Euclidean Voting

July 2, 2013

1 Preliminaries

Let A denote the set of m alternatives, and $\mathcal{L}(A)$ and $\mathcal{O}$ denote the space of rankings of alternatives and outcomes, respectively. Thus, $|\mathcal{L}(A)| = m!$, and let $k = |\mathcal{O}|$. A profile $\pi \in \mathcal{L}(A)^n$ is a collection of votes (rankings). For any profile π, let $n(\pi, \sigma)$ denote the number of times σ appears in π. Let $\sigma_1, \ldots, \sigma_{m!}$ denote a fixed reference order of the rankings in $\mathcal{L}(A)$.

For any two profiles π_1 and π_2, let $\pi_1 + \pi_2$ be the union profile such that $n(\pi_1 + \pi_2, \sigma) = n(\pi_1, \sigma) + n(\pi_2, \sigma)$ for every $\sigma \in \mathcal{L}(A)$. Similarly, for any profile π, let $c\pi$ be the profile such that $n(c\pi, \sigma) = c \cdot n(\pi, \sigma)$ for every $\sigma \in \mathcal{L}(A)$.

Voting Rule A voting rule (more technically, a social welfare function - SWF) $r : \mathcal{L}(A)^n \to \mathcal{P}(\mathcal{L}(A))$ is a function that maps every profile of votes to a set of tied rankings.

Anonymity A voting rule r is called *anonymous* if it only depends on the number of times each ranking appears in the profile: for every profiles π_1 and π_2 such that $n(\pi_1, \sigma) = n(\pi_2, \sigma)$ for every $\sigma \in \mathcal{L}(A)$, $r(\pi_1) = r(\pi_2)$.

Neutrality Given any profile $\pi = (\sigma_1, \ldots, \sigma_n)$, let $\tau\pi = (\tau\sigma_1, \ldots, \tau\sigma_n)$ be the profile where each vote is permuted according to τ. Similarly, given any set of rankings S, let $\tau S = \{\tau\sigma | \sigma \in S\}$. A voting rule r is called *neutral* if for every profile π and permutation τ, we have $r(\tau\pi) = \tau r(\pi)$.

Consistency A social welfare function r is called *consistent* (for rankings) if for every profiles π_1 and π_2 such that $r(\pi_1) \cap r(\pi_2) \neq \emptyset$, $r(\pi_1 + \pi_2) = r(\pi_1) \cap r(\pi_2)$.

Connectedness A voting rule r is called *connected* if for any two profiles π_1 and π_2 with $r(\pi_1) \cap r(\pi_2) \neq \emptyset$, there exist non-negative integers c and d such that $r(\pi_1) \cap r(c\pi_1 + d\pi_2) \neq \emptyset$ and $r(\pi_1) \neq r(c\pi_1 + d\pi_2)$.

Continuity Two profiles π_1 and π_2 satisfy $\pi_1 \approx \pi_2$ if they differ by one vote: for some σ and σ', $n(\pi_1, \sigma) = n(\pi_2, \sigma) - 1$, $n(\pi_1, \sigma') = n(\pi_2, \sigma') + 1$, and for every $\sigma \in \mathcal{L}(A) \setminus \{\sigma, \sigma'\}$, $n(\pi_1, \sigma) = n(\pi_2, \sigma)$. A voting rule r is called *continuous* if for every profile π and ranking σ, $\sigma \notin r(\pi)$ implies that there exists integer k such that for every profile $\pi' \approx k\pi$, $\sigma \notin r(\pi')$.

2 Background on Mean Proximity Rules and Generalized Scoring Rules

Mean Proximity Rules (Zwicker [4]) A voting rule is called a mean proximity rule if there exists an input embedding $\phi : \mathcal{L}(A) \to \mathbb{R}^k$ and an output embedding $\psi : \mathcal{O} \to \mathbb{R}^k$ such that for any profile π with n votes, $r(\pi) = \arg\min_{o \in \mathcal{O}} \|\psi(o) - mean(\pi)\|$, where $mean(\pi) = \sum_{\sigma \in \mathcal{L}(A)} (n(\pi, \sigma)/n) \cdot \phi(\sigma)$ is the mean of the input embeddings of the votes in π (along with multiplicity).

