University of Illinois at Urbana-Champaign
Department of Electrical and Computer Engineering

Aggregating Rankings with Positional Constraints

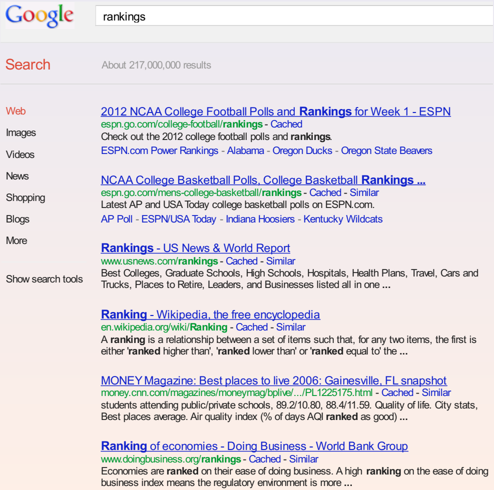
Farzad Farnoud and Olgica Milenkovic

Information Theory Workshop
September 10, 2013

Rankings

A ranking is an arrangement of objects based on a certain quality or set of properties.

- Sport teams
- Schools
- Companies
- Search results
- Politicians
- Movies
- Pets
- Anything really



Rankings

Rank processing used in:

- Statistics: Order statistics, statistical tests, measuring correlation.
- Economics and social sciences: Collective decision making (social choice); Individual decision making.
- Artificial Intelligence: Multi-agent automated systems.
- Recommender Systems: Measuring similarities of users, suggesting products based on ratings/rankings.
- Information Retrieval: Study of search engines.
- Coding Theory: Power-line communications; Flash memories.

Rank aggregation

Rank aggregation: One of the most important rank processing technique. Task is to combine a set of rankings such that the result is a ranking "representative" of the set.

Expert 1	**Expert 2**	**Expert 3**	**Aggregate**
Nadal	Djokovic	Djokovic	?
Djokovic	Nadal	Nadal	?
Murray	Del Potro	Del Potro	?
Federer	Murray	Wawrinka	?
Del Potro	Federer	Murray	?

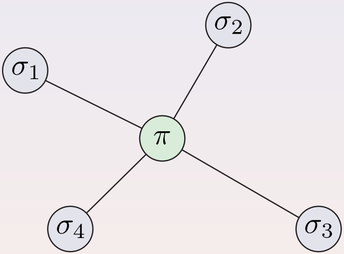
- Will represent rankings by permutations over $[n] = \{1, \ldots, n\}$.
- Partial orders, weak orders discussed elsewhere (Raisali, Farnoud, and M, 2013).

Rank aggregation

One of the most common approaches to rank aggregation is distance-based rank aggregation.

Given expert rankings $\sigma_1, \ldots, \sigma_m$, the rank aggregation problem can be stated as

$$\pi^* = \arg\min_{\pi} \sum_{i=1}^{m} \mathrm{d}(\pi, \sigma_i).$$



Equivalently, want the median of the votes.

How do we choose the distance?

Classical distances on rankings

- Cayley distance [Cayley, 1849]: minimum number of swaps; $\mathrm{d}_T(abc, cba) = 1$.
- Spearman's footrule: sum of differences of ranks; $\mathrm{d}_F(abc, cba) = 2 + 0 + 2 = 4$.
- Kendall τ [Kendall 1948]: minimum number of swaps of adjacent entries; $\mathrm{d}_K(abc, cba) = 3$.
- Hamming distance: number of positions with different entries; $\mathrm{d}_H(abc, cba) = 2$.
- Ulam distance [Beyer, Stein, Ulam 1972]: n minus length of longest common subsequence: $\mathrm{d}_U(abc, cba) = 2$.

Positional Constraints: When classical distances are not suitable

σ	π_1	π_2
Melbourne	Melbourne	Vienna
Vienna	Vienna	Melbourne
Vancouver	Vancouver	Vancouver
Toronto	Toronto	Toronto
Calgary	Calgary	Calgary
Adelaide	Adelaide	Adelaide
Sydney	Sydney	Sydney
Helsinki	Perth	Helsinki
Perth	Helsinki	Perth
Auckland	Auckland	Auckland

- People pay more attention to top of ranking rather than bottom.
- A voter with vote σ is likely to prefer π_1 to π_2.
- Unfortunately, all classical distances treat different positions in the rankings uniformly: $\mathrm{d}(\sigma, \pi_1) = \mathrm{d}(\sigma, \pi_2)$.

