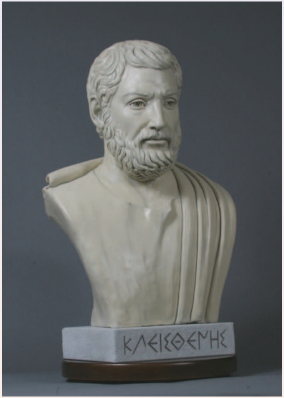
A Constrained Distance-Based Approach to Social Choice

Farzad Farnoud (Hassanzadeh)
with O. Milenkovic

10/15/12

Democracy: Greece, 2600 years ago

- Solon and Cleisthenis created the modern voting system.
- Black and white stones used to express preferences between candidates a and b: (a, b), (b, a).
- Winner declared based on simple plurality count.



Democracy: France, de Borda, 1784

- Simple plurality is "unquestionable" only for competitions between two candidates.

- Plurality voting versus pairwise majority rules

1	7	7	6	plur.	pmv
a	a	b	c	a	c
b	c	c	b	b	b
c	b	a	a	c	a

Democracy: France, de Borda, 1784

- Simple plurality is "unquestionable" only for competitions between two candidates.
- Borda's scoring method [1784]:
    - each voter ranks all candidates instead of voting for one;
    - score of vote equal to rank;
    - small total scores preferred;

a	b	a	d	a (8=1+2+1+4)
b	a	c	c	b (10=2+1+4+3)
d	c	d	b	c (11=4+3+2+2)
c	d	b	a	d (11=3+4+3+1)

Democracy: France, de Condorcet, 1785

Essay on the Application of Analysis to the Probability of Majority Decisions [1785]

Condorcet's paradox: Majority preferences may be intransitive with more than two options: majority prefers a to b, b to c, and c to a.
Example: three votes
(a, b, c),(b, c, a),(c, a, b)

Essay on the Application of Analysis to the Probability of Majority Decisions [1785]

Condorcet's jury theorem: two options to vote for, one of which is correct. Each voter has probability p of correct vote. How many independent voters are needed for correct decision via majority voting?

- $p > 1/2$, "the more the merrier".
- $p < 1/2$, best jury consists of one voter only.

Democracy: USA, Arrow, 1950's

- Arrow's impossibility theorem:
  Democracy is not possible!
- Axioms for Nobel prize:
  - No dictator
  - The preference ordering between two candidates does not depend on other candidates.
  - If all voters prefer A to B, the aggregate has to prefer A to B.
- No aggregate satisfies axioms!

It's not just political sciences

We often encounter rankings of:

- politicians, celebrities, performers, job candidates
- schools, teams in professional sports
- movies, products
- emotions, eligibility for marriage etc.

Ranking relevant in many CS/ENG applications

- Computer science (search engines, etc)
- Recommender systems, marketing
- Social sciences: competitions, voting
- Management and decision making

Rank Aggregation: Combining a set of rankings such that the result is a ranking "representative" of the set

Expert 1	**Expert 2**	**Expert 3**	**Aggregate**
GTech	UIUC	UCB	?
UIUC	UCB	UIUC	?
Stanford	GTech	MIT	?
MIT	MIT	Stanford	?
UCB	Standford	Gtech	?

Rank aggregation

Expert 1	Expert 2	Expert 3	Aggregate
			?
			?
			?

Rankings (permutations)

Mathematically, rankings are abstracted as permutations

Permutations are arrangements of a set of objects.

Example: (b, c, a) – a permutation over the set $\{a, b, c\}$

One common approach to rank aggregation is "distance-based" rank aggregation.

Given expert rankings $\sigma_1, \sigma_2, \cdots, \sigma_m$, the rank aggregation problem can be stated as

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

Related work

Rank aggregation requires a distance function over the space of permutations

Kemeny 59 Kemeny's axiomatic approach to determine appropriate distance function, use of Kendall's τ

Dwork 01 "Kemenization" is NP-hard, bipartite matching and Markov chain methods for aggregation, by Dwork et al.

