Distance Rationalizability of Scoring Rules

Burak Can

Abstract Collective decision making problems can be seen as finding an outcome that is "closest" to a concept of "consensus". Nitzan (1981) introduced "Closeness to Unanimity Procedure" as a first example to this approach and showed that the Borda rule is the closest to unanimity under the Kemeny (1959) distance. Elkind et al. (2009) generalized this concept as distance-rationalizability, and showed that all scoring rules can be distance rationalized via a class of distance functions, which we call scoring distances. In this paper, we propose another class of distances, i.e., weighted distances, introduced in Can (2014). This class is a generalization of the Kemeny distance that rationalizes the generalization of the Borda rule, i.e., scoring rules. Hence the results here extend those in Nitzan (1981) and reveal the broader connection between Kemeny-like distances and Borda-like voting rules.

Keywords Distance rationalizability • Scoring rules • Voting • Weighted distances

1 Introduction

Nitzan [9] introduced the *closeness to unanimity procedures* (CUPs) for collective decision making problems. Given a distance function as a measure of closeness over preference profiles, these procedures find "closest" unanimous preference profiles to the original preference profile at hand. This approach, in a sense, yields the outcome which requires the minimal total compromise towards a unanimous agreement from a utilitarian perspective.

Meskanen and Nurmi [8] use other consensus concepts such as the existence of a Condorcet winner in a profile. Then, the compromise needed is not to achieve a unanimous profile, but to achieve a profile in which a Condorcet winner exists. They show that if the consensus concept is not unanimity, but a Condorcet winner

B. Can (✉)
Department of Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: b.can@maastrichtuniversity.nl

C. Binder et al. (eds.), *Individual and Collective Choice and Social Welfare*,
Studies in Choice and Welfare, DOI 10.1007/978-3-662-46439-7_11

instead, then the Dodgson winner in a profile is the closest to being a Condorcet winner under a compromise defined as the Kemeny (swap) distance.

Elkind et al. [4] generalize the notion of closeness to various concepts of consensus as *distance-rationalization*.[1] They use many reasonable consensus classes apart from unanimity, and employ different distance functions to shed light on the existing voting rules and their relation to distance functions within a consensus approach.

Nitzan [9] showed that the simplest scoring rule, i.e. the Borda rule, is equivalent to *closeness to unanimity procedure* under the Kemeny distance. This means that the Borda rule is somewhat rationalized by the Kemeny distance. Elkind et al. [4] extend this result and show that non-degenerate[2] scoring rules are rationalized by a class which we shall call *scoring-distances*. They also show that degenerate scoring rules, e.g., scoring rules which may have equal scores for different positions in a ranking, can be rationalized by pseudo-distances.[3]

In this paper, we show that the non-degenerate scoring rules can also be rationalized by another class of distance functions introduced in [2], i.e., *weighted distances*. There it is shown that weighted distances are generalizations of the Kemeny distance. Hence, the connection between the "Borda" rule and the "Kemeny" distance revealed in [9], can be extended to the connection between the "scoring rules" and the "weighted distances". The main difference between weighted distances and scoring distances in [4, 5], is that the former class satisfy a condition called *decomposability*. This condition is a weakening of one of the Kemeny distance axioms, i.e., *betweenness*. Hence the rationalizability of the Borda rule (with the Kemeny distance) is naturally extended to rationalizability of scoring rules (with the weighted Kemeny distances). The results also extend to distance rationalization of degenerate scoring rules by weighted pseudo-distances.

2 Model

2.1 Preliminaries

Let N be a finite set of agents with cardinality n, and A be a finite set of alternatives with cardinality m. The set of all possible strict preferences, i.e., complete, transitive and antisymmetric binary relations over A, is denoted by $\mathscr{L}$. A generic preference is denoted by $R \in \mathscr{L}$ whereas the set of strict preferences with an alternative a at the top is denoted by $\mathscr{L}^a$. A preference profile is an n-tuple vector of preferences

[1]For a broad analysis of the connection between distance functions and voting rules see [5], for distance rationalizability of Condorcet-consistent voting rules see [6].

[2]Non-degenerate scoring rules are scoring rules that assign decreasing scores to the positions in a ranking, therefore these rules do not include plurality, k-approval rule etc.

[3]A pseudo-distance is a function which satisfies all metric conditions except identity of indiscernibles.

denoted by $p = (p(1), p(2), \dots, p(n)) \in \mathscr{L}^N$. Given an alternative $a \in A$, we denote the profiles with a as the top alternative in each individual preference by p^a.

