The Squared Kemeny Rule for Averaging Rankings

Patrick Lederer\textsuperscript{1} Dominik Peters\textsuperscript{2} Tomasz Was\textsuperscript{2,3}

\textsuperscript{1} TU Munich / UNSW\textsuperscript{2} CNRS, LAMSADE, Université Paris Dauphine - PSL\textsuperscript{3} University of Oxford

2024-07-10
ACM EC 2024
Input: a profile of (strict) rankings $R_1, \ldots, R_n$, possibly with weights $w_1, \ldots, w_n$. Output: a ranking $R$ (or several rankings if there is a tie)

Well-known: **Kemeny rule** [1959] – choose $R$ to minimize

$$ \sum_{i \in N} w_i \cdot d(R, R_i) $$

where $d(R, R_i)$ is the **Kendall-tau distance** $|\{(a, b) : aRb \text{ and } bR_i a\}|$.

---

Application: ranking hotels

Sort by: Price (lowest first)

- Top picks for solo travellers
- Homes & apartments first
- Price (lowest first)
- Best reviewed and lowest price
- Property rating (high to low)
- Property rating (low to high)
- Property rating and price
- Distance from city centre
- Top reviewed
- Genius discounts first
## Application: ranking hotels

<table>
<thead>
<tr>
<th>Price</th>
<th>90%</th>
<th>80%</th>
<th>70%</th>
<th>60%</th>
<th>50%</th>
<th>40%</th>
<th>30%</th>
<th>20%</th>
<th>10%</th>
<th>Score</th>
</tr>
</thead>
</table>
## Application: ranking hotels with Kemeny

<table>
<thead>
<tr>
<th>Price</th>
<th>90%</th>
<th>80%</th>
<th>70%</th>
<th>60%</th>
<th>50%</th>
<th>40%</th>
<th>30%</th>
<th>20%</th>
<th>10%</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>tied</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
</tr>
<tr>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>tied</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
</tr>
<tr>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>tied</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
</tr>
<tr>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>tied</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
</tr>
<tr>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>tied</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
</tr>
<tr>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>tied</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
</tr>
</tbody>
</table>

The table above shows the ranking of different hotels based on their price and quality. The Kemeny algorithm is used to determine the ranking, where Price is the ranking criterion. The hotels are ranked from highest to lowest price with 90% being the highest. The hotels that are tied share the same rank.
The Squared Kemeny Rule

Less well-known: Kemeny [1959] in the same article also introduced the “mean rule” which we call Squared Kemeny – choose $R$ to minimize

$$\sum_{i \in N} w_i \cdot d(R, R_i)^2$$

where $d(R, R_i)$ is the Kendall-tau distance $|\{(a, b) : aRb \text{ and } bR_i a\}|$.

## Application: ranking hotels with Squared Kemeny

<table>
<thead>
<tr>
<th>Price</th>
<th>90%</th>
<th>80%</th>
<th>70%</th>
<th>60%</th>
<th>50%</th>
<th>40%</th>
<th>30%</th>
<th>20%</th>
<th>10%</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Quinta</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
<td>Graduate</td>
</tr>
<tr>
<td>Graduate</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>The Study</td>
</tr>
<tr>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
<td>La Quinta</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>NHH</td>
<td>The Study</td>
<td>The Study</td>
<td>H. Marcel</td>
</tr>
<tr>
<td>Omni</td>
<td>Omni</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>La Quinta</td>
<td>The Study</td>
<td>The Study</td>
<td>NHH</td>
<td>NHH</td>
<td>NHH</td>
</tr>
<tr>
<td>H. Marcel</td>
<td>H. Marcel</td>
<td>Omni</td>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>La Quinta</td>
<td>Omni</td>
<td>Omni</td>
</tr>
<tr>
<td>The Study</td>
<td>The Study</td>
<td>The Study</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>Omni</td>
<td>La Quinta</td>
<td>La Quinta</td>
</tr>
</tbody>
</table>
Other applications

• lists of products in e-commerce (ranking by cost, rating, delivery time, etc.)
• newsfeeds of social networks (“for you” versus “following”)
• university rankings (student satisfaction, % of students employed after graduating, research output)
• voting: hiring committee needing to rank applicants (if you have too many ML colleagues, under Kemeny you might only ever hire more ML colleagues)
• groups of friends wanting to produce rankings of favorite music, restaurants, or travel destinations (majoritarian methods are weird in that context)
A rule $f$ satisfies 2-Rankings-Proportionality (2RP) if for all profiles $R$

- consisting of just two rankings $\succ_1$ and $\succ_2$
- that disagree on $d = \text{swap}(\succ_1, \succ_2)$ pairwise comparisons,
- that have weights $w_1$ and $w_2$ with $w_1 + w_2 = 1$,

we have

$$f(R) = \{ \triangleright \in \mathcal{R} : d - \text{swap}(\succ_i, \triangleright) \in \text{round}(w_i \cdot d) \text{ for } i \in \{1, 2\} \},$$

where $\text{round}(z)$ denotes the set of closest integers to $z$.

