

STATISTICAL OPTIMALITY AND ALGORITHMS FOR TOP-K AND TOTAL RANKING

CHAO GAO

Abstract: In this short course, we will discuss the problem of ranking with partially observed pairwise comparisons in the setting of Bradley-Terry-Luce (BTL) model. There are two fundamental problems: 1) top-K ranking, which is to select the set of K players with top performances; 2) total ranking, which is to rank the entire set of players. Both ranking problems find important applications in web search, recommender systems and sports competition.

In the first presentation, we will consider the top-K ranking problem. The statistical properties of two popular algorithms, MLE and rank centrality (spectral ranking) will be precisely characterized. In terms of both partial and exact recovery, the MLE achieves optimality with matching lower bounds. The spectral method is shown to be generally sub-optimal, though it has the same order of sample complexity as the MLE. Our theory also reveals the essentially only situation when the spectral method is optimal. This turns out to be the most favorable choice of skill parameter given the separation of the two groups.

The second presentation will be focused on total ranking. The problem is to find a permutation vector to rank the entire set of players. We will show that the minimax rate of the problem with respect to the Kendall's tau loss exhibits a transition between an exponential rate and a polynomial rate depending on the signal to noise ratio of the problem. The optimal algorithm consists of two stages. In the first stage, games with very high or low scores are used to partition the entire set of players into different leagues. In the second stage, games that are very close are used to rank the players within each league. We will give intuition and some analysis to show why the algorithm works optimally.