Generalized Scoring Rules (Zwicker [4]) A voting rule is called a generalized scoring rule if there exists a scoring function $s : \mathcal{L}(A) \times \mathcal{O} \to \mathbb{R}$ such that for any profile π, $r(\pi) = \arg\max_{o \in \mathcal{O}} \sum_{\sigma \in \mathcal{L}(A)} n(\pi, \sigma) \cdot s(\sigma, o)$.

For these two classes of voting rules, we have the following elegant equivalence theorem by Zwicker [4]. We reconstruct the proof since it is the foundation of several of our results.

Proposition 1 (Theorem 4.2.1, Zwicker [4]). *A voting rule is a mean proximity rule if and only if it is a generalized scoring rule.*

Proof (reconstructed). Take any mean proximity rule r. Let ϕ and ψ be any input and output embeddings that generate r. Now, for any profile π,

$$
\begin{aligned}
r(\pi) &= \arg\min_{o \in \mathcal{O}} \|\psi(o) - mean(\pi)\| \\
&= \arg\min_{o \in \mathcal{O}} \|\psi(o) - mean(\pi)\|^2 \\
&= \arg\min_{o \in \mathcal{O}} \sum_{\sigma \in \mathcal{L}(A)} n(\pi, \sigma) \cdot \|\psi(o) - \phi(\sigma)\|^2 \qquad (1) \\
&= \arg\max_{o \in \mathcal{O}} \sum_{\sigma \in \mathcal{L}(A)} n(\pi, \sigma) \cdot \left(-\|\psi(o) - \phi(\sigma)\|^2 \right),
\end{aligned}
$$

where Equation (1) follows by the well-known fact that the discrete mean, the member of a finite set that minimizes the sum of squares of Euclidean distances from given points, is the point in the finite set that is closest to the Euclidean mean (coordinate wise average) of the given points. This is given as Corollary 4.2.3 in [4]. Now, we can see that r is a generalized scoring rule by taking the score function $s(\sigma, o) = -\|\psi(o) - \phi(\sigma)\|^2$.

For the other direction, take any generalized scoring rule r and let s be any score function that generates r. Then, let $\phi(\sigma) = (s(\sigma, o_1), \ldots, s(\sigma, o_k))$ where $\{o_1, \ldots, o_k\}$ is some fixed enumeration of $\mathcal{O}$. Further, let $\psi(o_i) = e_i \in \mathbb{R}^k$ where

the ith coordinate is 1 and the rest are 0. Then, for any profile π, we have

$$o_i \in r(\pi) \Leftrightarrow \sum_{\sigma \in \mathcal{L}(A)} n(\pi, \sigma) \cdot s(\sigma, o_i) \geq \sum_{\sigma \in \mathcal{L}(A)} n(\pi, \sigma) \cdot s(\sigma, o), \forall o \in \mathcal{O}$$

$$\Leftrightarrow \langle mean(\pi), e_i \rangle \geq \langle mean(\pi), e_j \rangle, \forall 1 \leq j \leq k$$

$$\Leftrightarrow \|mean(\pi) - e_i\| \geq \|mean(\pi) - e_j\|, \forall 1 \leq j \leq k$$

$$\Leftrightarrow \|mean(\pi) - \psi(o_i)\|^2 \geq \|mean(\pi) - \psi(o_j)\|^2, \forall 1 \leq j \leq k,$$

$$\Leftrightarrow \|mean(\pi) - \psi(o_i)\| \geq \|mean(\pi) - \psi(o_j)\|, \forall 1 \leq j \leq k,$$

where the third transition follows since $\|e_j\| = 1$ for all j (and thus, $\|mean(\pi) - \psi(o_j)\|^2 - \langle mean(\pi), e_j \rangle$ is constant for all j). Thus, r is a mean proximity rule. $\qquad\square$

Further, Zwicker [3] shows that the set of voting rules that are *consistent* and *connected* is identical to the set of mean neat voting rules, which is a generalization of mean proximity rules. As a simple corollary, we have:

Proposition 2. *Any mean proximity rule (equivalently by Proposition 1, any generalized scoring rule) is consistent and connected.*

This implies that any voting rule that is not consistent (in the SWF sense) is not a mean proximity rule. In fact, we have the following.

