Maximum likelihood social choice rule *

Yuta Nakamura †

July 3, 2014

Abstract

This study is related to a Condorcetian problem of information aggregation that finds a "true" social ordering using individual orderings, that are supposed to partly contain the "truth". In this problem, we introduce a new maximum likelihood rule and analyze its performance. This rule selects an alternative that maximizes the probability of realizing individual orderings, conditional on the alternative being the top according to a true social ordering. We show that under a neutrality condition of alternatives, the probability that our rule selects the true top alternative is higher than that of any other rule.

1 Introduction

We consider the problem of searching for a "true" social ordering by aggregating individual orderings. Our purpose is to investigate properties of a new maximum likelihood rule, which is defined in line with the ideas of the maximum likelihood methods by Young (1988) and by Conitzer, Rognlie,

*The author thanks Yoko Kawada, Noriaki Okamoto, Shuhei Otani, and Toyotaka Sakai for helpful comments.

†Graduate School of Economics, Keio University, Tokyo 108-8345, Japan; yuta.nakamura0223@gmail.com

and Xia (2009). Our main result shows that under a neutrality condition of alternatives, the probability that our rule selects the true top alternative is higher than that of any other rule.

In his famous *Essai*, Condorcet investigated the way of breaking the so-called Condorcet cycle of alternatives yielded by pairwise voting, but it is known that his method does not work well when there are more than three alternatives (e.g., Black 1958). However, Young (1988) persuasively argued that what Condorcet is meant to say is in fact a maximum likelihood method.[1] He also studied that an alternative that is most likely to be the top of the true social ordering is not always the top of an ordering that is most likely to be true.[2] On the other hand, Young's maximum likelihood method finds an alternative that is the top of the ordering that is most likely to be true.

However, in the definition of Young's maximum likelihood method, voters only pairwisely compare alternatives, and the probability of being correct is the same among all pairwise comparisons. Conitzer and Sandholm (2005) and Conitzer, Rognlie, and Xia (2009) present a more general model in which an ordering submitted by each voter is an independent and identically distributed random variable. Conitzer and Sandholm (2005) examine which well-known social choice rules can be identified with a maximum likelihood method for some conditional probability distribution. Conitzer, Rognlie, and Xia (2009) offer a maximum likelihood method that finds an ordering, that is most likely to be true. On the other hand, our maximum likelihood rule selects an alternative that is most likely to be the top of the true social ordering.

One of the virtues of maximum likelihood methods is the statistical consis-

[1] Young points out that this method coincides with the Kemeny rule.

[2] For example, when the probability of voters being correct is close to 1/2, such an alternative is a Borda winner, which is not always the top of an ordering that is most likely to be true.

tency, that is, if there are sufficiently many voters, then maximum likelihood methods select the true outcome with a probability very close to one. In addition, we show that the probability that our maximum likelihood rule selects the top alternative is higher than that of any other neutral social choice function. Using this result, we show that if all non-top alternatives are equally undesirable, then our maximum likelihood rule maximizes an expected social welfare.

This paper is organized as follows: Section 2 provides our model. Section 3 presents our main result. Section 4 offers some discussion. Section 5 gives concluding comments. Appendix contains the proof of a lemma.

2 The model

Let $X = \{x_1, x_2, \ldots, x_m\}$ be the finite set of alternatives and $I = \{1, 2, \ldots, n\}$ the finite set of voters. An *ordering* $\succsim_i$ is a complete, transitive, and anti-symmetric binary relation on X.[3] Let $\mathscr{R}$ be the set of orderings on X. A *ranking* of $x \in X$ for $\succsim_i \in \mathscr{R}$ is

$$r(x, \succsim_i) = \mid \{y \in X : y \succsim_i x\} \mid .$$

An *(ordering) profile* of n voters is

$$\succsim = (\succsim_1, \succsim_2, \ldots, \succsim_n) \in \mathscr{R}^n.$$

Definition 1. *A social choice correspondence is a correspondence $F : \mathscr{R}^n \twoheadrightarrow X$ that maps each profile $\succsim \in \mathscr{R}^n$ to a nonempty subset $F(\succsim) \subset X$.*

Definition 2. *A social choice function is a function $f : \mathscr{R}^n \to X$ that maps*

[3]Completeness: for each $x, y \in X$, either $x \succsim_i y$ or $y \succsim_i x$, Transitivity: for every $x, y, z \in X$, $x \succsim_i y$ and $y \succsim_i z$ together imply $x \succsim_i z$, Anti-symmetry: for each $x, y \in X$, $x \succsim_i y$ and $y \succsim_i x$ implies $x = y$.

each profile $\succsim \in \mathscr{R}^n$ *to an alternative* $f(\succsim) \in X$.

Given any f and F, we say that f is a *selection* of F if $f(\succsim) \in F(\succsim)$ for all $\succsim \in \mathscr{R}^n$.

A *permutation* is a bijection π from X to itself. Let Π be the set of permutations. To simplify notation, for each $\succsim_i \in \mathscr{R}$ and $\pi \in \Pi$, $\pi(\succsim_i)$ denotes the ordering such that

$$x \succsim_i y \iff \pi(x)\pi(\succsim_i)\pi(y) \quad \forall x, y \in X.$$

Simirary, for each $\succsim \in \mathscr{R}^n$ and $\pi \in \Pi$, let

$$\pi(\succsim) = (\pi(\succsim_1), \pi(\succsim_2), \ldots, \pi(\succsim_n)).$$

We are interested in neutral social choice correspondences/functions, which do not discriminate alternatives in terms of their names.