For $l = 1, 2, \dots, m$, $R(l)$ denotes the alternative in the lth position in R, e.g., $R(1)$ denotes the top alternative. Given an alternative a and a preference R, we denote the position of a in R by $a(R)$, i.e., $a(R) = x$ if and only if $a = R(x)$. To denote the position of alternative a in the preference of ith individual in a profile, we abuse notation and write $a(i)$ instead of $a(p(i))$, as long as it is clear which preference profile we refer to. Two linear orders $(R, R') \in \mathscr{L}^2$ form an *elementary change*[4] in position k whenever $R(k) = R'(k+1)$, $R'(k) = R(k+1)$ and for all $t \notin \{k, k+1\}$, $R(t) = R'(t)$, i.e. $|R \setminus R'| = 1$. Given any two distinct linear orders $R, R' \in \mathscr{L}$, a vector of linear orders $\rho = (R_0, R_1, \dots, R_k)$ is called a *path* between R and R' if $k = |R \setminus R'|$, $R_0 = R$, $R_k = R'$ and for all $i = 1, 2, \dots k$, (R_{i-1}, R_i) forms an elementary change. For the special case where $R = R'$, we denote the unique path as $\rho = (R, R)$.

A vector $s = (s_1, s_2, \dots, s_m)$ over positions of alternatives in a preference is called a *scoring vector* whenever $s_1 \geq s_2 \geq \dots s_m \geq 0$. A scoring vector s is called *non-degenerate* if scores are strictly decreasing from s_1 to s_m, i.e., $s_1 > s_2 > \dots s_m \geq 0$. The score of an alternative a in a preference R is denoted by $score(a, R)$ and is equal to $s_{a(R)}$ in the scoring vector.

A *collective choice rule*, or a voting rule, is a correspondence $\alpha : \mathscr{L}^N \rightarrow 2^A \setminus \emptyset$, which assigns each preference profile a nonempty subset of alternatives. Given a preference profile $p \in \mathscr{L}^N$, a *scoring rule*, denoted by α_s, with scoring vector s is a choice rule that assigns a summed score to each alternative in A, $\sum_{i \in N} score(a, p(i))$, and assigns to each profile the alternatives with maximal total scores,

$$\alpha_s(p) = \max_{a \in A} \sum_{i \in N} score(a, p(i))$$

Example 1 Let $s = (m-1, m-2, \dots, 0)$, then *the Borda rule* on each preference profile is defined as:

$$\alpha_{Borda}(p) = \arg\max_{a \in A} \sum_{i \in N} score(a, p(i)) = \arg\max_{a \in A} \sum_{i \in N} (m - a(i)))$$

Let us now dwell upon the concepts of "closeness" between individual preferences and thereafter preference profiles. Let a function $\delta : \mathscr{L} \times \mathscr{L} \rightarrow \mathbf{R}$ assign a real number to each pair of preferences. A function over preferences is a *distance function* if it satisfies:

(i) Non-negativity: $\delta(R, R') \geq 0$ for all $R, R' \in \mathscr{L}$,
(ii) Identity of indiscernibles: $\delta(R, R') = 0$ if and only if $R = R'$ for all R, $R' \in \mathscr{L}$,

[4] We omit the parenthesis whenever it is clear and write R, R' instead.

(iii) Symmetry: $\delta(R, R') = \delta(R', R)$ for all $R, R' \in \mathscr{L}$.
(iv) Triangular inequality: $\delta(R, R'') \leq \delta(R, R') + \delta(R', R'')$ for all $R, R', R'' \in \mathscr{L}$.

Two well-known examples of distance functions are *the discrete distance*, and *the swap distance*. The former assigns 0 if the two preferences are identical, and 1 otherwise. The latter function was introduced by Kemeny [7] and re-characterized with logically independent conditions in [3]. The Kemeny distance counts the symmetric total number of disjoint ordered pairs in preferences, or simply the minimal number of "swaps of adjacent alternatives" required to transform one preference into another.[5] Elkind et al. [4] also refer to functions that satisfy i, iii, and iv. These functions, which lack the identity of indiscernibles condition, are called *pseudo-distance functions*. These functions may assign 0 to distances between distinct pair of rankings, e.g., $\delta(abc, cab) = 0$.

