Distance rationalizability of scoring rules

Maastricht University

Burak Can

Distance Rationalizability of Scoring Rules

RM/13/068

GSBE

Maastricht University School of Business and Economics
Graduate School of Business and Economics

P.O Box 616
NL- 6200 MD Maastricht
The Netherlands

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”. [1] introduced “Closeness to Unanimity Procedure” as a first example to this approach and showed that the Borda rule is the closest to unanimity under swap distance (a.k.a the [2] distance). [3] shows that the Dodgson rule is the closest to Condorcet under swap distance. [4, 5] generalized this concept as distance-rationalizability, where being close is measured via various distance functions and with many concepts of consensus, e.g., unanimity, Condorcet etc. In this paper, we show that all non-degenerate scoring rules can be distance-rationalized as “Closeness to Unanimity” procedures under a class of weighted distance functions introduced in [6]. Therefore, the results herein generalizes [1] and builds a connection between scoring rules and a generalization of the Kemeny distance, i.e. weighted distances.

JEL classification: C63, D71, D72, D74 Keywords: Strict preferences; Rankings; Distance-rationalizability

1 Introduction

[1] introduced the *closeness to unanimity procedures* (CUPs) for collective decision making problems. Given a distance function - for the concept of closeness - over preference profiles, these procedures finds “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.

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

[3] uses other consensus concepts such as the existence of Condorcet winner in a profile. Then the compromise needed is not to achieve a unanimous profile but to achieve a Condorcet winner as least costly as possible. They show that if the consensus concept is not unanimity but only a Condorcet winner, then the Dodgson winner in a profile is the closest to the Condorcet winner under a compromise defined as the Kemeny (swap) distance.

[4, 5] generalize the notion of closeness to various concepts of consensus as *distance-rationalization*. They use many reasonable consensus classes apart from unanimity, and employ different distance functions to shed light on the existing collective choice rules and their relation to distance functions within a consensus approach.

We focus on one particular class of rules, i.e., non-degenerate[1] scoring rules. [1] showed that the simplest scoring rule, i.e. the Borda rule, is equivalent to *closeness to unanimity procedure* under Kemeny distance. This means Borda rule is somewhat rationalized by Kemeny distance. [4] and [7] extend this result and show that non-degenerate scoring rules are rationalized by a class which we shall call *scoring-distances*. They also show that if for a scoring rule, the scores of positions are equal for some positions, i.e., it is a degenerate scoring rule, then it can be rationalized by a pseudo-distance.

In this paper, we show that the non-degenerate scoring rules can also be rationalized by another class of distance functions introduced in [6], i.e., *weighted distances*. There it is shown that weighted distances are generalizations of the Kemeny distance. Hence, the connection between "Borda" and the "Kemeny" distance revealed in [1], can be extended to the connection between "scoring rules" and the "weighted distances". The main difference of our results, with those in [4, 7], is that the weighted distances satisfy a condition called *decomposability* which is a weakening of one of the Kemeny distance axioms, i.e., betweenness. Hence the rationalizability of Borda with Kemeny distance is naturally extended to rationalizability of scoring rules with weighted Kemeny distances.

2 Model

2.1 Preliminaries

Let N be a finite set of agents with cardinality n, and A be the set of 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

[1] Non-degenerate scoring rules are rules that never assign same score to different positions in a ranking, therefore these rules do not include plurality, k-approval rule etc.

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

For $l = 1, 2, \ldots, m$, $R(l)$ denotes the alternative in the l^{th} 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 i^{th} 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*[2] 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 $\rho = (R_0, R_1, \ldots, R_k)$ is called a *path* between R and R' if $k = |R \setminus R'|$, $R_0 = R$, $R_k = R'$ and for all $i = 1, 2, \ldots k$, (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 $s = (s_1, s_2, \ldots, s_m)$ over positions of alternatives in a preferences is called a *scoring vector* whenever $s_1 \geq s_2 \geq \ldots s_m \geq 0$. A scoring vector s is called *non-degenerate* if scores are strictly decreasing from s_1 to s_m, i.e., $s_1 > s_2 > \ldots s_m \geq 0$. 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 $\alpha : \mathscr{L}^N \to 2^A \setminus \emptyset$, 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 as: $\sum_{i \in N} score(a, p(i))$ 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 $s = (m-1, m-2, \ldots, 0)$, 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 $\delta : \mathscr{L} \times \mathscr{L} \to \mathbb{R}$ 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 $R, R' \in \mathscr{L}$,
(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 $R, R', R'' \in \mathscr{L}$.

