The Japanese Economic Review
The Journal of the Japanese Economic Association

The Japanese Economic Review doi: 10.1111/jere.12065
Vol. 66, No. 2, June 2015

MAXIMUM LIKELIHOOD SOCIAL CHOICE RULE

By YUTA NAKAMURA
Keio University

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 analyse 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.
JEL Classification Numbers: D63, D71.

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 *et al*. (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 essay (Condorcet, 1785), Condorcet investigated the way of breaking the so-called Condorcet cycle of alternatives yielded by pairwise voting, but it has been known that his method does not work well when there are more than three alternatives (Black, 1958). However, Young (1988) persuasively argues that what Condorcet is meant to say is, in fact, a maximum likelihood method.[1] He also finds 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] In contrast, 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 *et al*. (2009) present a more general model in which an ordering submitted by each voter is an independent and identically distributed random element.[3] 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 *et al*. (2009) offer a maximum likelihood method that finds an ordering, that is most likely to be true. In contrast, our maximum likelihood rule selects an alternative that is most likely to be the top of the true social ordering.

---

[1] Young (1988) and (1995) point out that this method coincides with the Kemeny rule (Kemeny, 1959).

[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.

[3] Other generalizations can be found in, for example, Ben-Yashar and Paroush (2001) and Drissi-Bakhkhat and Truchon (2004).

In this paper, 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 also 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.

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.[4] Let $\mathscr{R}$ be the set of orderings on X. A *ranking* of $x \in X$ for $\succsim_i \in \mathscr{R}$ is

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

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 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 the notation, for each $\succsim_i \in \mathscr{R}$ and $\pi \in \Pi$, $\pi(\succsim_i)$ denotes the ordering such that

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

Similarly, 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* $\pi \in \Pi$,

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

[4] 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 $x = y$.

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 I. □

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 or she 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. $\mathrm{P}(\succsim_i | R_0 = R)$ 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 $\sum_{\succsim_i \in \mathscr{R}} \mathrm{P}(\succsim_i | R_0 = R) = 1$. For simplicity, we assume that $\mathrm{P}(\succsim_i | R_0 = R) > 0$ 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; that is, for any $\succsim \in \mathscr{R}^n$ and $R \in \mathscr{R}$, $\mathrm{P}(\succsim | R_0 = R) = \prod_{i=1}^n \mathrm{P}(\succsim_i | R_0 = R)$.

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

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

This condition means that the relationship between the true social ordering and voters' orderings is independent from the names of alternatives.[5]

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

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

[5] This assumption is also imposed by Conitzer *et al.* (2009).

where the third equality follows from Bayes' rule.

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

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

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

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

$R \in \mathscr{Q}(x)$ if and only if $\pi(R) \in \mathscr{Q}(\pi(x))$. Then, by Equation (1)

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

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.

3. Performance of the maximum likelihood rule

In this section, we analyse 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

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

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 $f : \mathscr{R}^n \to X$ that is not a selection of the maximum likelihood rule, and for every $x \in X$,*

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

Proof: Take any neutral social choice function f that is not a selection of the maximum likelihood rule, and any $x \in X$. 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

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

By definition of the maximum likelihood rule, for any $\succsim \in \mathscr{R}^n$,

$$\mathrm{P}(\succsim \mid r(f(\succsim), R_0) = 1) \leq \mathrm{P}(\succsim \mid r(f_M(\succsim), R_0) = 1). \tag{2}$$

Because f is not a selection of the maximum likelihood rule, there exists some $\succsim' \in \mathscr{R}^n$ such that

$$\mathrm{P}(\succsim' \mid r(f(\succsim'), R_0) = 1) < \mathrm{P}(\succsim' \mid r(f_M(\succsim'), R_0) = 1). \tag{3}$$

Let $\pi \in \Pi$ be such that $\pi(f_M(\succsim')) = x$. From Equation (3) and Lemma 2,

$$\mathrm{P}(\pi(\succsim') \mid r(\pi(f(\succsim')), R_0) = 1) < \mathrm{P}(\pi(\succsim') \mid r(\pi(f_M(\succsim')), R_0) = 1).$$

Moreover, by neutrality of f and f_M,

$$\mathrm{P}(\pi(\succsim') \mid r(f(\pi(\succsim')), R_0) = 1) < \mathrm{P}(\pi(\succsim') \mid r(f_M(\pi(\succsim')), R_0) = 1). \tag{4}$$

Note that $f_M(\pi(\succsim')) = \pi(f_M(\succsim')) = x$. In addition, because $f(\succsim') \neq f_M(\succsim')$, $f(\pi(\succsim')) = \pi(f(\succsim')) \neq \pi(f_M(\succsim')) = x$. Therefore, $\pi(\succsim') \in \mathscr{S}_{f_M}(x) \setminus C$. Then, by Equations (2) and (4),