Definition 3. *A social choice correspondence* $F : \mathscr{R}^n \twoheadrightarrow X$ *is neutral if for any* $\succsim \in \mathscr{R}^n$ *and* $\pi \in \Pi$,

$$F(\pi(\succsim)) = \pi(F(\succsim)).$$

Definition 4. *A social choice function* $f : \mathscr{R}^n \to X$ *is neutral if for any* $\succsim \in \mathscr{R}^n$ *and* $\pi \in \Pi$,

$$f(\pi(\succsim)) = \pi(f(\succsim)).$$

Lemma 1 ensures the existence of neutral selections of neutral social choice correspondences.

Lemma 1. *For any neutral social choice correspondence* $F : \mathscr{R}^n \twoheadrightarrow X$*, there exists a neutral selection* f *of* F.

Proof. See Appendix. $\square$

We consider situations in which there exists a unique "true" social ordering $R_0 \in \mathscr{R}$. Following Young (1988), we assume that the prior probability that an ordering is true is equal among all orderings, i.e., for any $R \in \mathscr{R}$, $\mathrm{P}(R_0 = R) = 1/m!$. Voters do not know which ordering is true, but each of them has an ordering $\succsim_i \in \mathscr{R}$ that he considers as the true social ordering. Our purpose is to find the top alternative of the true social ordering from voters' orderings, that is, to find $x \in X$ such that $r(x, R_0) = 1$.

In our analysis, $\succsim_i \in \mathscr{R}$ is treated as a random element, conditional on the true social ordering. $\mathrm{P}(\succsim_i \mid R_0 = R)$ denotes the probability that when $R \in \mathscr{R}$ is the true social ordering, i considers that $\succsim_i \in \mathscr{R}$ is true, where $\sum_{\succsim_i \in \mathscr{R}} \mathrm{P}(\succsim_i \mid R_0 = R) = 1$. For simplicity, we assume that $\mathrm{P}(\succsim_i \mid R_0 = R) > 0$ for any $\succsim_i \in \mathscr{R}$ and $R \in \mathscr{R}$. Each voter has an identical conditional probability of having $\succsim_i \in \mathscr{R}$, and the votes are statistically independent, i.e., for any $\succsim \in \mathscr{R}^n$ and $R \in \mathscr{R}$, $\mathrm{P}(\succsim \mid R_0 = R) = \prod_{i=1}^{n} \mathrm{P}(\succsim_i \mid R_0 = R)$.

In this paper, we throughout assume that conditional probability distributions satisfy that for any $\succsim_i \in \mathscr{R}$, $R \in \mathscr{R}$ and $\pi \in \Pi$,

$$\mathrm{P}(\succsim_i \mid R_0 = R) = \mathrm{P}(\pi(\succsim_i) \mid R_0 = \pi(R)).$$

This condition means that the relationship between the true social ordering and voters' orderings is independent from the names of alternatives.[4]

Let $\mathscr{Q}(x) \equiv \{R \in \mathscr{R} : r(x, R) = 1\}$. Denote the probability that an ordering profile $\succsim \in \mathscr{R}^n$ occurs when $x \in X$ is the top of the true social

[4]This assumption is also imposed by Conitzer, Rognlie and Xia (2009).

ordering by

$$
\begin{aligned}
\mathrm{P}(\succsim \mid r(x,R_0)=1) &= \mathrm{P}(\succsim \text{ and } r(x,R_0)=1 \mid r(x,R_0)=1) \\
&= \sum_{R\in\mathscr{Q}(x)} \mathrm{P}(\succsim \text{ and } R_0 = R \mid r(x,R_0)=1) \\
&= \sum_{R\in\mathscr{Q}(x)} \mathrm{P}(R_0 = R \mid r(x,R_0)=1)\mathrm{P}(\succsim \mid R_0 = R \text{ and } r(x,R_0)=1) \\
&= \sum_{R\in\mathscr{Q}(x)} \mathrm{P}(R_0 = R \mid r(x,R_0)=1)\mathrm{P}(\succsim \mid R_0 = R) \\
&= \frac{1}{(m-1)!}\sum_{R\in\mathscr{Q}(x)} \mathrm{P}(\succsim \mid R_0 = R) \\
&= \frac{1}{(m-1)!}\sum_{R\in\mathscr{Q}(x)} \prod_{i=1}^{n} \mathrm{P}(\succsim_i \mid R_0 = R),
\end{aligned}
$$

where the third equality follows from Bayes' rule.

