2002年日本オペレーションズ・リサーチ学会
春季研究発表会

距離関数最適化によるダブルス成績からの個人能力の推定

申請中 日本大学生産工学部 † 鬼頭 正浩
Nihon University Kito Masahiro
01205220 日本大学生産工学部 篠原 正明
Nihon University Shinohara Masaaki

1 はじめに

テニスにおけるダブルスの成績は、ダブルスを組んでいる各々のプレイヤーの成績となんらかの関係があると考えられる。本研究では、その関係を定式化することによって、ダブルスの成績（勝率）からシングルスにおける各々のプレイヤーの個人能力を推定する。

2 変数の定義

2.1 P、D、A について

- **各選手**（i、j、k、…）**の個人能力**を P_i、P_j、P_k、……というように表す。ここで P は、$0 \leq P \leq 1$ とする。
- **ダブルスチームの能力**は、(i と j) の組み合わせの場合 D_{ij}、(k と l) の組み合わせの場合 D_{kl} というように表す。D は、$0 \leq D \leq 1$ とする。
- **ダブルスチーム同士の対戦成績の理論値** (i と j) vs (k と l) の組み合わせの場合で (i と j) が勝つ確率を $Q_{ij,kl}$ と表し、$0 \leq Q \leq 1$ とする。
- **ダブルスチーム同士の対戦成績の参照値** (i と j) vs (k と l) の組み合わせの場合で (i と j) が勝つ確率データ値を $A_{ij,kl}$ と表し、$0 \leq A \leq 1$ とする。

2.2 ダブルスチーム能力 D_{ij} の定義

D_{ij} には、いくつかのモデルが考えられる。

- 直列モデル

$$D_{ij} = P_i \cdot P_j \tag{1}$$

- 並列モデル

$$D_{ij} = P_i + P_j - P_i \cdot P_j \tag{2}$$

本研究では、前者の直列モデルを採用する。

3 ダブルス成績理論値の定式化

ダブルスの能力 D とダブルスの成績 Q の関係モデルに関しても様々なモデルが考えられるが、ここでは次式のモデルを採用する。

3.1 割合比率モデル

$$\begin{aligned} Q_{ij,kl} &= \frac{D_{ij}}{D_{ij} + D_{kl}} \\ &= \frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} \end{aligned} \tag{3}$$

これを割合比率モデル（Bradly Terry モデル）という。

4 距離関数最適化による個人能力推定アルゴリズム

4.1 二乗和最小化

Q とダブルス成績のデータ値 A の差の二乗和を最小化をすることによって個人能力 P を推定する。

$$\begin{array}{ll} \text{minimize} & \sum \left(\frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - A_{ij,kl} \right)^2 \\ \text{subject to} & 0 \leq P \leq 1 \end{array} \tag{4}$$

※上式の $\sum$ は、関連するダブルス試合 (ij, kl) の組み合わせについての総和である（以下同様）。

4.2 絶対値和最小化

式 (4) において 2 乗していたところを絶対値に変えたモデルを、次式で与える。

$$\begin{array}{ll} \text{minimize} & \sum \left| \frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - A_{ij,kl} \right| \\ \text{subject to} & 0 \leq P \leq 1 \end{array} \tag{5}$$

4.3 最大値最小化

全ての組み合わせの中から、最大となる二乗偏差 $\left(\frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - A_{ij,kl}\right)^2$ を最小化するモデルを考える。

$$\begin{aligned} &\text{minimize} && \max\left(\frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - A_{ij,kl}\right)^2 \\ &\text{subject to} && 0 \le P \le 1 \end{aligned} \tag{6}$$

4.4 対数二乗和最小化

式 (4) の最小化する関数を対数に変えたモデルを、次のように考える。

$$\begin{aligned} &\text{minimize} && \sum\left(\log\frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - \log A_{ij,kl}\right)^2 \\ &\text{subject to} && 0 \le P \le 1 \end{aligned} \tag{7}$$

4.5 対数絶対値和最小化

式 (5) の関数を対数に変えたモデルを、次のように考える。

$$\begin{aligned} &\text{minimize} && \sum\left|\log\frac{P_i \cdot P_j}{P_i \cdot P_j + P_k \cdot P_l} - \log A_{ij,kl}\right| \\ &\text{subject to} && 0 \le P \le 1 \end{aligned} \tag{8}$$

5 例題

i、j、k、l の 4 人のプレイヤーの個人能力を求める。ダブルスチーム同士の対戦成績の参照データは次の通りである。

表 5.1 対戦成績の参照データ

1,2 vs 3,4 (1,2 の勝率) $A_{12,34}$	0.9
1,4 vs 2,3 (1,4 の勝率) $A_{14,23}$	0.4
1,3 vs 2,4 (1,3 の勝率) $A_{13,24}$	0.7

このデータと 4 で提案した各種モデルを、**NUOPT** (数理システム) に入力して非線形計画問題を解いた。求解結果は次の通りである。

5.1 二乗和最小化の場合

表 5.2 P_1, P_2, P_3, P_4 の能力

P_1	0.519363276
P_2	0.416417032
P_3	0.212029429
P_4	0.113334499

5.2 絶対値和最小化の場合

表 5.3 P_1, P_2, P_3, P_4 の能力

P_1	0.15999784
P_2	0.128283664
P_3	0.08690603
P_4	0.046453227

5.3 最大値最小化の場合

表 5.4 P_1, P_2, P_3, P_4 の能力

P_1	0.641267259
P_2	0.514319302
P_3	0.262420795
P_4	0.140317509

5.4 対数二乗和最小化の場合

表 5.5 P_1, P_2, P_3, P_4 の能力

P_1	0.502493782
P_2	0.402891414
P_3	0.205143716
P_4	0.109653939

5.5 対数絶対値和最小化の場合

表 5.6 P_1, P_2, P_3, P_4 の能力

P_1	0.158531313
P_2	0.119360881
P_3	0.118691685
P_4	0.059576639

6 まとめ

4.1〜4.5 のように少しずつ違う推定アルゴリズムを用いた場合でも、順位の逆転は見られなかった。また、$A_{ij,kl}$ の数値によって 2 人のプレイヤが同能力を表す数値となってしまうことがある。

今後は、大規模データや別の D-P 関係モデル、Q-D 関係モデルについて検討していきたい。とくに、テニスプレイヤーの個人能力を P などの 1 つのスカラーパラメータで代表するのではなく、攻撃能力 P_O、守備能力 P_D などの 2 つ以上のパラメータで特徴づけるアプローチも今後の研究課題である。また、ダブルス成績データの収集法も、セット単位、ゲーム単位、ポイント単位などが考えられる。