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

Let us show that

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

For each $y \in X$ with $y \neq x$, let a transposition $\tau^{yx} \in T$ be $\tau^{yx}(y) = x$ and $\tau^{yx}(x) = y$.[6] Now, let us show that for any $\succsim \in \mathscr{S}_{f_M}(x)\setminus C$, $\tau^{f(\succsim)x}(\succsim) \in \mathscr{S}_f(x)\setminus C$. Take any $\succsim \in \mathscr{S}_{f_M}(x)\setminus C$. By neutrality of f,

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

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

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

Therefore, $\tau^{f(\succsim)x}(\succsim) \notin \mathscr{S}_{f_M}(x)$, so $\tau^{f(\succsim)x}(\succsim) \notin C$. Now, we can define a function $g : \mathscr{S}_{f_M}(x)\setminus C \to \mathscr{S}_f(x)\setminus C$ by

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

Let us show that g is bijective. At first, we show that g is injective. Take any $\succsim, \succsim'' \in \mathscr{S}_{f_M}(x)\setminus C$ with $\succsim \neq \succsim''$. If $f(\succsim) = f(\succsim'')$, then because $\succsim \neq \succsim''$, we have

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

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

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

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

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

Therefore, g is injective.

Next, we show that g is surjective. Take any $\succsim \in \mathscr{S}_f(x)\setminus C$. At first, let us show $\tau^{f_M(\succsim)x}(\succsim) \in \mathscr{S}_{f_M}(x)\setminus C$. By neutrality of f_M, $f_M(\tau^{f_M(\succsim)x}(\succsim)) = \tau^{f_M(\succsim)x}(f_M(\succsim)) = x$. Thus, $\tau^{f_M(\succsim)x}(\succsim) \in \mathscr{S}_{f_M}(x)$. Because $f(\succsim) = x$ and $f_M(\succsim) \neq x$, by neutrality of f, $f(\tau^{f_M(\succsim)x}(\succsim)) = \tau^{f_M(\succsim)x}(f(\succsim)) = \tau^{f_M(\succsim)x}(x) = f_M(\succsim) \neq x$. Thus, $\tau^{f_M(\succsim)x}(\succsim) \in \mathscr{S}_{f_M}(x)\setminus C$. Then,

6 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.

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

where the first equality follows from the fact

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

Therefore, g is surjective.

Then,

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

where the first equality follows from Lemma 2, the second equality follows from definitions of g and $\tau^{f(\succsim)x}$, the third and the final equality, follows from the fact that g is a bijection from $\mathscr{S}_{f_M}(x) \setminus C$ to $\mathscr{S}_f(x) \setminus C$. □

As a corollary to Theorem 1, we can show that the decision by the maximum likelihood rule is more desirable than that 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 $x \in X$, if f_i is not a selection of the maximum likelihood rule, then*

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

Proof: Immediately follows from Theorem 1. □

4. Discussion

4.1 Note on Theorem 1

In Theorem 1, we showed that the maximum likelihood rule is 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 A difficulty of comparing correspondences

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.

TABLE 1
Winners of 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
Winners of 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*

To see this, consider a situation in which the maximum likelihood rule coincides with the Borda rule.[7] Suppose that there are three alternatives and two voters, and that the top alternative of the true social ordering is *x*. The Borda winners for all profiles are illustrated in Table 1. There, *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 winner is *y*. Similarly, Table 2 illustrates another neutral social choice correspondence, the "revised Borda rule." In the diagonal, the revised Borda winners are different from the Borda winners and the revised Borda rule selects two alternatives whereas the Borda rule selects only one alternative.

Because of the multiplicity of winners under correspondences, 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 both voters have orderings *yxz*. Then, the Borda rule selects *y*, so its outcome is not the top of the true social ordering. In contrast, 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, and vice versa. This argument shows the difficulty of comparing social choice correspondences.

[7] Conitzer and Sandholm (2005) show that any scoring rule can be identified with the maximum likelihood rule for some conditional probability distribution. Therefore, such a situation exists. In Subsection 4.2, we give a sufficient condition that the maximum likelihood rule coincides with some scoring rule.

TABLE 3
Winners of 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
Winners of 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*

4.1.2 Necessity of neutrality

In the proof of Theorem 1, neutrality plays an important role. We cannot derive the same result for the class of anonymous social choice functions. For example, suppose that x is the top alternative of the true social ordering and consider a social choice function that assigns x to all 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 where an anonymous function is more desirable than the maximum likelihood rule, consider social choice functions in Tables 3 and 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 selection of the revised Borda rule chooses x. Therefore, whenever the maximum likelihood rule coincides with the Borda rule, the maximum likelihood rule cannot be the most desirable one in the class of anonymous functions.