σ	π_1	π_2
Melbourne	Melbourne	Vienna
Vienna	Vienna	Melbourne
Vancouver	Vancouver	Vancouver
Toronto	Toronto	Toronto
Calgary	Calgary	Calgary
Adelaide	Adelaide	Adelaide
Sydney	Sydney	Sydney
Helsinki	Perth	Helsinki
Perth	Helsinki	Perth
Auckland	Auckland	Auckland

Kemeny's axioms

Kemeny proposed an axiomatic approach for determining appropriate distance on rankings [Kemeny, 1959]:

1. $\mathrm{d}(\cdot,\cdot)$ is a metric.
2. Relabeling of objects does not change the distance.
3. If two rankings agree except on a "segment," position of segment within ranking is not important: $\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe)$.
4. $\mathrm{d}(\pi,\sigma) = \mathrm{d}(\pi,\omega) + \mathrm{d}(\omega,\sigma)$ iff ω is "between π and σ.
   Betweenness: for $a, b \in \{1, \ldots, n\}$, if π and σ both rank a before b, then ω also ranks a before b.

Kendall τ

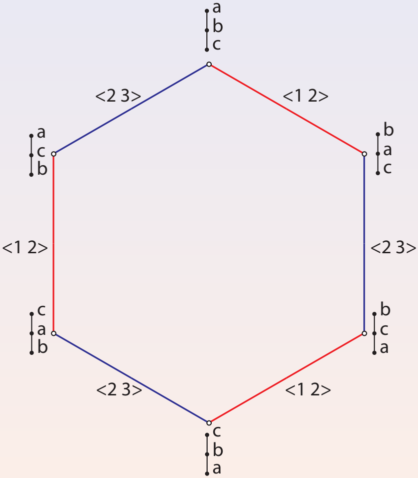
Theorem [Kemeny 1959]

The unique distance that satisfies Kemeny's axioms is Kendall τ.

Represent rankings by vertices in a graph. Neighboring vertices differ by an adjacent swap.

Swap of elements in positions i and j is denoted by $\langle i\ j \rangle$.

Distance is the length of the shortest path.



Generalizing the Kendall distance

Kendall τ ensures all adjacent swaps treated the same. How should the axioms be adapted for positional constraints?

– Let us remove the third axiom

1. Distance function is a pseudo-metric.
2. Relabeling of objects does not change distance.
3. ~~If two rankings agree except on a "segment," position of segment within ranking is not important:~~ $\cancel{\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe)}$.
4. $\mathrm{d}(\sigma, \pi) = \mathrm{d}(\pi, \omega) + d(\omega, \sigma)$ iff ω is between π and σ.

Generalizing the Kendall distance

Kendall τ ensures all adjacent swaps treated the same. How should the axioms be adapted for positional constraints?

– Let us remove the third axiom

1. Distance function is a pseudo-metric.
2. Relabeling of objects does not change distance.
3. ~~If two rankings agree except on a "segment," position of segment within ranking is not important:~~ ~~$\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe)$.~~
4. $\mathrm{d}(\sigma, \pi) = \mathrm{d}(\pi, \omega) + d(\omega, \sigma)$ iff ω is between π and σ.

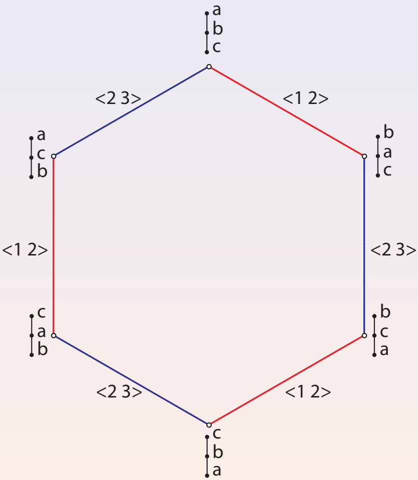
Theorem

For rankings without ties, the unique distance satisfying the above axioms is (again) the Kendall τ.

Why modify the fourth axiom?

Lemma

Axioms 2 and 4 imply that all adjacent swaps have same contribution to rank distance.



Why modify the fourth axiom?

Lemma

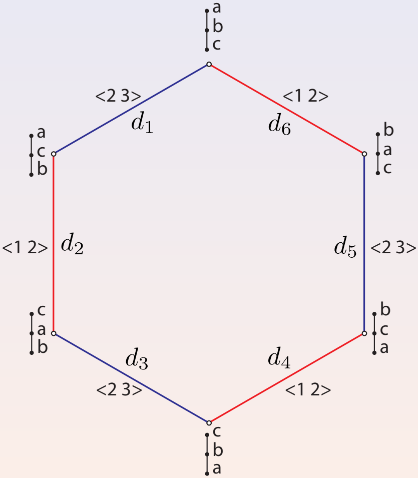
Axioms 2 and 4 imply that all adjacent swaps have same contribution to rank distance.

Proof sketch for $n = 3$

From axiom 3, $\mathrm{d}(abc, cba) = d_1 + d_2 + d_3 = d_4 + d_5 + d_6$.