For distance rationalizability we will mainly refer to distance functions between preference profiles. Given a distance function δ over preferences, a straightforward extension of δ over preference profiles, say $p, p' \in \mathscr{L}^N$, can be defined as a function $d : \mathscr{L}^N \times \mathscr{L}^N \to \mathbf{R}$ as follows:

$$d(p, p') = \sum_{i \in N} \delta(p(i), p'(i)).$$

Note that this is a very straightforward and common extension of distances over individual preferences to distances over preference profiles, e.g., see [1]. We abuse notation for the sake of simplicity by referring to δ instead of d as long as it is clear.

2.2 Distance Rationalizability

We only consider "unanimity" as a *consensus class*. The definitions below are adapted smoothly to our notation for simplicity. For a more general notation that would be applicable to many other consensus classes, we refer the reader to [4, 6].

Definition 1 ((U,δ)-Score) The unanimity-score of an alternative a in a preference profile p under the distance function δ is the minimal distance between the profile p and any profile p^a where a is unanimity winner. Formally:

$$(U, \delta){-}score(a, p) = \min_{p^a \in \mathscr{L}^N} \delta(p, p^a).$$

Roughly speaking, $(U, \delta) - score$ of an alternative in a profile tells us how costly it is to make this alternative the best alternative in each individual preference, i.e.,

[5] In the literature, the swap distance and the Kemeny distance are interchangeably used. Kemeny [7] originally assumes the distance for each swap in a ranking to be 2, whereas in many works, for convenience, this is normalized to 1. This occurs especially when the domain of preferences is strict and there is no indifference.

the unanimity winner. Obviously there are many possible preference profiles, p^a, where the alternative a is the unanimity winner. The aforementioned score assigns the total cost to convert the original profile to one of such profiles for which the total cost is minimal. Next we reproduce the definition of distance rationalizability. We adapt again from [6] to simplify our notation.

Definition 2 A collective choice rule α, is *distance-rationalizable* via unanimity and a distance function δ, or simply *(U,δ)-rationalizable*, if for all profiles $p \in \mathscr{L}^N$, we have:

$$\alpha(p) = \arg\min_{a \in A} [(U, \delta){-}score(a, p)]$$

To state verbally, a rule is (U,δ)-rationalizable if all outcomes the rule assigns to each profile are also the alternatives which have the minimal (U,δ)-scores for that profile, i.e., the least costly to make the unanimity winner with that distance function.

2.3 Weighted Distances

Can [2] introduced weighted distances as an extension of the Kemeny distance on strict rankings, which would allow for differential treatment of the position of elementary changes. For instance consider, $R = abc$, $R' = acb$, and $\bar{R} = bac$. The Kemeny distance between R and R' is 1 as well as the Kemeny distance between R and $\bar{R}$. However one might argue that the former two are less dissimilar than the latter two, i.e., $\delta_\omega(R, R') < \delta_\omega(R, \bar{R})$, because a swap at the top of rankings may be more critical than a swap at the bottom of thereof.

A weighted distance assigns weights to positions of such swaps with a weight vector on all possible swaps, e.g., $\omega = (\omega_1, \omega_2, \ldots, \omega_{m-1})$. For any two rankings that require more than a single swap, one would find the summation of sequential swaps on a shortest path between the two rankings (see Example 2 below for multiple paths). Hence a path between the two rankings is decomposed into elementary changes, and each elementary change is assigned its corresponding weight according to the weight vector.

Example 2 An example of the two possible shortest paths between $R = abc$ and $R' = cba$ would then be $\rho_1 = [abc, bac, bca, cba]$ and $\rho_2 = [abc, acb, cab, cba]$

For a technical description of the weighted distances, we refer the reader to [2]. Note that in the case of distance rationalizability, the complication regarding multiple paths between rankings do not occur. Hence, it is sufficient to illustrate a weighted distance with an example below:

Example 3 Let $R = abcd$, and $R' = dabc$. Consider the weight vector $\omega = (10, 3, 1)$ and a weighted distance δ_ω, i.e., a swap of alternatives at top creates a distance of 10, at the middle a distance of 3, and at the bottom a distance of 1. Then:

$$\delta(R, R') = \delta(ab\mathbf{cd}, ab\mathbf{dc}) + \delta(a\mathbf{bd}c, a\mathbf{db}c) + \delta(\mathbf{ad}bc, \mathbf{da}bc)$$
$$\delta(R, R') = \omega_3 + \omega_2 + \omega_1 = 10 + 3 + 1 = 14.$$

3 Results

Nitzan [9] proved that the plurality rule is $(U, \delta_{discrete})$-*rationalizable* and that the Borda rule is (U, δ_{Kemeny})-*rationalizable*. In this paper we extend the Borda result to all scoring rules via a new class of distance functions introduced in [2]. We show that any non-degenerate[6] scoring rule is (U, δ_ω)-rationalizable where δ_ω is a weighted distance with particular non-zero weights. For degenerate scoring rules, the rationalization still holds but with weighted pseudo-distances which allows for zero weights.