Lemma 1 (Proposition 1,2,5 and Theorem 3 of Conitzer et. al. [1])**.** *All positional scoring rules and the Kemeny rule are generalized scoring rules (i.e., mean proximity rules). However, Bucklin's rule, Copeland's rule, the maximin rule, the ranked pairs method, and STV are not generalized scoring rules since they do not satisfy consistency (under any tie-breaking scheme).*

3 Symmetric Mean Proximity Rules

In this paper, we are interested in social welfare functions (SWFs) that return a ranking, so $\mathcal{O} = \mathcal{L}(A)$. From here onwards, unless mentioned otherwise, any voting rule we mention will be a social welfare function. In this case, the scoring function $s : \mathcal{L}(A) \times \mathcal{L}(A) \to \mathbb{R}$ describes the *similarity* between two rankings. This special case was also defined by Conitzer et. al. [1], and was named *simple ranking scoring functions* (SRSFs).

Score Matrix For any SWF r that is a generalized scoring rule (and hence SRSF), let s be any score function that generates r. Let $\sigma_1, \ldots, \sigma_{m!}$ be any fixed enumeration of $\mathcal{L}(A)$. Then, the *score matrix* of r corresponding to the score function s, denoted S, is given by $S_{ij} = s(\sigma_i, \sigma_j)$, for $1 \leq i \leq m!$, $1 \leq j \leq m!$.

We use r_i to denote the ith row of the score matrix S. Now, given any profile π of n votes, we create the vector $y_\pi = [n(\pi, \sigma_1)/n, \ldots, n(\pi, \sigma_{m!})/n]^T$. It is easy to verify that the output of the rule would be σ_k such that $k = \arg\max_i (S \cdot y_\pi)_i = \arg\max_i r_i \cdot y_\pi$.

Since for SWFs the outcome space is identical to the input space, it is tempting to take the output embedding identical to the input embedding, i.e., $\psi = \phi$. Indeed, the well known mean proximity rules such as all positional scoring rules (when interpreted as SWFs) and the Kemeny rule can be achieved with $\psi = \phi$. We call such rules symmetric mean proximity rules.

Symmetric Mean Proximity Rules A voting rule is called symmetric mean proximity rule if there exists a mean proximity representation of the rule where the input and the output embeddings are identical, i.e., $\psi = \phi$.

We now show a characterization of symmetric mean proximity rules along the lines of Proposition 1.

Theorem 1. *A voting rule is a symmetric mean proximity rule if and only if it is a generalized scoring rule for which there exists a corresponding score matrix that is a negative of a Euclidean distance matrix.*

Proof. This follows simply by observing that the bijection between mean proximity rules and generalized scoring rules used by Zwicker [4] takes negation of square of the Euclidean distance between the embeddings of two rankings as the score between the two rankings. □

4 Neutrality in Symmetric Mean Proximity Rules and Linear Embeddings

Conitzer et. al. [1] showed that neutrality of a GSR is equivalent to neutrality of its scoring function.

Proposition 3 (Lemma 2, Conitzer et. al. [1])**.** *A generalized scoring rule is neutral if and only if it corresponds to a neutral scoring function $s : \mathcal{L}(A) \times \mathcal{L}(A) \to \mathbb{R}$ where $s(\tau \circ \sigma, \tau \circ \sigma') = s(\sigma, \sigma')$ for every $\tau, \sigma, \sigma' \in \mathcal{L}(A)$ and $\circ$ is the composition when rankings are seen as bijective functions, or equivalently the product operator of the symmetric group.*

Further, it is easy to show the following.

