Distance rationalization of anonymous and homogeneous voting rules

Mark C. Wilson
(joint with Benjamin Hadjibeyli, ENS Lyon)

Department of Computer Science
University of Auckland
www.cs.auckland.ac.nz/~mcw/

SCW 2014, Boston, 2014-06-20

Background

- The number of possible social choice rules is huge, but most are arbitrary and hard to understand.

Background

- The number of possible social choice rules is huge, but most are arbitrary and hard to understand.
- Researchers have tried many different axioms in order to classify and characterize these rules, sometimes leading to impossibility theorems.

Background

- The number of possible social choice rules is huge, but most are arbitrary and hard to understand.
- Researchers have tried many different axioms in order to classify and characterize these rules, sometimes leading to impossibility theorems.
- New rules are still being introduced, and the subject is far from tidy. "We have barely scratched the surface of the space of social choice rules." (W. Zwicker, 2009).

Background

- The number of possible social choice rules is huge, but most are arbitrary and hard to understand.
- Researchers have tried many different axioms in order to classify and characterize these rules, sometimes leading to impossibility theorems.
- New rules are still being introduced, and the subject is far from tidy. "We have barely scratched the surface of the space of social choice rules." (W. Zwicker, 2009).
- Distance rationalizability is a promising unifying framework that may allow us to find new rules with good properties. Studied in particular detail recently by Elkind, Faliszewski and Slinko.

Basic social choice setup

- A finite set V of voters, $n := |V|$, and a finite set A of alternatives, $m := |A|$.

Basic social choice setup

- A finite set V of voters, $n := |V|$, and a finite set A of alternatives, $m := |A|$.
- $L(A) :=$ set of linear orders on A, representing the possible strict preference orders. A profile is a map $V \to L(A)$ giving each voter's preference order.

Basic social choice setup

- A finite set V of voters, $n := |V|$, and a finite set A of alternatives, $m := |A|$.
- $L(A) :=$ set of linear orders on A, representing the possible strict preference orders. A profile is a map $V \to L(A)$ giving each voter's preference order.
- Let $\mathcal{E}$ denote the set of elections — triples (V, A, π) where π is a profile. A social choice rule is a mapping $\mathcal{E} \to 2^A \setminus \emptyset$.

Basic social choice setup

- A finite set V of voters, $n := |V|$, and a finite set A of alternatives, $m := |A|$.
- $L(A) :=$ set of linear orders on A, representing the possible strict preference orders. A profile is a map $V \to L(A)$ giving each voter's preference order.
- Let $\mathcal{E}$ denote the set of elections — triples (V, A, π) where π is a profile. A social choice rule is a mapping $\mathcal{E} \to 2^A \setminus \emptyset$.
- Voting situation: orbit of action of permutations of voters on set of profiles. Can be represented as a composition (ordered partition) of n with $m!$ parts. We usually normalize by dividing by n, yielding the preference simplex Δ.

Basic social choice setup

- A finite set V of voters, $n := |V|$, and a finite set A of alternatives, $m := |A|$.
- $L(A) :=$ set of linear orders on A, representing the possible strict preference orders. A profile is a map $V \to L(A)$ giving each voter's preference order.
- Let $\mathcal{E}$ denote the set of elections — triples (V, A, π) where π is a profile. A social choice rule is a mapping $\mathcal{E} \to 2^A \setminus \emptyset$.
- Voting situation: orbit of action of permutations of voters on set of profiles. Can be represented as a composition (ordered partition) of n with $m!$ parts. We usually normalize by dividing by n, yielding the preference simplex Δ.
- There are $m!^n$ profiles and $\binom{n+m!-1}{n}$ voting situations.

Some social choice rules

- The scoring rule defined by a fixed weight vector $(w_1, \ldots, w_m)$ gives w_i points to each candidate for each voter listing it in ith position. Highest total score wins. Special cases:

Some social choice rules

- The scoring rule defined by a fixed weight vector $(w_1, \ldots, w_m)$ gives w_i points to each candidate for each voter listing it in ith position. Highest total score wins. Special cases:
  - plurality $(1, 0, 0, \ldots, 0)$;

Some social choice rules

- The scoring rule defined by a fixed weight vector $(w_1, \ldots, w_m)$ gives w_i points to each candidate for each voter listing it in ith position. Highest total score wins. Special cases:
  - plurality $(1, 0, 0, \ldots, 0)$;
  - Borda $(m-1, m-2, \ldots, 1, 0)$;

Some social choice rules

- The scoring rule defined by a fixed weight vector $(w_1, \ldots, w_m)$ gives w_i points to each candidate for each voter listing it in ith position. Highest total score wins. Special cases:
  - plurality $(1, 0, 0, \ldots, 0)$;
  - Borda $(m-1, m-2, \ldots, 1, 0)$;
  - veto $(1, 1, \ldots, 1, 0)$.