The class of weighted distance functions in [2] are characterized by two conditions on top of the usual metric conditions: *positional neutrality* and *decomposability*. Both conditions[7] are in fact weakening of characterizing axioms of the Kemeny distance, which allow for differential treatment of positions in a ranking. Therefore to allow for scoring rules other than the Borda rule, some weakening on the conditions on the distance functions is necessary. The results herein, therefore extend the existing interconnectedness (of the Borda rule and the Kemeny distance) to that of "all scoring rules" and "weighted distances". Weighted distances are Kemeny-like metrics which assign weights on the position of the swaps required to convert one (strict) ranking to another. In that respect, the Kemeny distance is also a weighted distance where weights on all possible swaps, regardless of their positions, are identical. The scoring distances introduced in [4], however, are not decomposable[8] hence they do not follow a Kemeny-like pattern.

Let α_s be a scoring choice rule with the scoring vector $s = (s_1, s_2, \ldots, s_m)$. Then consider a weighted distance δ_ω with the weight vector $\omega = \Delta s = (s_1 - s_2, s_2 - s_3, \ldots, s_{m-1} - s_m)$, i.e., the weight assigned to each swap is the difference between the scores of the relevant consecutive positions. In the following theorem we explain the connection with the class of weighted distance functions and the distance rationalizability of non-degenerate scoring rules.

---

[6]By non-degenerate scoring rule we mean a non-degenerate scoring vector wherein $s_i > s_{i+1}$ for all $i = 1, 2, \ldots, m$.

[7]Positional neutrality is simply equal treatments of swaps of adjacent alternatives on same positions whereas decomposability requires additive summation of distances on at least one path as in Example 2.

[8]For instance consider the Borda score vector $s = (2, 1, 0)$. According to the scoring distance, the distance between $R = abc$ and $R' = cba$ would be 4, i.e.,$\delta_{scoring}(R, R') = |s_1 - s_3| + |s_2 - s_2| + |s_3 - s_1| = 2 + 0 + 2$. However when you consider the two paths between R and R' in Example 2, it is easy to see that the summation on each of the paths should add up to 6.

Theorem 1 *A non-degenerate scoring rule* α_s *is* (U, δ)*-rationalizable if* $\delta = \delta_\omega$ *is a weighted distance with* $\omega = \Delta s$.

Proof Let $\delta = \delta_\omega$ be a weighted distance function with a weight vector $\omega = \Delta s = (s_i - s_{i+1})_{i=1}^{m-1}$. We want to show that α_s is (U, δ_ω)-rationalizable which means for all profiles $p \in \mathscr{L}^n$, and for all alternatives $a \in A$, we have $a \in \alpha_s(p)$ if and only if (U, δ_ω)-score of a is minimal for all $a \in A$. Take any $p \in \mathscr{L}^n$ and any $a \in A$. Now for each $i \in N$, let $\bar{p}^a(i) \in \mathscr{L}^a$ be such that $\bar{p}^a(i)$ is identical to $p(i)$ except that alternative a is taken to the top while everything else remains the same. By triangular inequality of δ_ω, note that $\bar{p}^a(i) = \arg\min_{p^a \in \mathscr{L}^a} \delta_\omega(p(i), p^a)$, i.e., $\bar{p}^a(i)$ is the closest to $p(i)$ among all other preferences which have a at the top. This is simply because when constructing $\bar{p}^a(i)$, we leave everything unchanged except bringing a to the top. Hence, for the constructed preference profile $\bar{p}^a \in \mathscr{L}^N$, the alternative a is the unanimity winner and furthermore $\bar{p}^a$ is the closest to the original profile p among all other profiles $p^a \in \mathscr{L}^N$ where a is the unanimity winner.