4.2 Scoring rules and the maximum likelihood rule

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

A *score vector* is an m-dimensional vector $\alpha = (\alpha(1), \alpha(2), \ldots, \alpha(m)) \in \mathbb{R}^m$.[8] The score of $x \in X$ for $\succsim \in \mathscr{R}^n$ at $\alpha \in \mathbb{R}^m$ is defined by

[8] Later, we focus on descending score vectors; that is, $\alpha(1) \geq \alpha(2) \geq \cdots \geq \alpha(m)$.

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

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

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

Proposition 1: *Suppose that there exist some $R \in \mathcal{R}$ and $x \in X$ with $r(x, R) = 1$ such that $r(x, \succsim_i) = r(x, \succsim_i')$ implies $\mathrm{P}(\succsim_i | R_0 = R) = \mathrm{P}(\succsim_i' | R_0 = R)$. Then, the maximum likelihood rule coincides with a scoring rule with some score vector; that is, there exists some $\alpha \in \mathbb{R}^m$ such that*

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

Proof: Suppose that for $R \in \mathcal{R}$ and $x \in X$ with $r(x, R) = 1$, $r(x, \succsim_i) = r(x, \succsim_i')$ implies $\mathrm{P}(\succsim_i | R_0 = R) = \mathrm{P}(\succsim_i' | R_0 = R)$.

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

$$\mathrm{P}_k \equiv \sum_{\succsim_i' \in \mathcal{U}_k} P(\succsim_i' | R_0 = R), \quad \text{where } \mathcal{U}_k \equiv \{\succsim_i' \in \mathcal{R} : r(x, \succsim_i') = k\}. \tag{5}$$

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

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

$$\begin{aligned} \frac{1}{(m-1)!} \sum_{R' \in \mathcal{Q}(y)} \prod_{i=1}^{n} \mathrm{P}(\succsim_i | R_0 = R') &= \frac{1}{(m-1)!} \sum_{R' \in \mathcal{Q}(y)} \prod_{i=1}^{n} \mathrm{P}(\pi^{R'R}(\succsim_i) | R_0 = \pi^{R'R}(R')) \\ &= \frac{1}{(m-1)!} \sum_{R' \in \mathcal{Q}(y)} \prod_{i=1}^{n} \mathrm{P}(\pi^{R'R}(\succsim_i) | R_0 = R) \\ &= \frac{1}{(m-1)!} \sum_{R' \in \mathcal{Q}(y)} \prod_{i=1}^{n} \frac{1}{(m-1)!} \mathrm{P}_{r(x, \pi^{R'R}(\succsim_i))}. \end{aligned} \tag{6}$$

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

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

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

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

Thus, we obtain

$$(6) = \frac{1}{(m-1)!} \sum_{R' \in \mathcal{Q}(y)} \prod_{i=1}^{n} \frac{1}{(m-1)!} \mathrm{P}_{r(y, \succsim_i)}$$
$$= \frac{1}{(m-1)!} \prod_{i=1}^{n} \mathrm{P}_{r(y, \succsim_i)}. \tag{7}$$

Finally, let $\alpha = (\log \mathrm{P}_1, \log \mathrm{P}_2, \ldots, \log \mathrm{P}_m) \in \mathbb{R}^m$. Then,

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

□

Corollary 2: *Suppose that the assumption in the Proposition 1 is satisfied. Suppose also that* $(P_1, P_2, \ldots, P_m)$ *in Equation* (5) *in the proof of Proposition 1 satisfies that* $P_1 \geq P_2 \geq \cdots \geq P_\mathrm{m}$. *Then, the maximum likelihood rule coincides with a scoring rule with a score vector* $\alpha \in \mathbb{R}^m$ *that satisfies* $\alpha(1) \geq \alpha(2) \geq \cdots \geq \alpha(m)$*; that is,*

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

Proof: Immediately follows from the proof of Proposition 1. □

We give an example that the hypothesis of Proposition 1 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 1, the maximum likelihood rule can be identified with a scoring rule with a score vector:

$$(\log \mathrm{P}(xyz \mid R_0 = xyz), \log \mathrm{P}(yxz \mid R_0 = xyz), \log \mathrm{P}(yzx \mid R_0 = xyz)) \in \mathbb{R}^m.$$

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 certain welfare of society. To see this, denote the social welfare under $x \in X$ when $R_0 = R$ by $W(r(x, R)) \in \mathbb{R}$, where $W(1) \geq W(2) \geq \cdots \geq W(m)$. Then, our expected social welfare under a social choice function $f : \mathscr{R}^n \to X$ is defined by

