Computing Kemeny Rankings From d-Euclidean Preferences

Thekla Hamm Martin Lackner Anna Rapberger

TU WIEN · logics LOGICAL METHODS IN COMPUTER SCIENCE · ac ALGORITHMS AND COMPLEXITY GROUP · dbai Databases and Artificial Intelligence Group

Kemeny Ranking

An optimal Kemeny ranking $\succ$ minimizes $\sum_{v \in \mathcal{V}} \mathsf{K}(\succ, \succ_v)$ (Kemeny score), where $\mathsf{K}(\succ, \succ') = |\{\{x,y\} \subseteq \mathcal{C} \mid (x \succ y \wedge y \succ' x) \vee (y \succ x \wedge x \succ' y)\}|$ measures pairwise disagreements (Kendall-Tau distance).

- Computing an optimal Kemeny ranking is NP-hard [1].

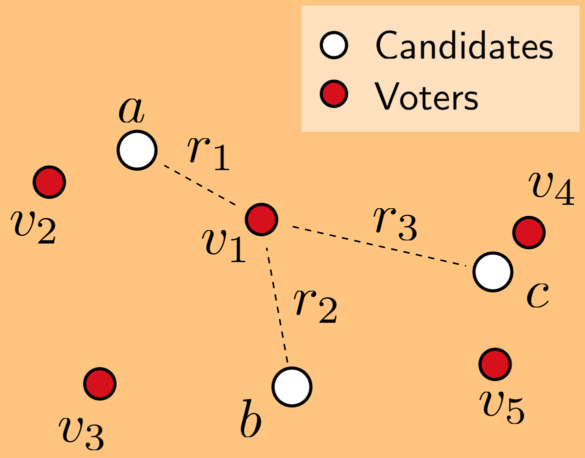
d-Euclidean Elections

Given $p : \mathcal{C} \cup \mathcal{V} \to \mathbb{R}^d$, an order $\succ_v$ is p-embeddable if for all $c, c' \in \mathcal{C}$, $c \succ c'$ if and only if $\|p(v) - p(c)\|_d < \|p(v) - p(c')\|_d$.
An election is d-Euclidean if it it is p-embeddable for some p.

- Recognizing d-Euclidean elections is NP-hard [2].

Example

- Preferences of voter v_1: $a \succ_1 b \succ_1 c$ because $r_1 < r_2 < r_3$
- Optimal Kemeny ranking: $b \succ a \succ c$ Kendall-Tau distance $\mathsf{K}(\succ, v_1) = 1$

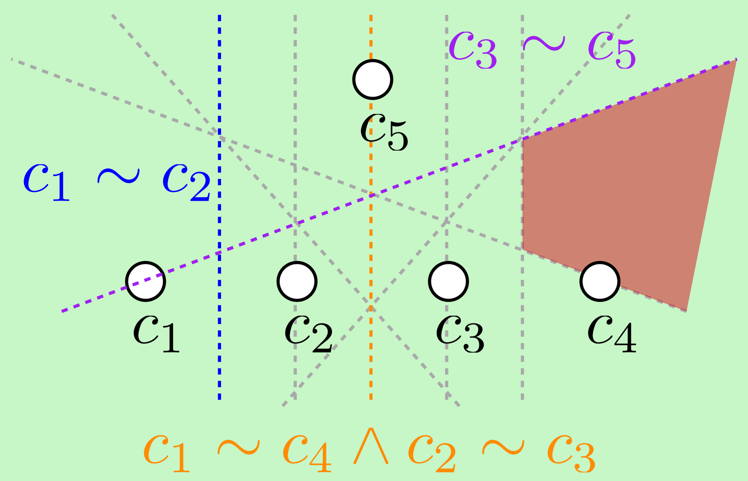


Theorem:

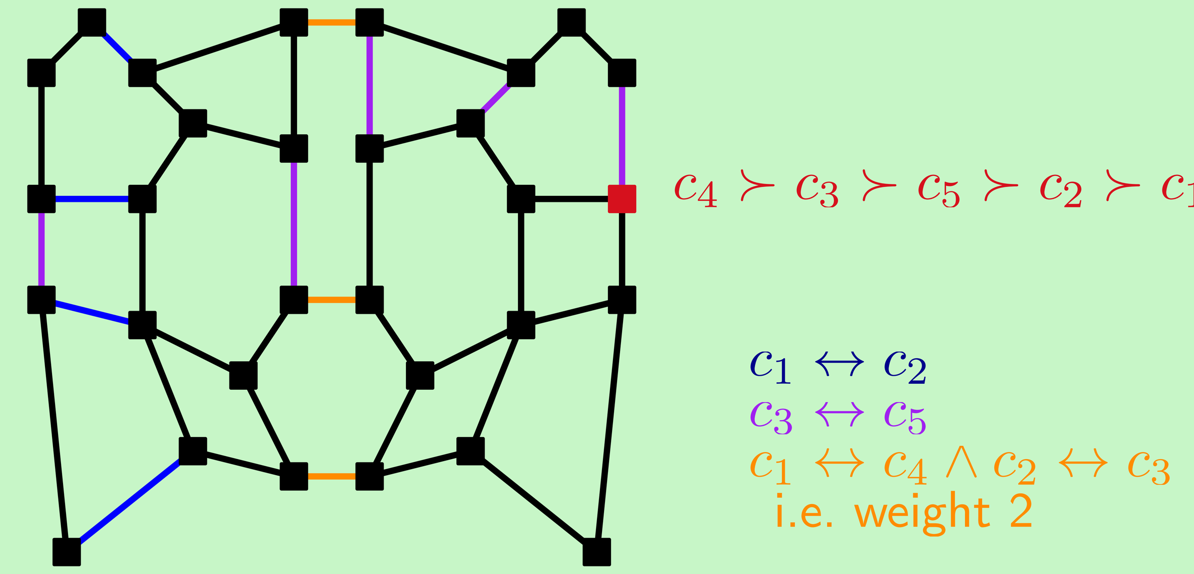
Determining all p-embeddable Kemeny rankings for a d-Euclidean election $(\mathcal{C}, \mathcal{V}, (\succ_v)_{v \in \mathcal{V}})$ given by $p : \mathcal{C} \cup \mathcal{V} \to \mathbb{R}^d$ is possible in time in $\mathcal{O}(|\mathcal{C}|^{4d})$.

[for weak voter preferences idea can be adapted and works in $\tilde{\mathcal{O}}(|\mathcal{C}|^{2(d\omega+1)})$, where ω is the matrix multiplication constant]

Idea: Embedding of $\mathcal{C}$ implies hyperplane arrangement



G_{pref}: vertices for p-emb. rankings; weighted edges for candidate-swaps between 'adjacent' rankings



→ Distances in G_{pref} $\hat{=}$ #disagreements between p-embed. rankings

→ want to find vertex minimising sum of distances to vertices with voter-rankings (scaled by resp. #voters)

Algorithm:

1. Compute G_{pref} $\mathcal{O}(|\mathcal{C}|^{2d})$
2. Compute scaled center of G_{pref} $\mathcal{O}(|\mathcal{C}|^{4d})$

p-Embeddable Kemeny Ranking

Input: Candidates $\mathcal{C}$, voters $\mathcal{V}$, embedding $p : \mathcal{C} \cup \mathcal{V} \to \mathbb{R}^d$, implied preferences $c \succ_v c'$ iff $\|p(v) - p(c)\| < \|p(v) - p(c')\|$

Output: p-embeddable Kemeny ranking $\succ$ such that
(1) $\exists x \in \mathbb{R}^d$ with $c \succ c'$ iff $\|x - c\| < \|x - c'\|$, and
(2) with this property $\succ$ minimises the Kemeny score, i.e. $\sum_{v \in \mathcal{V}} |\{(c, c') \in \mathcal{C}^2 \mid (c' \succ c \wedge c \succeq_v c')\}|$

Intuitively this captures a ranking that is (1) consistent with the embedding and (2) among these reflects the general consensus.