Lemma 2. *For any* $\succsim \in \mathscr{R}^n$, $x \in X$ *and* $\pi \in \Pi$,

$$\mathrm{P}(\succsim \mid r(x,R_0)=1) = \mathrm{P}(\pi(\succsim) \mid r(\pi(x),R_0)).$$

Proof. Take any $\succsim \in \mathscr{R}$, $x \in X$ and $\pi \in \Pi$. For any $R \in \mathscr{R}$, since

$$
\begin{aligned}
r(\pi(x),\pi(R)) &= \mid \{y \in X : y\pi(R)\pi(x)\} \mid = \mid \{y \in X : \pi^{-1}(y)Rx\} \mid \\
&= \mid \{y \in X : yRx\} \mid = r(x,R),
\end{aligned}
$$

$R \in \mathscr{Q}(x)$ if and only if $\pi(R) \in \mathscr{Q}(\pi(x))$. Then,

$$\begin{aligned}
\mathrm{P}(\pi(\succsim) \mid \pi(x)) &= \frac{1}{(m-1)!} \sum_{R \in \mathscr{Q}(\pi(x))} \prod_{i=1}^{n} \mathrm{P}(\pi(\succsim_i) \mid R_0 = R) \\
&= \frac{1}{(m-1)!} \sum_{R \in \mathscr{Q}(x)} \prod_{i=1}^{n} \mathrm{P}(\pi(\succsim_i) \mid R_0 = \pi(R)) \\
&= \frac{1}{(m-1)!} \sum_{R \in \mathscr{Q}(x)} \prod_{i=1}^{n} \mathrm{P}(\succsim_i \mid R_0 = R) \\
&= \mathrm{P}(\succsim \mid r(x, R_0) = 1).
\end{aligned}$$

□

We define the maximum likelihood rule as a social choice correspondence.

Definition 5. *The maximum likelihood rule is a correspondence $F_M : \mathscr{R}^n \twoheadrightarrow X$ defined by:*

$$F_M(\succsim) = \arg\max_{x \in X} \mathrm{P}(\succsim \mid r(x, R_0) = 1).$$

F_M is a social choice correspondence that maps each profile $\succsim \in \mathscr{R}^n$ to a nonempty subset $F_M(\succsim) \subset X$, each element of which maximizes the probability that $\succsim \in \mathscr{R}^n$ occurs given that such an element is the top of the true social ordering. By Lemma 2, the maximum likelihood rule F_M is clearly neutral. Throughout this paper, we take any neutral selection f_M of F_M and fix it.

We can show that if there are sufficiently many voters, then the maximum likelihood rule can select the top alternative with the probability very close to one.

Proposition 1. *Suppose that for any permutation $\pi \in \Pi$ that is not an identity mapping, there exists some $\succsim_i \in \mathscr{R}$ that satisfies $\mathrm{P}(\succsim_i \mid r(x, R_0) = 1) \neq \mathrm{P}(\tau(\succsim_i) \mid r(x, R_0) = 1)$. Then, the probability that the maximum* *likelihood rule selects the top alternative of the true social ordering converges to one, as the number of voters approaches infinity.*

Proof. Because the maximum likelihood rule is an extremum estimator, this result follows from the consistency theorem for extremum estimators (e.g., Hayashi 2000, Proposition 7.1). $\square$

3 Performance of the maximum likelihood rule

In this section, we analyze the performance of the maximum likelihood rule. To begin with, for each social choice function $f : \mathscr{R}^n \to X$ and $x \in X$, let $\mathscr{S}_f(x) \equiv \{\succsim \in \mathscr{R}^n : f(\succsim) = x\}$. Then, the probability that when $x \in X$ is the top of the true ordering, a social choice function $f : \mathscr{R}^n \to X$ selects x is

$$\begin{aligned} \mathrm{P}[f(\succsim) = x \mid r(x, R_0) = 1] &= \sum_{\succsim \in \mathscr{S}_f(x)} \mathrm{P}(\succsim \mid r(x, R_0) = 1) \\ &= \frac{1}{(m-1)!} \sum_{\succsim \in \mathscr{S}_f(x)} \sum_{R \in \mathscr{Q}(x)} \mathrm{P}(\succsim \mid R_0 = R). \end{aligned}$$

Our main result shows that f_M can select the top alternative with higher probability than any other neutral social choice function.

Theorem 1. *For every neutral social choice function $f : \mathscr{R}^n \to X$ and every $x \in X$*

$$\mathrm{P}[f(\succsim) = x \mid r(x, R_0) = 1] \leq \mathrm{P}[f_M(\succsim) = x \mid r(x, R_0) = 1].$$

Proof. Take any neutral social choice function f and $x \in X$. Let $C \equiv$

$\mathscr{S}_f(x) \cap \mathscr{S}_{f_M}(x)$. Then,

$$\mathrm{P}[f(\succsim) = x \mid r(x, R_0) = 1] = \sum_{\succsim \in \mathscr{S}_f(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1) + \sum_{\succsim \in C} \mathrm{P}(\succsim \mid r(x, R_0) = 1),$$
$$\mathrm{P}[f_M(\succsim) = x \mid r(x, R_0) = 1] = \sum_{\succsim \in \mathscr{S}_{f_M}(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1) + \sum_{\succsim \in C} \mathrm{P}(\succsim \mid r(x, R_0) = 1).$$

Therefore, it suffices to show that

$$\sum_{\succsim \in \mathscr{S}_f(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1) \leq \sum_{\succsim \in \mathscr{S}_{f_M}(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1).$$

For each $y \in X$ with $y \neq x$, let a transposition $\tau^{yx} \in T$ be $\tau^{yx}(y) = x$ and $\tau^{yx}(x) = y$.[5] Now, let us show that for any $\succsim \in \mathscr{S}_f(x) \setminus C$, $\tau^{f_M(\succsim)x}(\succsim) \in \mathscr{S}_{f_M}(x) \setminus C$. Take any $\succsim \in \mathscr{S}_f(x) \setminus C$. By neutrality of f_M,

$$f_M(\tau^{f_M(\succsim)x}(\succsim)) = \tau^{f_M(\succsim)x}(f_M(\succsim)) = \tau^{f_M(\succsim)x}(f_M(\succsim)) = x.$$

