with which team i wins team j at time point t.

oracle loss function:

$$f^\star(\boldsymbol{\beta}) = -\sum_{t=1}^{T}\sum_{i>j}(p_{ij}^{(t)}\boldsymbol{\beta}_i^{(t)} + p_{ji}^{(t)}\boldsymbol{\beta}_j^{(t)} - \log(\exp(\boldsymbol{\beta}_i^{(t)}) + \exp(\boldsymbol{\beta}_j^{(t)})))$$

oracle estimator:

$$\boldsymbol{\beta}^\star = \arg\min_{\boldsymbol{\beta}} f^\star(\boldsymbol{\beta})$$

objective for statistical analyses:

1. probabilistic bound on $f^\star(\hat{\boldsymbol{\beta}}) - f^\star(\boldsymbol{\beta}^*)$
2. probabilistic bound on $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^\star\|$ ($\|\cdot\|_2$ or $\|\cdot\|_\infty$)

9.3.2 Consistency in ranking

After we get probabilistic bound on $\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}^*\|$, under an assumption on a lower bound for $|{\boldsymbol{\beta}^\star}_i^{(t)} - {\boldsymbol{\beta}^\star}_j^{(t)}|$, consistency in ranking in terms that

$$\hat{\boldsymbol{\beta}}_i^{(t)} > \hat{\boldsymbol{\beta}}_j^{(t)} \iff {\boldsymbol{\beta}^\star}_i^{(t)} > {\boldsymbol{\beta}^\star}_j^{(t)}$$

from Bradley-Terry score would be easily get.

We might need arguments or references on nice properties of ranking from $\boldsymbol{\beta}^\star$.

9.3.3 Consistency in rank change detection

We might need assumptions on a lower bound for $|{\boldsymbol{\beta}^\star}'_i(t) - {\boldsymbol{\beta}^\star}'_j(t)|$ while ${\boldsymbol{\beta}^\star}'_i(t)$ is a derivative of ${\boldsymbol{\beta}^\star}_i(t)$ as a function of time t.