Journal of Mathematical Economics 51 (2014) 109–115

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Weighted distances between preferences

Burak Can *

Department of Economics, School of Business and Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands

ARTICLE INFO

Article history:
Received 21 November 2012
Received in revised form
2 December 2013
Accepted 11 January 2014
Available online 6 February 2014

Keywords:
Strict preferences
Rankings
Kemeny distance
Ideological distance

ABSTRACT

Individual disagreements are assumed to be reflected in the preferences. Distance functions, e.g., the well-known Kemeny (1959) metric, are used to measure these disagreements. However, a disagreement on how to rank the top two alternatives may be perceived more (or less) than a disagreement on how to rank the bottom two alternatives. We propose two conditions on functions which characterize a class of weighted semi-metric functions. This class of semi-metrics allows to quantify disagreements according to where they occur in preferences. It turns out one of these functions, "the path minimizing function", is the only metric which generalizes the Kemeny metric.

1. Introduction

In various contexts, the analysis of differences or dissimilarities between opinions is crucial. Consider, for instance, a situation where like-minded people form clusters, interest groups, or political parties to implement their agenda on some institution. Another example would be situations in which dissimilarities between the social preference and that of the individuals cause discontent. In that case the extent of the discontent is very much dependent on the model of dissimilarity. It is important, thus, to measure how similar (or dissimilar) two individuals are regarding their preferences.

Preferences are often modeled as orders/rankings over available alternatives. To compare two preferences, it is, therefore, plausible to look at the alternatives which are ranked oppositely. The well-known Kemeny metric (Kemeny, 1959) is commonly used in that way. Consider a strict preference $a \succ b \succ c$, which is interpreted as: a is preferred to b, and b to c, and by transitivity, a to c. The Kemeny metric between $a \succ b \succ c$ and another strict preference $b \succ a \succ c$ is 1, because the two preferences only disagree on how to order a and b. However, the Kemeny metric between $a \succ b \succ c$ and $a \succ c \succ b$ is also 1 (the disagreement is now on how to order b and c). Therefore, the dissimilarity between the former two and the dissimilarity between the latter two are given identical weights. It is not that difficult though to imagine situations where a disagreement at the top of a preference leads to a larger conflict/dissimilarity than a disagreement at the bottom of the preferences.

We believe the variation on the dissimilarities caused by the position of disagreements in preferences might be useful in many applications. For instance, consider three search engines, (G)*oogle*, (Y)*ahoo* and (B)*ing*. Given a word search, assume these engines give a strict ranking of the same millions of alternatives, i.e., websites that are relevant to the search term. Suppose that G and Y provide identical rankings in the first hundred websites and differ completely in the remaining millions of websites. Suppose also that G differs from B in the ranking of the first hundred websites but is identical in the remaining millions of websites. Nevertheless, it is natural to argue that G is closer to Y than it is to B, even if G and Y disagree on how to rank the remaining millions of alternatives after the first hundred websites. This is because what apparently matters most for internet users in website rankings is the first twenty–thirty websites (BBC,[1] 2006) that are ranked. Another branch of applications would occur in cases where at least two individuals need to find consensus by making concessions, such as in bargaining or collective decision making. The implicit cost of these concessions, then, might depend on the positions of the disagreement between the individuals.

A disagreement at the bottom of a ranking might also create more dissimilarity. For instance, a ranking can be interpreted as a priority list for a rescue operation during a catastrophic event, or a priority list of occupations to vaccinate during a pandemic. Given the limited supply of time/vaccines it is natural

* Tel.: +31 43 3883807.
E-mail address: b.can@maastrichtuniversity.nl.

[1] http://news.bbc.co.uk/2/hi/technology/4900742.stm.

http://dx.doi.org/10.1016/j.jmateco.2014.01.002
0304-4068/

that a pair of rankings which disagree at the bottom are more dissimilar than a pair which disagree at the top. Furthermore, the variation in dissimilarity may not always follow a monotonically decreasing pattern in the position of disagreement. In fact, in cases where certain positions in preferences are critical, the dissimilarity caused by a change in those positions might be more than changes in other positions. An example would be the ranking of football teams in a league, where the last, say 3, teams of the last week's ranking are to be relegated to another league. Then, a swap in the last 3rd and 4th positions might be much more critical, hence influential in the dissimilarity between two rankings, than a swap elsewhere. Therefore, it makes sense to assign more dissimilarity to a change at those critical positions.

In this paper, to model dissimilarity between preferences, we propose to use certain functions on strict preferences[2] in a similar spirit of the Kemeny metric, i.e., respectful to the number of disagreements, but also allow variation in the treatment of different pairs of disagreements. To that end, we distinguish between metrics and semi-metrics. The latter do not necessarily satisfy the triangular inequality condition hence allows for more functions to analyze preferential differences. Although triangular inequality condition is mathematically very relevant, when it comes down to the dissimilarity of rankings it may not always be that realistic. For instance, consider three parties (l)eft, (r)ight and (c)entre and three voters with preferences respectively: $v_l = l \succ c \succ r$, $v_r = r \succ c \succ l$, and $v_c = c \succ r \succ l$. Note that voters have single-peaked preferences on the left–right political dimension with v_l denoting a left wing, v_r a right wing, and v_c a center–right voter. Triangular inequality imposes that the distance between v_l and v_r cannot be more than the sums of distances between v_l, v_c and v_c, v_r. This may not be very realistic since distance between two extremist at times may surpass the sum of distance between each extremist and the median voter.