Sculley 07 Aggregation with similarity score, by Sculley

Kumar 10 Generalizing Kendall's τ and Spearman's footrule, by Kumar et al.

Kemeny's axioms

Kemeny's axiomatic approach for determining a distance function:

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

Kendall's τ

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

Kendall's τ: minimum number of swaps of adjacent elements needed to transform one into the other

A swap of two elements is called a transposition
Transposition of elements in positions i and j is denoted by (ij)

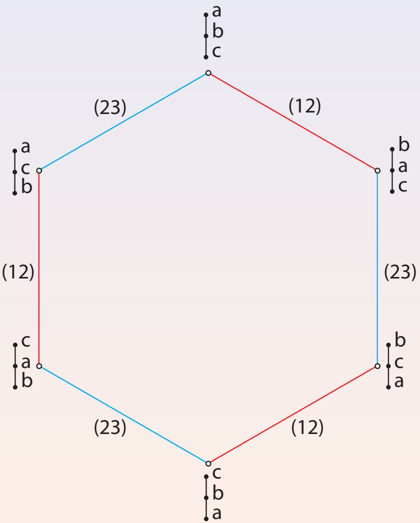
Example: $\mathrm{d}(abcde, cabde) = 2: \quad abcde \xrightarrow{(23)} acbde \xrightarrow{(12)} cabde$

Kendall's τ

Kendall's τ can be represented by a graph with $n!$ vertices.

Neighboring vertices differ by an adjacent transposition

Distance is the length of the shortest path



Rank aggregation: need for new distances

Need a new distance function that addresses shortcomings of Kendall's τ in terms of having following additional properties:

1. Top versus bottom

$$\mathrm{d}(abcd, bacd) > \mathrm{d}(abcd, abdc)$$

2. Similar items versus dissimilar items

$$\mathrm{d}(ab'ba', b'aba') > \mathrm{d}(aa'bb', a'abb')$$

Ease of calculation or approximation required!

Generalizing Kendall's distance

How should the axioms be changed?
- Let us remove the fourth axiom

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

Generalizing Kendall's distance

How should the axioms be changed?
- Let us remove the fourth axiom

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
4. ~~If two rankings agree except on a "segment," position of segment within ranking is not important:~~ $\cancel{\mathrm{d}(abcde, abdce) = \mathrm{d}(cdabe, dcabe).}$

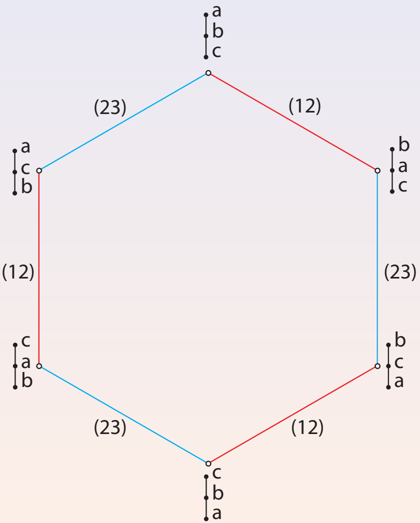
The solution is again Kendall τ!!
Removing the fourth axiom is not sufficient. Also need to modify the third axiom

Why modify the third axiom?

Lemma [F, Touri, Milenkovic]: For complete rankings, fourth axiom follows from third axiom.

Special case: $n = 3$

Consider the distinct paths between (a, b, c) and (c, b, a).