Hence $\tau^{f_M(\succsim)x}(\succsim) \in \mathscr{S}_{f_M}(x)$. Next, we shall show $\tau^{f_M(\succsim)x}(\succsim) \notin C$. Because $\succsim \in \mathscr{S}_f(x) \setminus C$, $f(\succsim) = x$ and $f_M(\succsim) \neq x$. So by the neutrality of f,

$$f(\tau^{f_M(\succsim)x}(\succsim)) = \tau^{f_M(\succsim)x}(f(\succsim)) = \tau^{f_M(\succsim)x}(x) = f_M(\succsim) \neq x.$$

Therefore $\tau^{f_M(\succsim)x}(\succsim) \notin \mathscr{S}_f(x)$, so $\tau^{f_M(\succsim)x}(\succsim) \notin C$.

Now, we can define a function $h : \mathscr{S}_f(x) \setminus C \to \mathscr{S}_{f_M}(x) \setminus C$ by

$$h(\succsim) = \tau^{f_M(\succsim)x}(\succsim).$$

Let us show that h is injective.[6] Take any $\succsim, \succsim' \in \mathscr{S}_f(x) \setminus C$ with $\succsim \neq \succsim'$. If

[5]A transposition is a permutation $\tau \in \Pi$ such that there exist $x, y \in X$ that satisfy $\tau(x) = y$ and $\tau(y) = x$, and for any $z \in X$ with $z \neq x$ and $z \neq y$, $\tau(z) = z$. We denote T as the set of transpositions.

[6]In fact, h is bijective. However, we here only need injectivity of h.

$f_M(\succsim) = f_M(\succsim')$, then since $\succsim \neq \succsim'$, we have

$$h(\succsim) = \tau^{f_M(\succsim)x}(\succsim) \neq \tau^{f_M(\succsim)x}(\succsim') = \tau^{f_M(\succsim')x}(\succsim') = h(\succsim').$$

Next, let us consider the case $f_M(\succsim) \neq f_M(\succsim')$. Because $\succsim \in \mathscr{S}_f(x)$ implies $f(\succsim) = x$, by the neutrality of f,

$$f(\tau^{f_M(\succsim)x}(\succsim)) = \tau^{f_M(\succsim)x}(f(\succsim)) = \tau^{f_M(\succsim)x}(x) = f_M(\succsim).$$

Similarly, we can prove $f(\tau^{f_M(\succsim')x}(\succsim')) = f_M(\succsim')$. Hence

$$h(\succsim) = \tau^{f_M(\succsim)x}(\succsim) \neq \tau^{f_M(\succsim')x}(\succsim') = h(\succsim').$$

Therefore, h is injective.

Then,

$$\begin{aligned}
\sum_{\succsim \in \mathscr{S}_f(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1) &= \sum_{\succsim \in \mathscr{S}_f(x) \setminus C} \mathrm{P}(\tau^{f_M(\succsim)x}(\succsim) \mid r(\tau^{f_M(\succsim)x}(x), R_0) = 1) \\
&= \sum_{\succsim \in \mathscr{S}_f(x) \setminus C} \mathrm{P}(h(\succsim) \mid r(f_M(\succsim), R_0) = 1) \\
&= \sum_{\succsim \in h(\mathscr{S}_f(x) \setminus C)} \mathrm{P}(h(h^{-1}(\succsim)) \mid r(f_M(h^{-1}(\succsim)), R_0) = 1) \\
&= \sum_{\succsim \in h(\mathscr{S}_f(x) \setminus C)} \mathrm{P}(\succsim \mid r(f_M(h^{-1}(\succsim)), R_0) = 1) \\
&\leq \sum_{\succsim \in \mathscr{S}_{f_M}(x) \setminus C} \mathrm{P}(\succsim \mid r(f_M(h^{-1}(\succsim)), R_0) = 1) \\
&\leq \sum_{\succsim \in \mathscr{S}_{f_M}(x) \setminus C} \mathrm{P}(\succsim \mid r(x, R_0) = 1),
\end{aligned}$$

where the first equality follows from Lemma 2, the second equality follows from the definition of h and $\tau^{f_M(\succsim)x}$, the third equality follows from the fact that h is injection from $\mathscr{S}_f(x) \setminus C$ to $\mathscr{S}_{f_M}(x) \setminus C$, the first weak inequality follows from the fact that $h(\mathscr{S}_f(x) \setminus C) \subset \mathscr{S}_{f_M}(x) \setminus C$, and the second weak inequality follows from the definitions of $\mathscr{S}_{f_M}(x)$, F_M, and f_M. $\square$

As a corollary to Theorem 1, we can show that the decision by the maximum likelihood rule is more desirable than the decision by any one individual. To see this, for each $i \in I$, define a social choice function $f_i : \mathscr{R}^n \to X$ by

$$f_i(\succsim) = x \text{ with } r(x, \succsim_i) = 1.$$

Corollary 1. *For all $i \in I$ and all $x \in X$,*

$$\mathrm{P}[f_i(\succsim) = x \mid r(x, R_0) = 1] \leq \mathrm{P}[f_M(\succsim) = x \mid r(x, R_0) = 1].$$

Proof. Immediately follows from Theorem 1. $\square$

4 Discussion

4.1 Note on Theorem 1

In Theorem 1, we showed that the maximum likelihood rule is the most desirable in the class of neutral social choice functions. We explain why our analysis focuses on the class of neutral social choice functions.