Lemma 2. *Any neutral voting rule has all $m!$ rankings tied on the profile where each ranking appears once.*

Proof. To see this, observe that this profile does not change by applying any (but the same) permutation to all of its ranking. If the set of rankings returned by the neutral rule is non-empty and not the set of all rankings, then it would be easy to find a permutation which changes this output set. □

Now, let $\phi_{avg} = (1/m!) \cdot \sum_{\sigma \in \mathcal{L}(A)} \phi(\sigma)$ denote the average of all the embeddings. Then, Lemma 2 indicates that $\|\phi(\sigma) - \phi_{avg}\|$ must be independent of σ. Further, note that translating and scaling all embeddings in the Euclidean space

does not change the voting rule. We can first apply the translation which brings ϕ_{avg} to the origin. This makes $\|\phi(\sigma)\|$ constant for all σ. Next, we can scale the embeddings such that $\|\phi(\sigma)\| = 1$ for all σ. Thus, we have the following.

Lemma 3. *For any neutral symmetric mean proximity rule, there exists a corresponding embedding ϕ such that $\|\phi(\sigma)\| = 1$ for every $\sigma \in \mathcal{L}(A)$, and $\phi_{avg} = (1/m!) \cdot \sum_{\sigma \in \mathcal{L}(A)} \phi(\sigma) = 0$.*

This shows that embeddings that map rankings to the unit sphere are sufficient to generate almost all mean proximity rules of interest. For this specific condition, we have the following.

Lemma 4. *If $\|\phi(\sigma)\| = 1$ for all $\sigma \in \mathcal{L}(A)$, then*

$$\underset{\sigma \in \mathcal{L}(A)}{\arg\min} \|\phi(\sigma) - mean(\pi)\| = \underset{\sigma}{\arg\max} \langle \phi(\sigma), mean(\pi) \rangle.$$

Using this, we give a refinement of Theorem 1.

Lemma 5. *A voting rule is a symmetric mean proximity rule for which some corresponding embedding ϕ maps all rankings to points on the unit sphere if and only if the rule is a generalized scoring rule for which there exists a corresponding positive semidefinite score matrix.*

Proof. Gramian matrix if and only if positive semidefinite. □

Using Lemma 3 and Lemma 5, we obtain the following.

Corollary 1. *Any neutral mean proximity rule is a generalized scoring rule for which there exists a positive semidefinite score matrix.*

5 Linear Embeddings and Neutrality

Linear Embeddings $\phi(\tau \circ \sigma) = R_\tau \cdot \phi(\sigma)$.

Motivation: While mean proximity rules / generalized scoring rules capture many nice voting rules, they capture bad ones too. In particular, if we have $s(\sigma, \sigma') > s(\sigma, \sigma)$ for any σ and σ', then the rule will not return σ on the profile where all votes are σ, thus violating strong unanimity. While linear embeddings solve this problem and achieve many other desirable properties, they are quite unrestrictive - they still capture all positional scoring rules and the Kemeny rule.

Theorem 2. *Any mean proximity rule that has a representation using a linear embedding satisfies strong unanimity and neutrality.*

In addition, we can now define a distance function $d(\sigma, \sigma') = \|\phi(\sigma) - \phi(\sigma')\|$. It is easy to show that this, in addition to being a distance metric, is also left-invariant. Additionally, if we take the linear embedding that generates the
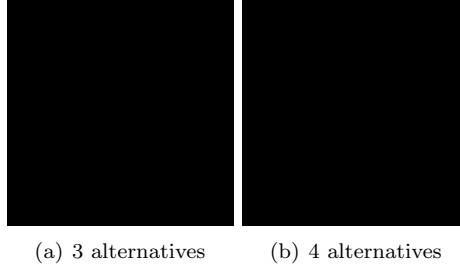
<div align="center">(a) 3 alternatives   (b) 4 alternatives</div>

<div align="center">Figure 1: Embeddings for the Borda Rule</div>

Kemeny rule, then the corresponding distance function becomes the squaure root of the KT distance. This demonstrates that the KT distance is actually the square of a more natural Euclidean distance metric. In this sense, the Kemeny rule is a mean rule, rather than a median rule, and so are all positional scoring rules.

