Ranking Individuals by Group Comparisons

Chih-Jen Lin
Department of Computer Science
National Taiwan University

Joint work with Tzu-Kuo Huang and Ruby C. Weng
Problem
Ranking individuals by group comparisons
Existing approaches
Our approaches
Real applications:
  Ranking bridge partnerships
Discussion and conclusions
The Problem

- Many sports are team comparisons
  - How to rank individuals?
- Rank a basketball player by average points
  - But ignore teammates’/opponents’ abilities.
- In bridge
  - Two partnerships vs. two partnerships
  - Match record shows which two are better
  - But how to rank partnerships?
Multi-class classification by error-correcting codes
[Dietterich and Bakiri, 1995, Allwein et al., 2001]
Some classes vs. some others
Finding the winning class (individual)
A Naive Approach: SUM

- Summing # of winning games

\[
\sum_{i:s \in I'} n_i^+ + \sum_{i:s \in I'} n_i^- \\
\sum_{i:s \in I'} 1
\]

- Not consider opponents’ abilities
  Susceptible to individuals playing very few (or many) games

- Not consider teammates’ abilities
  Strong and weak players: the same credits

- Ranking by SUM similar to that of teams.
Games: \{1,3\} vs. \{2\}; \{1,4\} vs. \{2,3,5\}, etc.

Individual j’s ability: \( p_j \geq 0 \)

The i th setting: team \( I_i^+ \) vs. team \( I_i^- \)

\[
P(I_i^+ \text{ beats } I_i^-) = \frac{\sum_{j: j \in I_i^+} p_j}{\sum_{j: j \in I_i} p_j} = \frac{\sum_{j: j \in I_i^+} p_j}{\sum_{j: j \in I_i^+} p_j + \sum_{j: j \in I_i^-} p_j}
\]
Our 1st Approach [Huang et al., 2005]

- Games: \{1,3\} vs. \{2\}; \{1,4\} vs. \{2,3,5\}, etc.
- Individual j’s ability: \( p_j \geq 0 \)
- The \( i \)th setting: team \( I_i^+ \) vs. team \( I_i^- \)

\[
P(I_i^+ \text{ beats } I_i^-) = \frac{\sum_{j: j \in I_i^+} p_j}{\sum_{j: j \in I_i} p_j} = \frac{\sum_{j: j \in I_i^+} p_j}{\sum_{j: j \in I_i^+} p_j + \sum_{j: j \in I_i^-} p_j}
\]

- Extension of Bradley-Terry model [Bradley and Terry, 1952]

\[
P(\text{individual } i \text{ beats individual } j) = \frac{p_i}{p_i + p_j}
\]
Minimizing the negative log-likelihood

\[
\min_p \quad - \sum_{i=1}^{m} \left( n_i^+ \log \frac{\sum_{j: j \in I_i^+} p_j}{\sum_{j: j \in I_i} p_j} + n_i^- \log \frac{\sum_{j: j \in I_i^-} p_j}{\sum_{j: j \in I_i} p_j} \right)
\]

subject to \( \sum_{j=1}^{k} p_j = 1, 0 \leq p_j, j = 1, \ldots, k. \)

- \( n_i^+ \) and \( n_i^- \): \# wins by \( I_i^+ \) and \( I_i^- \); \( n_i \equiv n_i^+ + n_i^- \)
- May be non-convex
Estimation by an iterative method
stationary point

Extensions
weighted individual skill
home-field advantage
ties
comparisons with more than two teams.
Our 2nd Approach: [Huang et al., 2006]

- An exponential model
- $k$ individuals’ abilities: a vector $\mathbf{v} \in \mathbb{R}^k$, $-\infty < v_s < \infty$, $s = 1, \ldots, k$.
- Ability of team $I_i^+$ and $I_i^-$: sum of members’

\[
T_i^+ \equiv \sum_{s: s \in I_i^+} v_s \quad \text{and} \quad T_i^- \equiv \sum_{s: s \in I_i^-} v_s.
\]

- Teams’ actual performances: random variables $Y_i^+$ and $Y_i^-$

\[
P(I_i^+ \text{ beats } I_i^-) \equiv P(Y_i^+ - Y_i^- > 0).
\]
Distribution unknown, assume doubly-exponential extreme-value distribution

\[ P(Y_i^+ \leq y) = \exp(-e^{-(y - T_i^+)}) , \]

Probability that \( I_i^+ \) wins

\[ P(I_i^+ \text{ beats } I_i^-) = \frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} . \]

Can assume normal; but model more complex
Estimation: Regularized Least Square

- $n_i^+$ and $n_i^-$: \# games teams $I_i^+$ and $I_i^-$ win

\[
\frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} \approx \frac{n_i^+}{n_i^+ + n_i^-} \Rightarrow e^{T_i^+ - T_i^-} \approx \frac{n_i^+}{n_i^-}.
\]

- Solve

\[
\min_v \sum_{i=1}^m ((T_i^+ - T_i^-) - \log(n_i^+/n_i^-))^2
\]

- Unique solution:

Adding a regularized term $\mu v^T v$
Estimation: Maximum Likelihood (ML)

- Negative log-likelihood function:
  \[
  \arg \min l(v)
  \]
  where

\[
l(v) \equiv - \sum_{i=1}^{m} \left( n_i^+ \log \frac{e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} + n_i^- \log \frac{e^{T_i^-}}{e^{T_i^+} + e^{T_i^-}} \right)
\]

- Convex
Maximum Likelihood: Iterative Scaling

- Standard optimization methods: Newton’s etc.
Maximum Likelihood: Iterative Scaling

- Standard optimization methods: Newton’s etc.
- Simple iterative scaling
  
  Update the $s$th component;
  