4.1.1 Comparing functions

In Theorem 1, we compared any neutral social choice function with a neutral selection of the maximum likelihood rule. The reason why we compare social choice functions in Theorem 1 comes from the difficulty of comparing social choice correspondences.

To see this, consider a situation in which the maximum likelihood rule coincides with the Borda rule.[7] Suppose that there are three alternatives

[7]Conitzer and Sandholm (2005) show that any scoring rule can be identified with the

Table 1: The Borda rule

voter 1\2	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x,y	y	x	x,y,z
xzy	x	x	x	x,y,z	x,z	z
yxz	x,y	x	y	y	x,y,z	y
yzx	y	x,y,z	y	y	z	y,z
zxy	x	x,z	x,y,z	z	z	z
zyx	x,y,z	z	y	y,z	z	z

Table 2: The revised Borda rule

voter 1\2	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x,y	x	x,y	y	x	x,y,z
xzy	x	x,z	x	x,y,z	x,z	z
yxz	x,y	x	x,y	y	x,y,z	y
yzx	y	x,y,z	y	y,z	z	y,z
zxy	x	x,z	x,y,z	z	x,z	z
zyx	x,y,z	z	y	y,z	z	y,z

and two voters with the top alternative x of the true social ordering. Then, outcomes of the Borda rule for all ordering profiles are illustrated in Table 1. xyz means that x is ranked higher than y, y is ranked higher than z, and x is ranked higher than z. For example, Table 1 shows that if voter 1's ordering is yxz and voter 2's ordering is zyx, then the Borda outcome is y. Similarly, Table 2 illustrates another neutral social choice correspondence, say the "revised Borda rule." In the diagonal, the revised Borda rule is different from the Borda rule and the revised Borda rule selects two alternatives whereas the Borda rule selects only one alternative there.

Because we are looking for the top alternative of the true social ordering,

---

maximum likelihood rule for some conditional probability distribution. Therefore, such a situation exists. In Section 4.2, we give a sufficient condition that the maximum likelihood rule can be identified with some scoring rule.

we can only ambiguously compare these two rules. To see this point, at first, look at a profile where both voters have ordering xyz. Here, the Borda rule selects x, and the revised Borda rule selects x and y. Hence the Borda rule is more precise for this profile. Next, look at a profile where voter 1's ordering is yxz and voter 2's ordering is yxz. Then, the Borda rule selects y, so its outcome is not the top of the true social oredring. On the other hand, the revised Borda rule selects x and y, and it includes the top alternative x of the true social ordering. So the revised Borda rule is more precise for this profile. Therefore, we cannot simply conclude that the Borda rule is superior to the revised Borda rule.

This argument shows the difficulty of comparing social choice correspondences.

4.1.2 Necessity of neutrality

In the proof of Theorem 1, the neutrality plays an important role. Therefore, we cannot derive a same result for the class of anonymous social choice functions. For example, suppose that x is the top alternative and consider a social choice function that assigns x to all ordering profiles. This social choice function is anonymous. However, this social choice correspondence is obviously more desirable than the maximum likelihood rule because for some profiles, the maximum likelihood rule selects a non-top alternative of the true social ordering.

To construct another example that an anonymous function is more desirable than the maximum likelihood rule, consider social choice functions in Table 3 and Table 4 with the top alternative x of the true social ordering. These functions are anonymous selections of the Borda rule and the revised Borda rule, respectively. Their outcomes differ only in the profiles (yxz, yxz) and (zxy, zxy). In these profiles, the selection of the revised Borda rule is apparently more desirable than the selection of the Borda rule because the

Table 3: An anonymous selection of the Borda rule

voter 1\2	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	y	x	x
xzy	x	x	x	x	x	z
yxz	x	x	y	y	x	y
yzx	y	x	y	y	z	y
zxy	x	x	x	z	z	z
zyx	x	z	y	y	z	z

Table 4: An anonymous selection of the revised Borda rule

voter 1\2	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	y	x	x
xzy	x	x	x	x	x	z
yxz	x	x	x	y	x	y
yzx	y	x	y	y	z	y
zxy	x	x	x	z	x	z
zyx	x	z	y	y	z	z

selection of the revised Borda rule selects x in the profiles. Therefore, for any conditional probability distribution such that the maximum likelihood rule becomes the Borda rule, then the maximum likelihood rule cannot be the most desirable one in the class of anonymous functions.

However in reality, we cannot construct such correspondences because of the lack of information about the true social ordering. Therefore, Theorem 1 justifies the use of the maximum likelihood rule.

4.2 Scoring rules and the maximum likelihood rule

Here, we study a condition where the maximum likelihood rule is identified with a scoring rule.

