Dynamic Ranking and Translation Synchronization

Ernesto Araya* Inria Lille

Eglantine Karlé* Inria Lille

Hemant Tyagi* Inria Lille

1 Introduction

In many applications, such as sport tournaments or recommendation systems, we are given outcomes of pairwise comparisons between n items. The goal is to then infer the latent strength of each item and/or their ranking. This has been thorougly studied both in theory and in practice for a single comparison graph G. However, many applications (e.g., sports tournaments) involve time-evolving data, for which only limited theoretical results exist, under local (Lipschitz-type) smoothness assumptions [2, 3]. We add to this recent line of work by considering a more general setting with a global smoothness assumption [1].

2 Problem setup

The data consist of pairwise comparisons on a set of items $[n] = \{1, 2, \ldots, n\}$ at different times t on a uniform grid $\mathcal{T} = \left\{\frac{k}{T} : k = 0, \ldots, T\right\} \subset [0, 1]$. At each $t \in \mathcal{T}$, we denote the comparison graph $G_t = ([n], \mathcal{E}_t)$ where $\mathcal{E}_t$ is the set of undirected edges. Let $z_t^* = (z_{t,1}^*, \ldots, z_{t,n}^*)^\top \in \mathbb{R}^n$, where $z_{t,i}^*$ denotes the latent strength of item i at time t. For each $t \in \mathcal{T}$ and $\{i, j\} \in \mathcal{E}_t$, we observe the noisy measurement

$$y_{ij}(t) = z_{t,i}^* - z_{t,j}^* + \epsilon_{ij}(t),$$

where $\epsilon_{ij}(t)$ are i.i.d. centered subgaussian random variables. This corresponds to the Translation Synchronization model at each time $t \in \mathcal{T}$. It can also be linked to the classic BTL model, see for e.g., [1]. For meaningful estimation of z_t^*, we assume z_t^* is centered (i.e., $\sum_{i=1}^n z_{t,i}^* = 0$) and moreover evolves smoothly with t.

Assumption 1 (Global ℓ_2-smoothness)**.** *Let $C \in \mathbb{R}^{n \times \binom{n}{2}}$ be the edge incidence matrix of the complete graph. We assume that*

$$\sum_{k=0}^{T-1} \left\| C^\top (z_k^* - z_{k+1}^*) \right\|_2^2 \leq S_T. \tag{2.1}$$

Denoting $z^* = (z_0^{*\top} \cdots z_T^{*\top})^\top$, we can write (2.1) as $\|Ez^*\|^2 \leq S_T$, where E is a suitable smoothness operator [1], indicating that z^* lies close to the null space of E. We propose the following estimators for estimating z_t^*.

1. **Smoothness-penalized least-squares.**

$$\hat{z} = \underset{\substack{z_0, \ldots, z_T \in \mathbb{R}^n \\ z_k^\top \mathbf{1}_n = 0}}{\operatorname{argmin}} \sum_{t \in \mathcal{T}} \sum_{(i,j) \in \vec{\mathcal{E}}_t} \left(y_{ij}(t) - (z_{t,i} - z_{t,j})\right)^2 + \lambda \left\| Ez \right\|_2^2,$$

where $\hat{z} = \begin{pmatrix} \hat{z}_0^\top & \ldots & \hat{z}_T^\top \end{pmatrix}^\top \in \mathbb{R}^{n(T+1)}$.

2. **Projection method.** For each $t \in \mathcal{T}$, let $\check{z}_t \in \mathbb{R}^n$ be the least-squares solution of

$$y_{ij}(t) = z_{t,i} - z_{t,j} \quad \forall \{i, j\} \in \mathcal{E}_t$$

and denote $\check{z} = \begin{pmatrix} \check{z}_0^\top & \ldots & \check{z}_T^\top \end{pmatrix}^\top$. For $\tau > 0$, let P_τ be the projection matrix on the eigenspace of $E^\top E$ corresponding to eigenvalues smaller than τ. Then define

$$\hat{z} = \begin{pmatrix} \hat{z}_0^\top & \ldots & \hat{z}_T^\top \end{pmatrix}^\top = P_\tau \check{z} \in \mathbb{R}^{n(T+1)}.$$

3 Main results

Our main results are summarized below, see [1] for details.

Theorem A (Penalized least-squares)**.** *Suppose that G_t is connected for each $t \in \mathcal{T}$ and $\lambda = (\frac{T}{S_T})^{2/5}$, then it holds w.h.p.*

$$\|\hat{z} - z^*\|_2^2 = O(T^{4/5} S_T^{1/5}).$$

We show in [1] that the bound in Theorem A can be improved to $O(T^{2/3} S_T^{1/3})$ under additional technical assumptions on the graphs $(G_t)_{t \in \mathcal{T}}$.

Theorem B (Projection method)**.** *Suppose that G_t is connected for each $t \in \mathcal{T}$ and $\tau = (\frac{S_T}{T})^{2/3}$, then it holds w.h.p.*

$$\|\hat{z} - z^*\|_2^2 = O(T^{2/3} S_T^{1/3}).$$

Note that if $S_T = o(T)$, then both theorems imply that the mean squared error $\frac{1}{T+1} \|\hat{z} - z^*\|_2^2 = o(1)$ as $T \to \infty$. Moreover, if the strengths are Lipschitz functions of time, then $S_T = O(1/T)$, and Theorem B leads to the optimal empirical L_2-norm rate for the estimation of Lipschitz functions on a uniform grid [1, Remark 4]. We also assess the performance of our estimators via experiments on synthetic data, and show that they acheive similar results for different choices of S_T. They also display comparable performance to existing estimators in the Dynamic BTL setting [2, 3]. We finally test our estimators on the Netflix Prize and the English Premier League datasets, and show that the dynamic setup leads to an improvement in the performance compared to the static case.

References

[1] E. Araya, E. Karlé, and H. Tyagi. Dynamic ranking and translation synchronization. *arXiv preprint arXiv:2207.01455*, 2022.

[2] H. Bong, W. Li, S. Shrotriya, and A. Rinaldo. Nonparametric estimation in the dynamic Bradley-Terry model. In *AISTATS*, pages 3317–3326, 2020.

[3] E. Karlé and H. Tyagi. Dynamic ranking with the btl model: A nearest neighbor based rank centrality method. *arXiv preprint arXiv:2109.13743*, 2021.

*Emails: ernesto-javier.araya-valdivia@inria.fr, eglantine.karle@inria.fr,hemant.tyagi@inria.fr