Then, $(U, \delta_\omega) - score(a, p)$ is $\sum_{i=1}^n \delta(p(i), \bar{p}^a(i))$. By definition of a weighted distance and construction of ω, this equals to $\sum_{i=1}^n \sum_{t=1}^{a(i)-1} \omega_t = \sum_{i=1}^n \sum_{t=1}^{a(i)-1} (s_t - s_{t+1})$, which[9] in turn equals to $\sum_{i=1}^n (s_1 - s_{a(i)}) = n \times s_1 - \sum_{i=1}^n s_{a(i)}$. Note that the score of a in α_s is $\sum_{i=1}^n s_{a(i)}$. Obviously, $n \times s_1 - \sum_{i=1}^n s_{a(i)}$ is minimal if and only if $\sum_{i=1}^n s_{a(i)}$ is maximal. Hence $(U, \delta_\omega) - score(a, p)$ is minimal if and only if $a \in \alpha_s(p)$. This completes the proof as the choice of p and a is arbitrary.

An immediate corollary is on the extension of the result to degenerate scoring rules via weighted pseudo-distances. The proof follows identical reasoning with the theorem above, except where an equal score assigned by the degenerate scoring rule to two adjacent positions leads to a zero weight. This leads to violation of "identity of indiscernibles" condition hence δ_ω is a pseudo distance.

Corollary 1 *A degenerate scoring rule* α_s *is* (U, δ)*-rationalizable if* $\delta = \delta_\omega$ *is a weighted pseudo-distance with* $\omega = \Delta s$.

Let us finally dwell upon the significance of these results. In Example 3, one can see "positional neutrality" leading to assigning the same value so long as the swaps are at the same position. "Decomposability" is also seen in the example via the additivity of distances on pairs that require a single swap. Decomposability is a natural weakening of the original Kemeny [7] betweenness condition. This particular weakening of characterizing conditions lead to the class of weighted distances which rationalize scoring rules. As we already know "Kemeny" and "Borda" are very interconnected, it is interesting to see that a natural "generalization" of the former, i.e., the weighted distances, helps us rationalize the "generalization" of the latter, i.e., the scoring rules.

---

[9]Note that if a is already at the top of $p(i)$, then this formulation gives 0. The equation $\sum_{i=1}^n \sum_{t=1}^{a(i)-1} \omega_t$ sums the weights (costs) of carrying alternative a to the top in each individual preference.

4 Conclusion

In this paper we show that the relation between the Borda rule and the Kemeny distance is further extended to a relation between all scoring rules and all weighted distances. In fact the relation even spans the degenerate scoring rules in case we extend the weighted functions to pseudo-distances.

The distance rationalization of scoring rules, as mentioned in the introduction, has already been shown in [4], albeit the metrics therein do not resemble the Kemeny distance. The scoring distances proposed in that paper fails to satisfy an additivity condition, i.e., decomposability. This condition is essential in the axiomatisation of the Kemeny distance, as shown [2, 3]. This paper shows in fact that distance rationalization of the scoring rules can be achieved via the weighted distances which mimic the features of the Kemeny distance. Hence, the rationalization result of the Borda rule with the Kemeny distance is carried over naturally to a rationalization result on Borda-like rules with Kemeny-like distances.

Acknowledgements The author appreciates the crucial feedback from two anonymous referees. This research has benefited from Netherlands Organisation for Scientific Research (NWO) grant with project nr. 400-09-354.

References

1. Baigent N (1987) Preference proximity and anonymous social choice. Q J Econ 102(1):161–169
2. Can B (2014) Weighted distances between preferences. J Math Econ 51:109–115
3. Can B, Storcken T (2013) A re-characterization of the Kemeny distance. METEOR Research Memoranda RM/13/009
4. Elkind E, Faliszewski P, Slinko A (2009) On distance rationalizability of some voting rules. In: Proceedings of the 12th conference on theoretical aspects of rationality and knowledge, New York, pp 108–117. doi:10.1145/1562814.1562831. http://doi.acm.org/10.1145/1562814.1562831
5. Elkind E, Faliszewski P, Slinko A (2010) On the role of distances in defining voting rules. In: Proceedings of the 9th international conference on autonomous agents and multiagent systems: volume 1 - volume 1, International Foundation for Autonomous Agents and Multiagent Systems, AAMAS'10, Richland, SC, pp 375–382. http://dl.acm.org/citation.cfm?id=1838206.1838259
6. Elkind E, Faliszewski P, Slinko A (2012) Rationalizations of condorcet-consistent rules via distances of hamming type. Soc Choice Welf 39(4):891–905
7. Kemeny J (1959) Mathematics without numbers. Daedalus 88(4):577–591
8. Meskanen T, Nurmi H (2008) Closeness counts in social choice. In: Braham M, Steffen F (eds) Power, freedom, and voting. Springer, Berlin, pp 289–306
9. Nitzan S (1981) Some measures of closeness to unanimity and their implications. Theory Decis 13(2):129–138