Theorem 3. *A symmetric mean proximity rule is neutral if and only if there exists a representation ϕ of the rule that is linear.*

6 Euclidean Embeddings and Dimensions

In this section, we analyze the following general question. *Given a symmetric mean proximity rule r, what is the minimum dimension of any embedding that represents r?.*

Conjecture 1. *Let S and S' be two score matrices with rows $\{r_i\}$ and $\{r'_i\}$. Then, S and S' correspond to the same generalized scoring rule if and only if there exist $a \in \mathbb{R}$ and $b \in \mathbb{R}^{m!}$ such that for every i, $r'_i = a \cdot r_i + b$.*

Theorem 4. *The minimum dimension required for any positional scoring rule is $m - 1$.*

6.1 Borda Rule

In this section, we analyze the $m - 1$ dimension embeddings of the Borda rule.

3 alternatives: Zwicker [4] informally demonstrated that for 3 alternatives, any embedding that represents the Borda rule embeds the 6 possible rankings at the corners of a regular hexagon as shown in Figure 6.1. In fact, it is exactly the *permutahedron* of the symmetric group S_3 where each ranking is connected to two rankings that are obtained by two possible adjacent swaps. Thus, neighbouring rankings have *Kendall Tau* distance 1 from each other. Note that there are hexagons that connect rankings at KT distance greater than 1, but those do not correspond to the Borda rule.

4 alternatives: Interestingly, the observation that the Borda rule embeds rankings to the vertices of the regular polytope of the permutahedron carries over to the case of 4 alternatives. Thus, each ranking is connected to 3 rankings that are obtained by performing one of the three possible adjacent swaps. The polytope of the permutahedron of S_4, shown in Figure 6.1, consists of 8 hexagons and 6 squares. Each hexagon contains the 6 rankings that are obtained by either keeping the first alternative constant or the last alternative constant. Note that we can forget the alternative that is constant, and the hexagon exactly matches the embeddings of the rankings over the remaining 3 alternatives. The 6 squares each contain 4 rankings that are obtained by swapping the first two or the last two alternatives. Thus, each square has the following form:

$$a \succ b \succ c \succ d \ \longrightarrow \ a \succ b \succ d \succ c \ \longrightarrow \ b \succ a \succ d \succ c \ \longrightarrow \ b \succ a \succ c \succ d.$$

Replace this by a figure of a square, similarly draw a hexagon

For better understanding of permutahedrons, refer to Crisman [2].

7 Connections to other approaches

7.1 Axiomatic

Consistency, continuity, anonymity, neutrality already described above.

Question: When is it PM-c (related to the Euclidean distance being MC?), monotonic, majority for rankings?

7.2 DR

Consensus = Strong unanimity. Distance = Square of Euclidean distance. Votewise DR rules.

Square is not always a distance metric $\rightarrow$ Bad. More meaningfully, define votewise DR rules by sum of squares of distances rather than sum of distances. .

7.3 MLE

Take $\Pr[\sigma|\sigma^*] \propto e^{\|\phi(\sigma)-\phi(\sigma^*)\|^2}$. Not always Mallows since square is not always a metric. But actually better $\rightarrow$ Gaussian.

Neutrality $\Rightarrow$ Normalization independent of the true ranking $\Rightarrow$ Linear mean proximity rule is the MLE for this model.

Question: Efficient sampling?

Question: Anything interesting for the Gaussian distributions that connect to PSR?

8 Research Questions

1. What shape, and what corresponding voting rule do we get if we replace the KT distance by some other distance in the permutahedron?

2. Proving a lower bound on the dimensions for the Kemeny rule

3. Conjecture by Conitzer et. al. [1]: Consistent + continuous + neutral ⇔ Neutral SRSF

4. What are the equivalence classes of ϕ that lead to the same voting rule?

5. What about notions of consensus in Euclidean spaces other than mean → e.g., minimize maximum distance (everyone "lets go" equally)?

References

[1] V. Conitzer, M. Rognlie, and L. Xia. Preference functions that score rankings and maximum likelihood estimation. In *Proc. of 21st IJCAI*, pages 109–115, 2009.

[2] KARL-DIETER CRISMAN. The borda count, the kemeny rule, and the permutahedron.

[3] William S Zwicker. A characterization of the rational mean neat voting rules. *Mathematical and Computer Modelling*, 48(9):1374–1384, 2008.

[4] William S Zwicker. Consistency without neutrality in voting rules: When is a vote an average? *Mathematical and Computer Modelling*, 48(9):1357–1373, 2008.