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

Two well-known examples of distance function are *the discrete distance*, and *the swap distance*, a.k.a., the Kemeny distance. The former assigns 0 if the two preferences are identical, and 1 otherwise. The latter, characterized by [2] and [8], 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.[3][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 $d : \mathscr{L}^N \times \mathscr{L}^N \to \mathbb{R}$ as follows:

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

Note that it is a very straightforward and common extension of distances over individual preferences to distances over preference profiles, e.g., see [9]. 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 consider only the "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, 5].

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:*

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

Roughly speaking, $(U, \delta) - score$ of an alternative in a profile tells us how costly it is -in terms of a distance function- to make this alternative the best alternative in each individual preference, i.e., 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 [5] to simplify our notation.

[3] In the literature, the swap distance and the Kemeny distance is interchangeably used. [2] originally assumes the distance for elementary changes 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.

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 each outcome the rule assigns to each profile is also an alternative which have the minimal (U,δ)-score for that profile, i.e., the least costly to make the unanimity winner with that distance function.

2.3 Weighted distances

[6] 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 is 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, then, that require more than a single swap, one would find the summation of sequential swaps on a shortest[4] path between the two rankings. 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.

For a technical description of the weighted distances, we refer the reader to [6]. 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 2. 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:

$$\delta(R,R') = \delta(ab\mathbf{cd}, ab\mathbf{dc}) + \delta(a\mathbf{bd}c, a\mathbf{db}c) + \delta(\mathbf{ad}bc, \mathbf{da}bc)$$

$$\delta(R,R') = \omega_3 + \omega_2 + \omega_1 = 10 + 3 + 1 = 14.$$

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

3 Results

Nitzan (1981) proved that the plurality rule is $(U, \delta_{discrete})$*-rationalizable* and that Borda's rule is (U, δ_{Kemeny})*-rationalizable*. In this paper we extend the Borda result to all non-degenerate[5] scoring rules. In the sequel, we shall not use the term "non-degenerate" anymore to avoid repetition as long as it is clear. We show that any scoring rule is (U, δ_ω)-rationalizable where δ_ω is a weighted distance with particular weights. Note that the class of weighted distance functions are characterized by two conditions on top of the usual metric conditions, i.e., *positional neutrality* and *decomposability* (see [6]). Both conditions 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 Borda, such weakening of the distance functions is necessary.

The results, herein, extend the existing interconnectedness (of the Borda rule and the Kemeny distance) to those between "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.

We first state the result then discuss its implications and its relation to the distance functions that also rationalize scoring rules as in [7]. Let α_s be a scoring choice rule with the scoring vector $s = (s_1, s_2, \dots, s_m)$. Then consider a weighted distance δ_ω with the weight vector $\omega = \Delta s = (s_1 - s_2, s_2 - s_3, \dots, s_{m-1} - s_m)$, 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.

Theorem 1. *A scoring rule* α_s *is* (U, δ)*-rationalizable if* $\delta = \delta_\omega$ *with* $\omega = \Delta s$*, i.e.,* δ *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 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 construction $\bar{p}^a(i)$, 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

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

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[6] 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.

Let us finally dwell upon the significance of these results. The scoring distances introduced in [7] are not decomposable hence they are not weighted distances. This implies that a comparison between two ranking cannot be additively decomposed into elementary changes as in Example 2. This is slightly not fitting to the spirit of the Kemeny distance. However the main advantage of the framework applied in [7] is the functions therein can also be used to pseudo-distance rationalize many popular degenerate rules, e.g., Plurality, Borda, Veto, and k-approval. Weighted distances, by construction however, cannot rationalize such rules.

In Example 2, 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. Note that these are weakening of the original [2] axioms. This particular weakening of those conditions lead to the class of weighted distances. 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.

References

1. S. Nitzan, Theory and Decision **13**(2), 129 (1981)
2. J. Kemeny, Daedalus **88**(4), 577 (1959)
3. T. Meskanen, H. Nurmi, in *Power, Freedom, and Voting*, ed. by M. Braham, F. Steffen (Springer, 2008), pp. 289–306
4. E. Elkind, P. Faliszewski, A. Slinko, in *Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge* (ACM, New York, NY, USA, 2009), TARK ’09, pp. 108–117. DOI 10.1145/1562814.1562831. URL http://doi.acm.org/10.1145/1562814.1562831
5. E. Elkind, P. Faliszewski, A. Slinko, Social Choice and Welfare pp. 1–15 (2012)
6. B. Can, Weighted distances between preferences (2012). METEOR Research Memoranda RM/12/056
7. E. Elkind, P. Faliszewski, A. Slinko, 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, Richland, SC, 2010), AAMAS ’10, pp. 375–382. URL http://dl.acm.org/citation.cfm?id=1838206.1838259

[6] 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.

8. B. Can, T. Storcken, A re-characterization of the kemeny distance (2013). METEOR Research Memoranda RM/13/009
9. N. Baigent, The Quarterly Journal of Economics pp. 161–169 (1987)