Our relaxation of Kemeny's axioms:

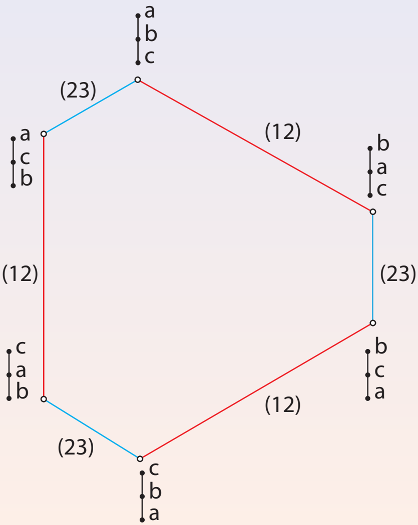
1. Distance function is a pseudo-metric
2. Relabeling of objects does not change distance.
3. $\mathrm{d}(\sigma,\pi) = \mathrm{d}(\pi,\omega) + d(\omega,\sigma)$ ~~iff ω is "between" π and σ~~ for some ω between π and σ and distinct from them if π and σ disagree on more than one pair of elements.
4.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

Unique solution: weighted (cost constrained) Kendall's τ! [F, Touri, Milenkovic, 2012]

New Distance: weighted Kendall distance

Weighted Kendall distance: minimum cost of transforming one permutation into the other using adjacent transpositions where each adjacent transposition has a given cost

Cost of transposition (ij) is denoted by $\varphi(i, j)$



Computing distance

Computing Kendall's τ is straightforward: count the number of out of order pairs.

How to compute the weighted Kendall distance for general cost function is not known, but is known for a very important case.

Monotonic cost function: φ is monotonic if $\varphi(i, i+1)$ is monotonic in i.

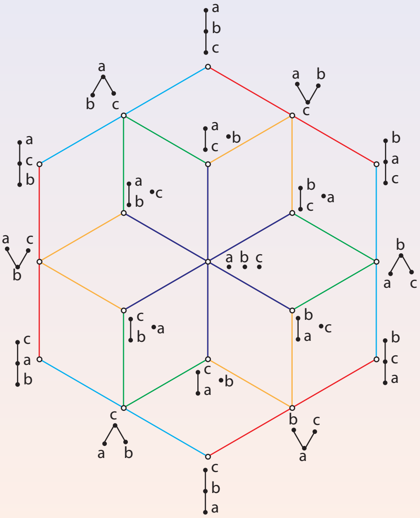
Theorem [F, Touri, Milenkovic]: Weighted Kendall distance with monotonic cost can be computed in time $O(n^4)$.

Theorem [F, Milenkovic]: 2-approximation for weighted Kendall distance with general cost can be computed in time $O(n^2)$.

Partial rankings and partially ordered sets

Bogart [1973] generalizes Kemeny's approach to partially ordered sets.

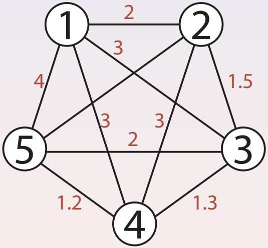
Our relaxation also generalizes to partially ordered sets (work in progress)



New Distance: weighted transposition distance

We assign cost $\varphi(i, j) \geq 0$ to any transposition (ij), φ is called a cost function

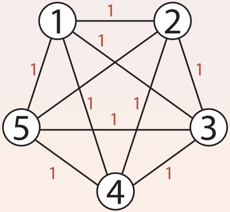
Weighted transposition distance: between two permutations π and σ is the minimum cost of transforming π into σ using transpositions $= \mathrm{d}_{\varphi}(\pi, \sigma)$



Rank aggregation: common distance functions

Several distance functions used for rank aggregation [Diaconis and Graham 88]:

- Kendall's τ: $K(\pi, \sigma) = \#$ of transpositions of adjacent ranks. Equivalent to $\varphi_K(i, i+1) = 1$.
- Spearman's Footrule: $F(\pi, \sigma) = \sum_i |\pi(i) - \sigma(i)|$. Equivalent to the path cost function $\varphi_F(i, j) = |i - j|$.
- Cayley's distance:
  $T(\pi, \sigma) = \#$ of transpositions
  Equivalent to $\varphi_T(i, j) = 1$.



Three stage algorithm for computing distance

1. Distance of a single transposition from identity: e.g. d(12345, 42315)
   - A Viterbi-style algorithm on a trellis, or
   - Bellman-Ford type algorithm over graphs
2. Distance of a single cycle from identity
   - A dynamic program finds an approximation,
   - Proof of approximation using "h-transpositions"
3. Distance of a general permutation from identity
   - By extending results for single cycle permutations

See F.Farnoud and O.Milenkovic, Sorting of permutations by weighted transpositions. IT Transaction, 58(1):3–23, Jan. 2012.