A *score vector* is an m-dimensional vector $\alpha = (\alpha(1), \alpha(2), \dots, \alpha(m)) \in$

$\mathbb{R}^m$. Let A be the set of score vectors. The score of $x \in X$ for $\succsim \in \mathscr{R}^n$ at $\alpha \in A$ is defined by

$$S_\alpha(x, \succsim) = \sum_{i=1}^{n} \alpha(\succsim_i (x)).$$

For each $\alpha \in A$, the scoring rule with $\alpha \in A$ is a social choice correspondence $F_\alpha : \mathscr{R}^n \twoheadrightarrow X$ such that for all $\succsim \in \mathscr{R}^n$,

$$F_\alpha(\succsim) = \arg\max_{x \in X} S_\alpha(x, \succsim).$$

Proposition 2. *Suppose that there exists some $R \in \mathscr{R}$ such that the conditional probability distribution $\mathrm{P}(\cdot \mid R_0 = R)$ satisfies that whenever $r(x, \succsim_i) = r(x, \succsim_i')$ for some $x \in X$ with $r(x, R) = 1$, $\mathrm{P}(\succsim_i \mid R_0 = R) = \mathrm{P}(\succsim_i' \mid R_0 = R)$. Then, the maximum likelihood rule is identified with a scoring rule with some score vector, that is, there exists some $\alpha \in A$ such that*

$$F_M(\succsim) = F_\alpha(\succsim) \quad \textit{for all } \succsim \in \mathscr{R}^n.$$

Proof. Suppose that for $R \in \mathscr{R}$ and $x \in X$ with $r(x, R) = 1$, the conditional probability distribution $\mathrm{P}(\cdot \mid R_0 = R)$ satisfies that $r(x, \succsim_i) = r(x, \succsim_i')$ implies $\mathrm{P}(\succsim_i \mid R_0 = R) = \mathrm{P}(\succsim_i' \mid R_0 = R)$ for all $\succsim_i, \succsim_i' \in \mathscr{R}$.

Take any $\succsim \in \mathscr{R}^n$. For each $k \in \{1, 2, \ldots, m\}$, let

$$\mathrm{P}_k = \sum_{\succsim_i \in \mathscr{U}_k} \mathrm{P}(\succsim_i \mid R_0 = R), \quad \text{where } \mathscr{U}_k = \{\succsim_i \in \mathscr{R} : r(x, \succsim_i) = k\}.$$

By assumption, if $r(x, \succsim_i) = k$, then $\mathrm{P}(\succsim_i \mid R_0 = R) = 1/(m-1)!\mathrm{P}_k$.

For each $R' \in \mathscr{R}$, let $\pi^{R'R} \in \Pi$ be such that $\pi^{R'R}(R') = R$. Take any

$y \in X$. Then,

$$\begin{aligned}
\frac{1}{(m-1)!}\sum_{R'\in\mathscr{Q}(y)}\prod_{i=1}^{n}\mathrm{P}(\succsim_i|\ R_0=R') &= \frac{1}{(m-1)!}\sum_{R'\in\mathscr{Q}(y)}\prod_{i=1}^{n}\mathrm{P}(\pi^{R'R}(\succsim_i)\mid R_0=\pi^{R'R}(R'))\\
&= \frac{1}{(m-1)!}\sum_{R'\in\mathscr{Q}(y)}\prod_{i=1}^{n}\mathrm{P}(\pi^{R'R}(\succsim_i)\mid R_0=R)\\
&= \frac{1}{(m-1)!}\sum_{R'\in\mathscr{Q}(y)}\prod_{i=1}^{n}\frac{1}{(m-1)!}\mathrm{P}_{r(x,\pi^{R'R}(\succsim_i))}.
\end{aligned} \tag{1}$$

If $R' \in \mathscr{Q}(y)$, then by definition of $\mathscr{Q}(y)$, $r(y, R') = 1$. Therefore, for any $R' \in \mathscr{Q}(y)$,

$$r(\pi^{R'R}(y), R) = r(\pi^{R'R}(y), \pi^{R'R}(R')) = r(y, R') = 1 = r(x, R),$$

hence $\pi^{R'R}(y) = x$. Then, for any $R' \in \mathscr{Q}(y)$,

$$r(x, \pi^{R'R}(\succsim_i)) = r(\pi^{R'R}(y), \pi^{R'R}(\succsim_i)) = r(y, \succsim_i).$$

Thus, we obtain

$$\begin{aligned}
(1) &= \frac{1}{(m-1)!}\sum_{R'\in\mathscr{Q}(x)}\prod_{i=1}^{n}\frac{1}{(m-1)!}\mathrm{P}_{r(y,\succsim_i)}\\
&= \frac{1}{(m-1)!}\prod_{i=1}^{n}\mathrm{P}_{r(y,\succsim_i)}.
\end{aligned} \tag{2}$$

Finally, let $\alpha = (\log \mathrm{P}_1, \log \mathrm{P}_2, \ldots, \log \mathrm{P}_{\mathrm{m}}) \in A$. Then,

$$\begin{aligned}
F_M(\succsim) &= \arg\max_{x' \in X} \mathrm{P}(\succsim \mid r(x, R_0) = 1') \\
&= \arg\max_{x' \in X} \frac{1}{(m-1)!} \sum_{R' \in \mathscr{Q}(x')} \prod_{i=1}^{n} \mathrm{P}(\succsim_i \mid R_0 = R') \\
&= \arg\max_{x' \in X} \frac{1}{(m-1)!} \prod_{i=1}^{n} \mathrm{P}_{r(x', \succsim_i)} \\
&= \arg\max_{x' \in X} \sum_{i=1}^{n} \log \mathrm{P}_{r(x', \succsim_i)} \\
&= \arg\max_{x' \in X} \sum_{i=1}^{n} \alpha(r(x', \succsim_i)) \\
&= \arg\max_{x' \in X} S_\alpha(x', \succsim) = F_\alpha(\succsim).
\end{aligned}$$

□

We give an example that the hypothesis of Proposition 2 is satisfied.