In this paper, we provide two conditions that characterize a class of semi-metric functions. First one, "positional neutrality" is a neutrality condition towards the position of disagreement between two *adjacent* preferences, i.e., preferences which have only one disagreement. The second one, "decomposability" is a additivity-like condition which requires that the dissimilarity between any two preferences can be decomposed into a *path* of adjacent preferences between the two. We show some examples from the class of weighted semi-metric functions and also prove under which conditions these functions become metrics, i.e., they also satisfy the triangular inequality condition. We restrict our attention to strict preferences only and employ some group-theoretic results.

In Section 2, we introduce the notation and the basic conditions for semi-metric functions and define the two aforementioned conditions; positional neutrality and decomposability. In Section 3 we introduce the class of weighted semi-metrics which is characterized by these two conditions. Our main result characterizes one of these weighted functions, *the path minimizing function*, to be the only metric regardless of how weights are distributed on the positions. This result generalizes the Kemeny metric on strict preferences. In Section 4, we discuss some other examples of semi-metrics: the Kemeny metric, the Lehmer function, the inverse Lehmer function. We discuss under what type of weight distributions, these functions become also metrics, i.e., they satisfy the triangular inequality condition. Section 5 concludes the paper with some possible applications and future research.

[2] A strict preference on a set of alternatives is a complete, transitive and antisymmetric binary relation over that set of alternatives.

2. The model

2.1. Notation

Let A be the set of alternatives with cardinality $m \geq 3$. Strict preferences are modeled by linear orders[3] over A, and the set of all linear orders is denoted by $\mathcal{L}$. Given $R \in \mathcal{L}$, aRb is interpreted as a is strictly preferred to b, i.e., the ordered pair $(a, b) \in R$. We sometimes write $R = \cdots a \cdots b \cdots$ if aRb, and $R = \cdots ab \cdots$ if aRb and there exists no $c \in A \setminus \{a, b\}$ such that aRc and cRb, i.e., a and b are adjacent in R. Given any $a \in A$, $UC(a, R) = \{b \in A \mid bRa\}$ is the "upper contour set" of a in R, i.e., the set of alternatives that are ranked above a in the linear order R. Correspondingly, $LC(a, R) = \{b \in A \mid aRb\}$ is the "lower contour set" of a in R.

For $l = 1, 2, \ldots, m$, $R(l)$ denotes the alternative in the lth position in R. For some subset $B \subseteq A$, $R|_B$ denotes the preference reduced to B, i.e., $R|_B = R \cap (B \times B)$. Given any two linear orders $R, R' \in \mathcal{L}$, the set difference $R \setminus R'$ denotes the set of ordered pairs that exist in R and not in R', i.e., $\{(x, y) \in A \times A \mid xRy \text{ and } yR'x\}$. Two linear orders $(R, R') \in \mathcal{L}^2$ 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 $R, R' \in \mathcal{L}$, 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 bijection $\pi : \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$ is called a *permutation* and the set of all permutations is denoted by Π. We use $\pi(R)$ (or $\pi \cdot R$) to denote the permutation of the linear order R by π, i.e., $\pi(R) = R'$ if and only if $R(i) = R'(\pi(i))$ for all $i = 1, 2, \ldots, m$. Given $R, R' \in \mathcal{L}$, a permutation $\pi \in \Pi$ is called the *corresponding permutation*[5] for R, R', if $\pi(R) = R'$. We denote the *conjugate* of a permutation π by $\tilde{\pi} \in \Pi$, i.e., $\tilde{\pi}(R') = R$ if and only if $\pi(R) = R'$. A permutation that swaps the kth alternative of a linear order with $(k+1)$th is called an *elementary permutation* and is denoted by σ_k. Hence, σ_k is the corresponding permutation for any $R, R' \in \mathcal{L}$ that form an elementary change in position k. The set of all elementary permutations is denoted by $S = \{\sigma_1, \sigma_2, \ldots, \sigma_{m-1}\} \subseteq \Pi$. The identity permutation is denoted by σ_0.

Note that the set of all permutations Π over the set of alternatives A forms a symmetric group (also known as a permutation group) with the group operator "$\cdot$", which implies any permutation $\pi \in \Pi$ can be obtained by composition of some other permutations with the group operator, e.g., $\pi'' \cdot \pi' \cdot R = \pi \cdot R$ refers to the situation where R is first permuted via π' and then π'', and $\pi'' \cdot \pi' = \pi$. Note, however, that unless $m \leq 2$, the group fails commutativeness, e.g., for $R = abc$; note that $\sigma_1 \cdot \sigma_2 \cdot R = cab$ whereas $\sigma_2 \cdot \sigma_1 \cdot R = bca$.

In this paper, we will especially make use of compositions of elementary permutations in S. Since Π is a permutation group it has S, as the generator set, which means every permutation $\pi \in \Pi$, including the identity permutation σ_0, can be expressed by some composition of elements of S. Given $R, R' \in \mathcal{L}$, and a corresponding permutation $\pi \in \Pi$, let $I(\pi)$ denote the size of π, which is the number of minimal inversions required to obtain R' from R by elementary permutations. Note that as π is the corresponding permutation for R, R', we have that $I(\pi) = |R \setminus R'|$. Note also that for the identity permutation, we have $I(\sigma_0) = 0$. Next we define compositions of a permutations via elementary/identity permutations.

[3] Complete, transitive and antisymmetric binary relations.

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

[5] We omit this expression whenever it is clear which permutation we employ.