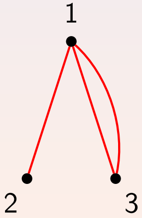
Transformations and multigraphs

A sequence of transpositions that transforms one permutation to another is called a transformation.

A transformation can be represented by a multigraph

The vertices of the multigraph are $\{1, \cdots, n\}$, while the edges are the transpositions of the transformation.

Example: Transformation $(13), (12), (13)$ transforms $\pi = (a, c, b)$ into $e = (a, b, c)$

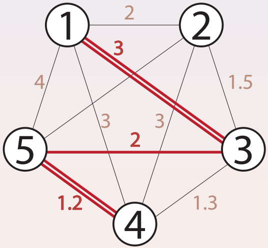


Computing distance: transpositions

Theorem: For $i, j \in [n]$, let τ be the minimum cost transformation of $\pi = (ij)$ to identity. The multigraph of τ is of the form

$$2 \times (\text{a path between } i \text{ and } j) - (\text{one edge}).$$

Example: Transformation of $\pi = 42315$



Theorem: Distance between a permutation $\pi = (ij)$ and identity can be computed in time $O(n^4)$.

A modification of the Bellman-Ford shortest path algorithm can be used.

Computing distance

- Constant approximation algorithm (const=4) for arbitrary cost functions φ in $O(n^4)$ operations
- Constant approximation algorithm (const=2), if cost function φ is a metric, in $O(n^4)$ operations
- Constant approximation algorithm (const=2) for path cost functions (e.g. weighted Kendall) in $O(n^4)$ operations
- Exact algorithms for metric-path cost functions (e.g. weighted Spearman's Footrule) in $O(n^2)$ operations.

See F.Farnoud and O.Milenkovic, Sorting of permutations by weighted transpositions. IT Transaction, 58(1):3–23, Jan. 2012.

Rank aggregation

- Weighted Kendall distance generalizes Kendall's τ to model the significance of top vs bottom of a ranking, e.g.,

$$\varphi(1,2) > \varphi(9,10)$$

- Weighted transposition distance generalizes Cayley's distance to model similarities and dissimilarities of elements, e.g.,

$$\begin{aligned}&\varphi(\text{Godfather I}, \text{Godfather II})\\ &\qquad < \varphi(\text{Godfather I}, \text{Goodfellas})\\ &\qquad\qquad < \varphi(\text{Godfather I}, \text{Star Wars})\end{aligned}$$

Rank aggregation

Recall: viven voter rankings $\sigma_1, \sigma_2, \cdots, \sigma_m$, rank aggregation solves

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

For many distance functions, problem is NP-hard.

Alternative ways to find reasonable solutions

- Approximation: 2-approximation or 4-approximation (depending on type of φ)
- Using Matching algorithms and finding local optimum instead of global optimum
- Markov chain[Dwork et al. 01] methods

Rank aggregation: approximation

For general cost function φ, to find

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

we approximate d_φ by D such that

$$(1/2)D(\pi, \sigma) \le \mathrm{d}_\varphi(\pi, \sigma) \le 2D(\pi, \sigma).$$

We can find

$$\pi' = \arg\min_{\pi} \sum_{i=1}^{m} D(\pi, \sigma_i)$$

and can show that

$$\sum_{i=1}^{m} \mathrm{d}_\varphi(\pi', \sigma_i) \le 4 \sum_{i=1}^{m} \mathrm{d}_\varphi(\pi^*, \sigma_i).$$

Rank aggregation

Students were asked to rank the following items in order of importance in their academic life:

1. Campus friendliness and inclusiveness
2. Availability of recreational and cultural facilities
3. Quality of classrooms and dorms
4. Extracurricular student groups and activities
5. Geographical proximity to your family/boyfriend/girlfriend
6. Commitment of campus to build a diverse community
7. Being able to express one's personal identity freely
8. Being able to make friends on campus
9. Safety and security
10. Availability of financial support/scholarship
11. Availability of personal counseling/academic tutoring
12. Friendliness/academic prowess of faculty members/instructors

We used the cost function $\varphi(i, i+1) = (3/4)^{i-1}$.