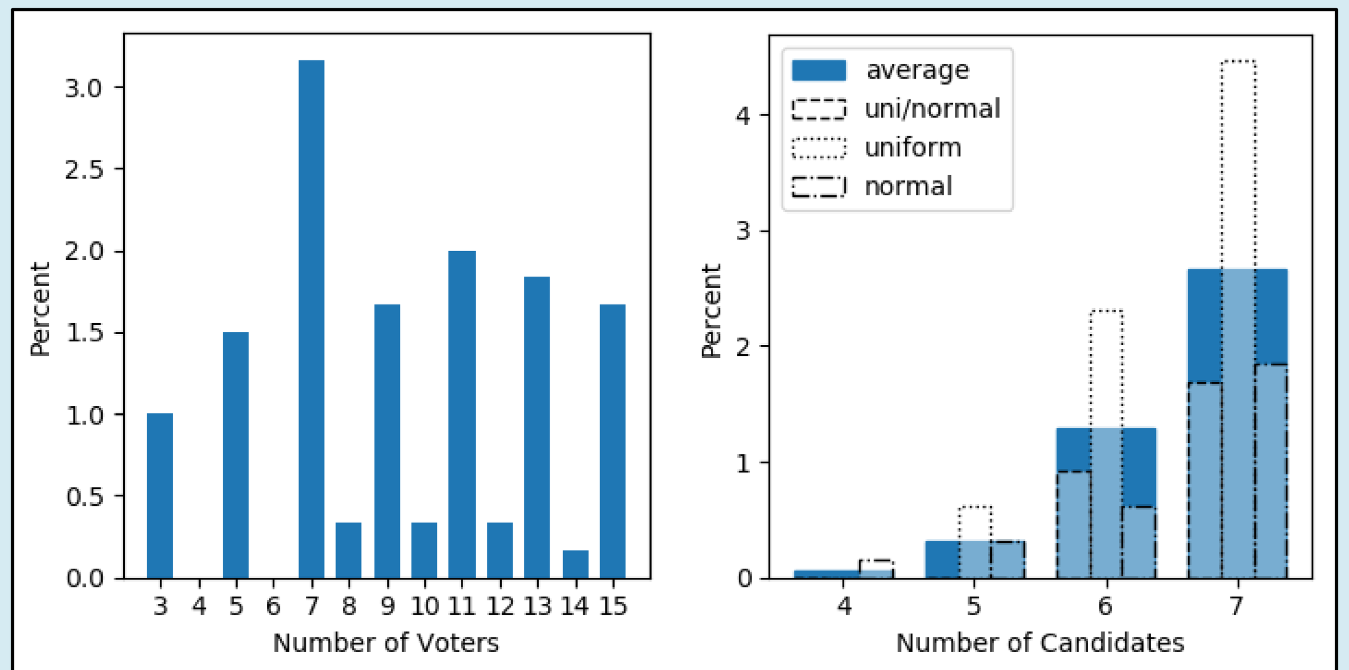
Approximation and experimental Results

- A p-embeddable Kemeny ranking 2-approximates any optimal Kemeny ranking (theoretical worst case bound).
- Known largest ratio so far is $8/7$ - cf. Example →

Experiments

Computation and comparison of the ratio r between the p-embeddable and the optimal Kemeny score

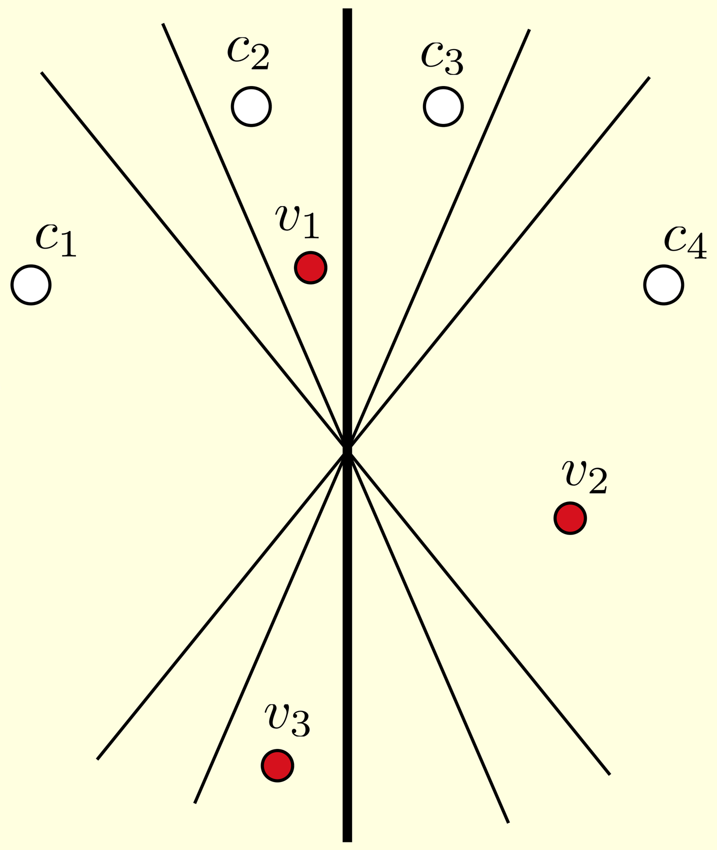
Instances: Randomly generated 2-Euclidean elections with n voters, $3 \leq n \leq 15$, m candidates, $4 \leq m \leq 7$



Percentage of instances with ratio $r > 1$

Example

- *Preferences of voter:*
  v_1: $c_2 \succ_1 c_3 \succ_1 c_1 \succ_1 c_4$
  v_2: $c_4 \succ_2 c_3 \succ_2 c_2 \succ_2 c_1$
  v_3: $c_1 \succ_3 c_4 \succ_3 c_2 \succ_3 c_3$
- *Unique optimal Kemeny ranking*
  κ_{opt}: $c_4 \succ c_2 \succ c_3 \succ c_1$
  - Kemeny score $= 14$
  - $\succ$ is not embeddable!
- *p-embeddable Kemeny rankings:* v_1, v_2, and v_3 (Score $= 16$)



Experimental Results:

The d-embeddable Kemeny ranking is optimal in $\approx 98.9\%$ - only 84 out of 7800 instances yield suboptimal results. The ratio r lies between 1.0077 and 1.11 ($< \frac{8}{7}$).

Further Questions:

- Characterisation of instances for which Kemeny and p-embeddable Kemeny ranking coincide?
- What are examples in which the ratio between Kemeny and p-embeddable Kemeny score is $> \frac{8}{7}$?
- Can we decrease the dependency on d in an algorithm for p-Embeddable Kemeny Ranking?
- What is the complexity of Kemeny Ranking for d-Euclidean preferences?

[1] Bartholdi, J., Tovey, C.A., Trick, M.A.: Voting schemes for which it can be difficult to tell who won the election. In *Social Choice and Welfare* 6(2),157–165 (Apr 1989)
[2] Peters, D.: Recognising multidimensional euclidean preferences. In *Proc. AAAI.* pp. 642–648 (2017)