Example 1. Suppose that there are n voters who want to find the biggest ball from $X = \{x, y, z\}$. Suppose also that x has a diameter of 11.0 cm, y has a diameter of 10.1 cm and z has a diameter of 10.0 cm. Then, the true social ordering is $R_0 = xyz$. In this case, each voter may not be able to distinguish y from z. So, we can approximately assume that $\mathrm{P}(xyz \mid R_0 = xyz) = \mathrm{P}(xzy \mid R_0 = xyz), \mathrm{P}(yxz \mid R_0 = xyz) = \mathrm{P}(zxy \mid R_0 = xyz)$, and $\mathrm{P}(yzx \mid R_0 = xyz) = \mathrm{P}(zyx \mid R_0 = xyz)$. Then, by Proposition 2, the maximum likelihood rule can be identified with a scoring rule with a score vector

$$(\log \mathrm{P}(xyz \mid R_0 = xyz), \log \mathrm{P}(yxz \mid R_0 = xyz), \log \mathrm{P}(yzx \mid R_0 = xyz)) \in A.$$

4.3 Expected social welfare

From Theorem 1, we can show that if all non-top alternatives are equally undesirable, then the maximum likelihood rule can maximize the expected welfare of society. To see this, denote the social welfare of $x \in X$ when $R_0 = R$ by $W(r(x, R)) \in \mathbb{R}$, where $W(1) \geq W(2) \geq \cdots \geq W(m)$. Then, the expected social welfare under a social choice function $f : \mathscr{R}^n \to X$ is

$$\mathrm{E}[W(r(f(\succsim), R_0))] = \sum_{R \in \mathscr{R}^n} \sum_{\succsim \in \mathscr{R}^n} W(r(f(\succsim), R))\mathrm{P}(\succsim \mid R_0 = R)\mathrm{P}(R_0 = R).$$

Proposition 3 states that if all non-top alternatives are equally undesirable, then the expected social welfare under f_M is greater than that under any other neutral social choice function.

Proposition 3. *Suppose that $W(1) > W(2) = \cdots = W(m)$. Then, for any neutral social choice function $f : \mathscr{R}^n \to X$,*

$$\mathrm{E}[W(r(f(\succsim), R_0))] \leq \mathrm{E}[W(r(f_M(\succsim), R_0))].$$

Proof. Take any neutral social choice fuction $f : \mathscr{R}^n \to X$. Let $\bar{W} \equiv W(2) = \cdots = W(m)$.

Take any $x \in X$. Then, the expected social welfare with f conditional on $r(x, R_0) = 1$ is

$$\begin{aligned}
&\mathrm{E}[W(r(f(\succsim), R_0)) \mid r(x, R_0) = 1] \\
&= \sum_{y \in X} \mathrm{E}[W(r(y, R_0)) \mid f(\succsim) = y \,\text{and}\, r(x, R_0) = 1] \cdot \mathrm{P}[f(\succsim) = y \mid r(x, R_0) = 1] \\
&= W(1)\,\mathrm{P}[f(\succsim) = x \mid r(x, R_0) = 1] + \bar{W}\,\{1 - \mathrm{P}[f(\succsim) = x \mid r(x, R_0) = 1]\}
\end{aligned} \tag{3}$$

By Theorem 1,

$$(3) \leq W(1)\,\mathrm{P}[f_M(\succsim) = x \mid r(x, R_0) = 1] + \bar{W}\,\{1 - \mathrm{P}[f_M(\succsim) = x \mid r(x, R_0) = 1]\}$$
$$= \mathrm{E}[W(r(f_M(\succsim), R_0)) \mid r(x, R_0) = 1].$$

Therefore,

$$\mathrm{E}[W(r(f(\succsim), R_0)) \mid r(x, R_0) = 1] \leq \mathrm{E}[W(r(f_M(\succsim), R_0)) \mid r(x, R_0) = 1],$$

which in turn implies

$$\begin{aligned}
\mathrm{E}[W(r(f(\succsim), R_0))] &= \sum_{x \in X} \mathrm{E}[W(r(f(\succsim), R_0)) \mid r(x, R_0) = 1]\mathrm{P}(r(x, R_0) = 1) \\
&\leq \sum_{x \in X} \mathrm{E}[W(r(f_M(\succsim), R_0)) \mid r(x, R_0) = 1]\mathrm{P}(r(x, R_0) = 1) \\
&= \mathrm{E}[W(r(f_M(\succsim), R_0))].
\end{aligned}$$

□

5 Conclusion

We studied the maximum likelihood rule that selects an alternative that is most likely to be the top of the true social ordering with assuming that voters' ordering is an i.i.d. random variable. We showed that the probability that the maximum likelihood rule chooses the top alternative of the true social ordering is higher than that of any other neutral social choice function. This result justifies the use of our maximum likelihood rule for information aggregation in our Condorcetian problem. Relaxing the i.i.d. assumption is left to the future research.