From axiom 2,
$d_1 = d_3 = d_5, \; d_2 = d_4 = d_6$.

Thus,
$d_1 = d_2 = d_3 = d_4 = d_5 = d_6$.



Kendall distance with positional constraints

Changing Kemeny's axioms

1. Distance function is a pseudo-metric.
2. Relabeling of objects does not change distance.
3. $\mathrm{d}(\sigma, \pi) = \mathrm{d}(\pi, \omega) + d(\omega, \sigma)$ ~~iff ω is "between" π and σ~~ for some ω between π and σ if π and σ disagree on more than one pair of elements.
4. ~~If two rankings agree except on a "segment," position of segment within ranking is not important: $\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe)$.~~

Kendall distance with positional constraints

Changing Kemeny's axioms

1. Distance function is a pseudo-metric.
2. Relabeling of objects does not change distance.
3. $\mathrm{d}(\sigma, \pi) = \mathrm{d}(\pi, \omega) + d(\omega, \sigma)$ ~~iff ω is "between" π and σ~~ for some ω between π and σ if π and σ disagree on more than one pair of elements.
4. ~~If two rankings agree except on a "segment," position of segment within ranking is not important: $\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe)$.~~

Theorem

The solution to the above axioms is a unique class of distances, termed the weighted Kendall distances.

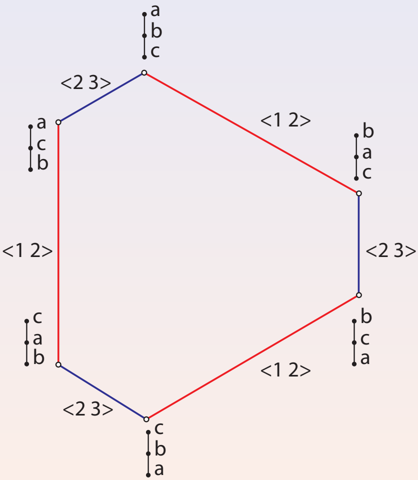
New distance family: Weighted Kendall distances

Weighted Kendall distance

minimum weight of transforming one permutation into the other using adjacent transpositions that have given non-negative weights.

Weight of transposition $\langle i\ j \rangle$ is denoted by $\varphi_{\langle i\ j \rangle}$.

$\mathrm{d}(abc, cba) = 2\varphi_{\langle 2\ 3 \rangle} + \varphi_{\langle 1\ 2 \rangle}$.



Decreasing weight functions

- For weighted Kendall distance,
  - $d(\sigma, \pi_1) = \varphi_{\langle 8\ 9\rangle}$
  - $d(\sigma, \pi_2) = \varphi_{\langle 1\ 2\rangle}$
- Choose $\varphi_{\langle i\ i+1\rangle}$ to be decreasing in i, so as to have $d(\sigma, \pi_1) < d(\sigma, \pi_2)$.

σ	π_1	π_2
Melbourne	Melbourne	Vienna
Vienna	Vienna	Melbourne
Vancouver	Vancouver	Vancouver
Toronto	Toronto	Toronto
Calgary	Calgary	Calgary
Adelaide	Adelaide	Adelaide
Sydney	Sydney	Sydney
Helsinki	Perth	Helsinki
Perth	Helsinki	Perth
Auckland	Auckland	Auckland

Weighted Kendall distance: Additional Properties

Consider the set of rankings

$$\Sigma = \left(\begin{array}{cccc} 1 & 4 & 2 & 3 \\ \hline 1 & 4 & 3 & 2 \\ \hline 2 & 3 & 1 & 4 \\ \hline 4 & 2 & 3 & 1 \\ \hline 3 & 2 & 4 & 1 \end{array} \right).$$

The Kemeny aggregate (using Kendall τ) is $(4, 2, 3, 1)$.

The optimum aggregate ranking for the weight function φ with $\varphi_{\langle i \ i+1 \rangle} = (2/3)^{i-1}, i \in [4]$, equals $(1, 4, 2, 3)$.

A candidate with both strong showings and weak showings beats a candidate with a rather average performance.

Computing the weighted Kendall distance

Computing Kendall τ is straightforward: count the number of (mutual) inversions.

Theorem

Weighted Kendall distance with decreasing weights can be computed in polynomial time $O(n^4)$.

Theorem

A 2-approximation for weighted Kendall distance with general weights can be computed in time $O(n^2)$.

Weighted Kendall distance

Aggregating with weighted Kendall: LP programming (Raisali, Farnoud, and M, 2013); Weighted bipartite matching (Farnoud and M, 2012).

Farnoud, M, Touri, A Novel Distance-Based Approach to Constrained Rank Aggregation.

Applications of weighted Kendall: Machine learning, Bioinformatics (Gene Prioritization), Rank Modulation (precise charge drop models).