

# D-08 BTモデルのオンライン推定に対するリグレット解析

松本 一成・畑埜 晃平・瀧本 英二(九州大学)

## 各時刻 $t = 1, 2, \dots, T$ のプロトコル

1  $\theta_t \in \Theta$  を予測

環境



2

$(i_t, j_t)$  をもらう

3

損失

$$f_t(\theta_t) = -\ln P((i_t, j_t) | \theta_t)$$

をこうむる

凸関数でない!

## Bradley-Terry(BT) モデル

各チームの強さパラメータを基とした  
2チーム間の勝敗に関する確率モデル

$S = \{1, 2, \dots, n\}$  : チームの集合

$\Theta = \{\theta \in \mathbb{R}_+^n | \theta_i: \text{チーム}i\text{の強さ}\}$

$(i, j)$  : チーム*i*がチーム*j*に勝つ事象

$$P((i, j) | \theta) = \theta_i / (\theta_i + \theta_j)$$

目標: Regretを小さくする

$$\text{Regret} = \sum_{t=1}^T f_t(\theta_t) - \sum_{t=1}^T f_t(\theta^*)$$

$$\theta^* = \arg \min_{\theta \in \Theta} \sum_{t=1}^T f_t(\theta)$$

## 既存手法と提案手法

$$K = \max_{i,j} \theta_i^* / \theta_j^* \text{とする}$$

### 既存手法

損失関数を凸関数になるように  
パラメータ変換し,  
凸最適化手法を利用

Kが既知

既存手法

$$\text{Regret} \leq n K \ln T$$

提案手法1

$$\text{Regret} \leq n \ln K \ln T$$

Kが未知

既存手法

動作しない

提案手法2

$$\text{Regret} \leq n^{\frac{3}{2}} \sqrt{\ln n} \ln K \sqrt{T}$$