**Example**

Rankings $\succ_1$ and $\succ_2$ have weight 70% and 30% respectively, and they disagree on 10 pairwise comparisons. Then 2RP says the outcome should agree with $\succ_1$ on 7 of these comparisons, and with $\succ_2$ on the other comparisons.
<table>
<thead>
<tr>
<th>Theorem (Young–Levenglick, 1979)</th>
<th>Theorem (this paper)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>The Kemeny rule is the only rule that satisfies</em></td>
<td></td>
</tr>
<tr>
<td>• anonymity and neutrality</td>
<td></td>
</tr>
<tr>
<td>• reinforcement</td>
<td></td>
</tr>
<tr>
<td>• “Condorcet” property</td>
<td></td>
</tr>
<tr>
<td><em>The Squared Kemeny rule is the only rule that satisfies</em></td>
<td></td>
</tr>
<tr>
<td>• anonymity and neutrality</td>
<td></td>
</tr>
<tr>
<td>• reinforcement</td>
<td></td>
</tr>
<tr>
<td>• continuity</td>
<td></td>
</tr>
<tr>
<td>• 2-ranking proportionality</td>
<td></td>
</tr>
</tbody>
</table>

Axiomatic Comparison

Kemeny rule

- ✓ Pareto efficiency
- ✓ participation
- ✓ strategyproofness (weak form)
- • behaves like median on single-crossing profiles

Squared Kemeny rule

- ✓ Pareto efficiency
- ✓ participation
- ✗ strategyproofness
- • behaves like mean on single-crossing profiles
More than 2 rankings

Figure 1: The simplex of profiles in which the rankings $\succ_1 = abcdefgh$, $\succ_2 = fedcbahg$, and $\succ_3 = bahgfdec$ occur. Each point of the simplex is colored according to the swap distance of the (a) Kemeny and (b) Squared Kemeny ranking to the input rankings.
Comparing the Rules

Figure 2: The maximum swap distance (normalized to $[0, 1]$) between the output of the Kemeny or Squared Kemeny rule to an input ranking, as a function of the weight $\alpha$ of the input ranking, for $m = 6$ alternatives.
Comparing the Rules

Figure 2: The maximum swap distance (normalized to $[0, 1]$) between the output of the Kemeny or Squared Kemeny rule to an input ranking, as a function of the weight $\alpha$ of the input ranking, for $m = 6$ alternatives.
Comparing the Rules

**Theorem**

If \( \succ^* \in \mathcal{R} \) is an input ranking with weight \( \alpha \), and \( \triangleright \) is a Squared Kemeny output ranking, then

\[
\text{swap}(\succ^*, \triangleright) \leq \sqrt{\frac{1 - \alpha}{\alpha}} \cdot \binom{m}{2}.
\]

In the paper: We also consider groups of voters who may have different rankings and prove that the SqK ranking cannot be too far away from them on average.
Computational Complexity

- Squared Kemeny is NP-complete to compute, even for $n = 4$ rankings. Similar reduction as for egalitarian Kemeny.


- It can be computed using an ILP using the same trick as for maximizing Nash welfare in the allocation of indivisible goods.


- The Kemeny ranking provides a 2-approximation to Squared Kemeny.

- For every constant $\varepsilon > 0$, there exists a polynomial-time $(2 + \varepsilon)$-approximation to the Squared Kemeny rule.
Figure 3: The median running time of computing the Squared Kemeny using Gurobi for a given number of alternatives and \( n = \{3, 4, 5\} \) rankings occurring in the profile with equal weights, drawn uniformly at random. The values are based on 50 samples.
Figure 4: Euclidean embeddings. For the Kemeny rule, the positions of its outputs are denoted with a red diamond, and for Squared Kemeny with a green square.
• Draw “$\alpha$-curves” for other social choice models.

• Figure out if there is a difference between “averaging” and “proportional aggregation”, philosophically or formally.

• What about other distances such as the Spearman footrule distance?

• What about other $p$-Kemenys?

• What about using other types of rules? PAV? Multiple binary issues with a transitivity constraint?

• Can one make an epistemic argument for Squared Kemeny, e.g. as the MLE of a “normal” Mallows model?