Rank aggregation

	Aggregate Ranking
Graduate (28)	10, 12, 9, 8, 1, 3, 2, 11, 7, 5, 4, 6
Undergrad (73)	12, 9, 8, 1, 3, 10, 4, 2, 11, 7, 5, 6

1. Campus friendliness and inclusiveness
2. Availability of recreational and cultural facilities
3. Quality of classrooms and dorms
4. Extracurricular student groups and activities
5. Geographical proximity to your family/boyfriend/girlfriend
6. Commitment of campus to build a diverse community
7. Being able to express one's personal identity freely
8. Being able to make friends on campus
9. Safety and security
10. Availability of financial support/scholarship
11. Availability of personal counseling/academic tutoring
12. Friendliness/academic prowess of faculty members/instructors

Rank aggregation

	Aggregate Ranking
Undergrad/Female (32)	12, 9, 1, 8, 3, 4, 2, 5, 10, 11, 7, 6
Undergrad/Male (31)	12, 9, 3, 1, 8, 10, 4, 2, 11, 7, 5, 6
Undergrad/? (10)	8, 12, 4, 1, 3, 9, 7, 2, 10, 11, 6, 5

1. Campus friendliness and inclusiveness
2. Availability of recreational and cultural facilities
3. Quality of classrooms and dorms
4. Extracurricular student groups and activities
5. Geographical proximity to your family/boyfriend/girlfriend
6. Commitment of campus to build a diverse community
7. Being able to express one's personal identity freely
8. Being able to make friends on campus
9. Safety and security
10. Availability of financial support/scholarship
11. Availability of personal counseling/academic tutoring
12. Friendliness/academic prowess of faculty members/instructors

Distributed rank aggregation

Borda's aggregation over networks:
Network modeled by a connected graph $G = ([m], E)$.
Voters are vertices and edges indicate connectivity.
$b_i(t)$ is the vector of Borda scores according to voter i at time t.

1. At time $t \geq 0$, pick an edge $\{i, i'\} \in E$ with probability $P_{ii'} > 0$, where $\sum_{\{i,i'\}\in E} P_{ii'} = 1$,
2. i, i' exchange their estimate $b_i(t), b_{i'}(t)$ and they let $b_i(t+1) = b_{i'}(t+1) = \frac{1}{2}(b_i(t) + b_{i'}(t))$,
3. voters $\ell \neq i, i'$ let $b_\ell(t+1) = b_\ell(t)$.

Well-known: Almost surely $\lim_{t\to\infty} b_i(t) = \bar{b}$ where $\bar{b} = \frac{1}{m}\sum_{i=1}^{m} \bar{b}_i(0)$ [Boyd et al.].

Distributed rank aggregation

Definition: t is a consensus time if the ordering of $b_i(t)$ matches ordering of $\bar{b}$, for all $i \in [m]$.

$$T = \min\{t \geq 0 \mid t \text{ is a consensus time for the ordering.}\}$$

$r^j = \min(\bar{b}^{j+1} - \bar{b}^j, \bar{b}^j - \bar{b}^{j-1})$: minimum distance of average rating of j from neighboring candidates.

$d^j = \max_i b_i^j(0) - \min_i b_i^j(0)$: spread of the initial ratings of the object j among voters.

$$P(T > t) \leq 4m\lambda_2^t(W) \sum_{j=1}^{n} \left(\frac{d^j}{r^j}\right)^2,$$

$W = \sum_{\{i,j\}\in E} P_{ij}(I - \frac{1}{2}(e_i - e_j)(e_i - e_j)^T)$, $\lambda_2(W) =$ second largest eigenvalue of W.

Thank you!