Appendix: Proof of Lemma 1

Take any neutral social choice correspondence F. Let

$$B_1 \equiv \{\succsim \in \mathscr{R}^n : |F(\succsim)| > 1\}.$$

If $B_1 = \emptyset$, then the unique selection of F is a desired selection. So we suppose $B_1 \neq \emptyset$. Since F is neutral, for any $\succsim \in B_1$ and $\pi \in \Pi$, $\pi(\succsim) \in B_1$. Take some $\succsim' \in B_1$ and $x \in F(\succsim')$, and define the social choice correspondence $F_1 : \mathscr{R}^n \twoheadrightarrow X$ by

$$F_1(\succsim) \equiv \begin{cases} F(\succsim) & \text{if } \nexists \pi \in \Pi, \succsim = \pi(\succsim') \\ \{\pi(x)\} & \text{if } \exists \pi \in \Pi, \succsim = \pi(\succsim') \end{cases}.$$

To show that F_1 is neutral, take any $\succsim \in \mathscr{R}^n$ and $\pi' \in \Pi$. If $\nexists \pi \in \Pi, \succsim = \pi(\succsim')$, then clearly $\nexists \pi \in \Pi, \pi'(\succsim) = \pi(\succsim')$, so that

$$F_1(\pi'(\succsim)) = F(\pi'(\succsim)) = \pi'(F(\succsim)) = \pi'(F_1(\succsim)).$$

If $\exists \pi \in \Pi, \succsim = \pi(\succsim')$, then $\pi'(\succsim) = \pi'(\pi(\succsim'))$, so

$$F_1(\pi'(\succsim)) = \{\pi'(\pi(x))\} = \pi'(\{\pi(x)\}) = \pi'(F_1(\succsim)).$$

Therefore F_1 is neutral.

Now, let

$$B_2 \equiv \{\succsim \in \mathscr{R}^n : |F_1(\succsim)| > 1\}.$$

By definition of $F_1, \succsim' \notin B_2$. If $B_2 = \emptyset$, then the unique selection of F_1 is a desired selection. So we suppose $B_2 \neq \emptyset$. Since F_1 is neutral, for any $\succsim \in B_2$ and $\pi \in \Pi$, We have $\pi(\succsim) \in B_2$. Again, take some $\succsim'' \in B_2$ and $x' \in F(\succsim'')$

and define the social choice correspondence $F_2 : \mathscr{R}^n \twoheadrightarrow X$ by

$$F_2(\succsim) \equiv \begin{cases} F_1(\succsim) & \text{if } \nexists \pi \in \Pi, \succsim = \pi(\succsim'') \\ \{\pi(x')\} & \text{if } \exists \pi \in \Pi, \succsim = \pi(\succsim'') \end{cases}$$

By a similar way, we can show that F_2 is neutral.

We can define B_k and the neutral social choice correspondence F_k by the same manner. Then, clearly

$$B_1 \supsetneq B_2 \supsetneq \cdots \supsetneq B_k.$$

Since B_1 is finite, there exists some k such that $B_k = \emptyset$, and the unique selection of F_{k-1} is a desired selection. □

References

Barbera, S. (1977). “The manipulation of social choice mechanisms that do not leave “ too much ” to chance.” *Econometrica*, *45*(7), 1573–1588.

Ben-Yashar, R., and Paroush, J. (2001). “Optimal decision rules for fixed-size commttees in polychotomous choice situations.” *Social Choice and Welfare*, *18*(4), 737–746.

Black, D. (1958). *The theory of committees and elections.* Springer.

Borda, J. C. d. (1784). “Mémoire sur les élections au scrutin.” *Histoire de l’Académie Royale des Sciences*, 657–664.

Condorcet, M. d. (1785). *Essai sur l’ application de l’ analysea la probabilité des décisions renduesa la pluralité des voix.* Chelfea Publishing Company(1972).

Conitzer, V., Rognlie, M., and Xia, L. (2009). "Preference functions that score rankings and maximum likelihood estimation." In *IJCAI*, vol. 9, 109–115.

Conitzer, V., and Sandholm, T. (2005). "Common voting rules as maximum likelihood estimators." *UAI*, 145–152.

Drissi-Bakhkhat, M., and Truchon, M. (2004). "Maximum likelihood approach to vote aggregation with variable probabilities." *Social Choice and Welfare*, *23*(2), 161–185.

Grofman, B., and Feld, S. L. (1988). "Rousseau's general will: a condorcetian perspective." *American Political Science Review*, *82*(02), 567–576.

Hayashi, F. (2000). *Econometrics.* Princeton University Press.

Kemeny, J. G. (1959). "Mathematics without numbers." *Daedalus*, *88*(4), 577–591.

Ladha, K. K. (1992). "The condorcet jury theorem, free speech, and correlated votes." *American Journal of Political Science*, *36*(3), 617–634.

McLean, I., Hewitt, F., et al. (1994). *Condorcet: foundations of social choice and political theory.* Edward Elgar Publishing.

McLean, I., Urken, A. B., and Hewitt, F. (1995). *Classics of social choice.* University of Michigan Press.

Okamoto, N., and Sakai, T. (2013). "The borda rule and the pairwise-majority-loser revisited." *unpublished manuscript, Keio University.*

Young, H. P. (1988). "Condorcet's theory of voting." *American Political Science Review*, *82*(04), 1231–1244.

Young, H. P. (1995). "Optimal voting rules." *Journal of Economic Perspectives*, *9*, 51–51.

Young, H. P., and Levenglick, A. (1978). "A consistent extension of condorcet's election principle." *SIAM Journal on Applied Mathematics*, *35*(2), 285–300.