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

Proposition 2 implies 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 2: *Suppose that $W(1) \geq W(2) = \cdots = W(m)$. Then, for any neutral social choice function $f : \mathscr{R}^n \to X$ that is not a selection of the maximum likelihood rule,*

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

Proof: Take any neutral social choice function $f : \mathscr{R}^n \to X$ that is not a selection of the maximum likelihood rule. 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

$$\begin{aligned} &\mathrm{E}[W(r(f(\succsim), R_0)) \,|\, r(x, R_0) = 1] \\ &= \sum_{y \in X} \mathrm{E}[W(r(y, R_0)) \,|\, f(\succsim) = y \text{ and } r(x, R_0) = 1] \cdot \mathrm{P}[f(\succsim) = y \,|\, r(x, R_0) = 1] \\ &= W(1)\mathrm{P}[f(\succsim) = x \,|\, r(x, R_0) = 1] + \bar{W}\{1 - \mathrm{P}[f(\succsim) = x \,|\, r(x, R_0) = 1]\}. \end{aligned} \tag{8}$$

By Theorem 1,

$$\begin{aligned} (8) &< W(1)\mathrm{P}[f_M(\succsim) = x \,|\, r(x, R_0) = 1] + \bar{W}\{1 - \mathrm{P}[f_M(\succsim) = x \,|\, r(x, R_0) = 1]\} \\ &= \mathrm{E}[W(r(f_M(\succsim), R_0)) \,|\, r(x, R_0) = 1]. \end{aligned}$$

Therefore,

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

which, in turn, implies

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

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 under the assumption that voters' orderings are i.i.d. random elements. 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.

Acknowledgements

The author gratefully acknowledges Toyotaka Sakai for his continuous encouragement and support, and an anonymous referee and Shuhei Otani for their very valuable comments. I am also grateful for comments from Yoko Kawada, Noriaki Okamoto, and participants at SAET Satellite Conference "Normative Economics from the Traditional and Nontraditional View" held at Keio University in August 2014, at the Summer Meeting on Game Theory held at Shimosuwa in August 2014, and at the 20th Decentralization Conference in Japan held at Fukuoka University in October 2014.

Appendix I

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 $B_1 \neq \emptyset$. Because 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 a social choice correspondence $F_1 : \mathscr{R}^n \twoheadrightarrow X$ by

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

To show that F_1 is neutral, take any $\succsim \in \mathscr{R}^n$ and $\pi' \in \Pi$. If there is no $\pi \in \Pi$ such that $\succsim = \pi(\succsim')$, then clearly there is no $\pi \in \Pi$ such that $\pi'(\succsim) = \pi(\succsim')$. Therefore,

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

Similarly, if there exists $\pi \in \Pi$ with $\succsim = \pi(\succsim')$, then $\pi'(\succsim) = \pi'(\pi(\succsim'))$. Therefore,

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

Thus, 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 $B_2 \neq \emptyset$. Because 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 a social choice correspondence $F_2 : \mathscr{R}^n \twoheadrightarrow X$ by

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

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

We can define, for each $k \in \mathbb{N}$, B_k and a neutral social choice correspondence F_k by the same manner. Then, clearly for any $k \in \mathbb{N}$,

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

Because B_1 is a finite set, there exists $k' \in \mathbb{N}$ such that $B_{k'} = \emptyset$, and the unique selection of $F_{k'-1}$ is a desired selection. □

Final version accepted 29 November 2014.

REFERENCES

Ben-Yashar, R. and J. Paroush (2001) "Optimal Decision Rules for Fixed-Size Committees in Polychotomous Choice Situations", *Social Choice and Welfare*, Vol. 18, No. 4, pp. 737–746.

Black, D. (1958 [2011]) *The Theory of Committees and Elections*, Cambridge: Cambridge University Press.

Conitzer, V., M. Rognlie and L. Xia (2009) "Preference Functions That Score Rankings and Maximum Likelihood Estimation", *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence*, Vol. 9, pp. 109–115.

—— and T. Sandholm (2005) "Common Voting Rules as Maximum Likelihood Estimators", *Proceedings of the Twenty-First Conference on Uncertainty in Artificial Intelligence*, pp. 145–152.

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

Drissi-Bakhkhat, M. and M. Truchon (2004) "Maximum Likelihood Approach to Vote Aggregation with Variable Probabilities", *Social Choice and Welfare*, Vol. 23, No. 2, pp. 161–185.

Kemeny, J. G. (1959) "Mathematics without Numbers", *Daedalus*, Vol. 88, No. 4, pp. 577–591.

Young, H. P. (1988) "Condorcet's Theory of Voting", *American Political Science Review*, Vol. 82, No. 4, pp. 1231–1244.

—— (1995) "Optimal Voting Rules", *Journal of Economic Perspectives*, Vol. 9, pp. 51–64.