  $\delta \equiv [0, \ldots, 0, \delta_s, 0, \ldots, 0]^T$

$$l(v + \delta) - l(v) \leq - \left( \sum_{i:s \in I_+^i} n_i^+ + \sum_{i:s \in I_-^i} n_i^- \right) \delta_s +$$

$$\left( \sum_{i:s \in I_+^i} \frac{n_i e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} + \sum_{i:s \in I_-^i} \frac{n_i e^{T_i^-}}{e^{T_i^+} + e^{T_i^-}} \right) (e^{\delta_s} - 1)$$
Updating rule

\[ \nu_s \leftarrow \nu_s + \log \frac{\sum_{i: s \in I_i^+} n_i^+ + \sum_{i: s \in I_i^-} n_i^-}{\sum_{i: s \in I_i^+} \frac{n_i e^{T_i^+}}{e^{T_i^+} + e^{T_i^-}} + \sum_{i: s \in I_i^-} \frac{n_i e^{T_i^-}}{e^{T_i^+} + e^{T_i^-}}} . \]

Under a minor assumption,

any limit point a global minimum
Ranking Bridge Partnerships

- A match setting:

  Team A: two partnerships \((A_1, A_2), (A_3, A_4)\)
  Team B: two partnerships \((B_1, B_2), (B_3, B_4)\)
  
  **Same board** for \((B_1, B_2), (A_3, A_4)\)

  Avoid uneven hands
Bridge Scoring

- First three boards; India vs. Portugal in Bermuda Bowl 2005

<table>
<thead>
<tr>
<th>Board</th>
<th>Table I</th>
<th>Table II</th>
<th>IMPs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NS</td>
<td>EW</td>
<td>NS</td>
</tr>
<tr>
<td>1</td>
<td>1510</td>
<td></td>
<td>1510</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>650</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>630</td>
<td></td>
<td>630</td>
</tr>
</tbody>
</table>

- International Match Points (IMPs): difference in two teams’ total scores
- IMPs of all rounds ⇒ Victory Points (VP)

Overall results between two teams; used as $n_i^+$, $n_i^-$
Experimental Settings

- Bermuda Bowl 2005, the most prestigious bridge event
- 22 teams, round robin
  \[
  \binom{22}{2} = 231 \text{ matches}
  \]
- Most teams: six players \(\Rightarrow\) three fixed partnerships playing similar \# matches
- Total 69 partnerships

Other details not shown here
Results:
HNG: generalized Bradley-Terry model
RLS, ML: estimations of the exponential model

<table>
<thead>
<tr>
<th>Team</th>
<th>Partnership rankings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RLS</td>
</tr>
<tr>
<td>IT</td>
<td>21</td>
</tr>
<tr>
<td>US2</td>
<td>63</td>
</tr>
<tr>
<td>US1</td>
<td>9</td>
</tr>
<tr>
<td>SE</td>
<td>2</td>
</tr>
<tr>
<td>IN</td>
<td>14</td>
</tr>
<tr>
<td>AR</td>
<td>33</td>
</tr>
<tr>
<td>EG</td>
<td>47</td>
</tr>
<tr>
<td>AU</td>
<td>44</td>
</tr>
<tr>
<td>NZ</td>
<td>68</td>
</tr>
<tr>
<td>UK</td>
<td>10</td>
</tr>
</tbody>
</table>

Chih-Jen Lin (National Taiwan Univ.)
## Results of U.S.A.2

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>Score</th>
<th>vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>0 25</td>
<td>NL(10)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>12 18</td>
<td>ZA(12)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>11 19</td>
<td>UK(17)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>14 16</td>
<td>TW(19)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>25  5</td>
<td>IN(5)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>22  8</td>
<td>NZ(16)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>22  8</td>
<td>PL(20)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>12 18</td>
<td>US1(3)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>22  8</td>
<td>SE(4)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>19 11</td>
<td>EG(7)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>19 11</td>
<td>BR(8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>Score</th>
<th>vs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>24  6</td>
<td>RU(13)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>20 10</td>
<td>PT(14)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>20 10</td>
<td>CA(18)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>25  3</td>
<td>JO(22)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>15 15</td>
<td>AU(15)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>12 18</td>
<td>IT(1)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>13 17</td>
<td>AR(6)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>14 16</td>
<td>JP(9)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>13 17</td>
<td>CN(11)</td>
</tr>
<tr>
<td>⋆</td>
<td>⋆</td>
<td></td>
<td>14 16</td>
<td>GP(21)</td>
</tr>
</tbody>
</table>

- **(P1,P3): significantly win 3, but lose 1**
- **(P2,P3): win 7 matches but lose 5**
- **P2 is not better than P1**
Ranking Many Objects with Limited Resources

- New research we are doing
  See talk by Ting-Fan Wu in the workshop next Friday

- Current method: scoring
  Conference reviewing
  Hot or Not http://www.hotornot.com
Our New Website: Hotter or Notter

[ See Pairs | Hall of Fame | Add/Delete Photos | FAQ ]

Who is HOTTER?

This is a research website maintained by the Machine Learning group at Computer Science Dept., National Taiwan Univ.

It is built with HOTorNOT API to study users' behaviors on comparing photos.

Please direct your questions to {b89096@clin}@csie.ntu.edu.tw
It takes 0.112572 seconds to process your request
Please help to cask votes. We need more data

http://svm.csie.ntu.edu.tw/~ranker

Thank You
References I

Reducing multiclass to binary: a unifying approach for margin classifiers.
*JMLR*, 1:113–141.

The rank analysis of incomplete block designs: I. the method of paired comparisons.

Solving multiclass learning problems via error-correcting output codes.

Ranking individuals by group comparisons.
In *Proceedings of the Twenty Third International Conference on Machine Learning (ICML)*.

A generalized Bradley-Terry model: From group competition to individual skill.