Some social choice rules

- The scoring rule defined by a fixed weight vector $(w_1, \ldots, w_m)$ gives w_i points to each candidate for each voter listing it in ith position. Highest total score wins. Special cases:
  - plurality $(1, 0, 0, \ldots, 0)$;
  - Borda $(m-1, m-2, \ldots, 1, 0)$;
  - veto $(1, 1, \ldots, 1, 0)$.
- Condorcet rules: if there is a Condorcet winner (preferred to each other alternative by some majority of voters), choose it. Otherwise choose something else. Most famous is Copeland rule (chess scoring).

Consensus

- We define a set of profiles for which the result is “obvious”.

Consensus

- We define a set of profiles for which the result is “obvious”.
- Example: if every voter has the same preference, this is the social preference.

Consensus

- We define a set of profiles for which the result is "obvious".
- Example: if every voter has the same preference, this is the social preference.
- We define a consensus to be a mapping $\mathcal{K}$ defined on a subset D of profiles, which always returns a unique social choice. We denote by $\mathcal{K}_a$ the subset on which a is that choice.

Consensus

- We define a set of profiles for which the result is "obvious".
- Example: if every voter has the same preference, this is the social preference.
- We define a consensus to be a mapping $\mathcal{K}$ defined on a subset D of profiles, which always returns a unique social choice. We denote by $\mathcal{K}_a$ the subset on which a is that choice.
- This is strongly related to various axioms (e.g. unanimity) and to the topic of domain restrictions.

Examples: consensus

- The strong unanimity consensus $\mathbf{S}$: all voters have the same preference order; (the top of) this order is chosen.

Examples: consensus

- The strong unanimity consensus $\mathbf{S}$: all voters have the same preference order; (the top of) this order is chosen.
- Weak unanimity consensus $\mathbf{W}$: all voters have the same top-ranked candidate; it is chosen.

Examples: consensus

- The strong unanimity consensus $\mathbf{S}$: all voters have the same preference order; (the top of) this order is chosen.
- Weak unanimity consensus $\mathbf{W}$: all voters have the same top-ranked candidate; it is chosen.
- Condorcet consensus $\mathbf{C}$: there is a Condorcet winner (majority tournament has a source node); it is chosen. This is more controversial than the previous ones.

Examples: consensus

- The strong unanimity consensus $\mathbf{S}$: all voters have the same preference order; (the top of) this order is chosen.
- Weak unanimity consensus $\mathbf{W}$: all voters have the same top-ranked candidate; it is chosen.
- Condorcet consensus $\mathbf{C}$: there is a Condorcet winner (majority tournament has a source node); it is chosen. This is more controversial than the previous ones.
- Condorcet ranking consensus $\mathbf{C}^*$: the majority tournament is transitive.

Examples: consensus

- The strong unanimity consensus $\mathbf{S}$: all voters have the same preference order; (the top of) this order is chosen.
- Weak unanimity consensus $\mathbf{W}$: all voters have the same top-ranked candidate; it is chosen.
- Condorcet consensus $\mathbf{C}$: there is a Condorcet winner (majority tournament has a source node); it is chosen. This is more controversial than the previous ones.
- Condorcet ranking consensus $\mathbf{C}^*$: the majority tournament is transitive.
- Many other choices are possible, e.g. single peaked preferences, Lorenz consensus.

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.
  - $d(x, y) \geq 0$

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.
  - $d(x, y) \geq 0$
  - $d(x, x) = 0$

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.
  - $d(x, y) \geq 0$
  - $d(x, x) = 0$
  - $d(x, z) \leq d(x, y) + d(x, z)$

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.
  - $d(x, y) \geq 0$
  - $d(x, x) = 0$
  - $d(x, z) \leq d(x, y) + d(x, z)$
- A distance that distinguishes points ($d(x, y) = 0 \Rightarrow x = y$) is called a quasimetric. If it is also symmetric ($d(x, y) = d(y, x)$) it is a metric.

Distances

- By a distance (or hemimetric) on $\mathcal{E}$ we mean a function $d : \mathcal{E} \times \mathcal{E} \to \mathbb{R} \cup \{+\infty\}$ that satisfies the following identities.
  - $d(x, y) \geq 0$
  - $d(x, x) = 0$
  - $d(x, z) \leq d(x, y) + d(x, z)$
- A distance that distinguishes points $(d(x, y) = 0 \Rightarrow x = y)$ is called a quasimetric. If it is also symmetric $(d(x, y) = d(y, x))$ it is a metric.
- A distance d is standard if $d(E, E') = \infty$ whenever E and E' have different sets of voters or candidates.

Examples: distance

- Define a graph by stipulating edges between some pairs of elections. Then define $d(E, E')$ to be the length of a shortest path (geodesic) in this graph. This includes:

distance	create an edge when we
Hamming d_H	change one preference order
Kemeny d_K	swap two candidates in a pref. order
insertion d_{ins}	add a preference order
deletion d_{del}	delete a preference order
tournament d_{RT}	reverse arrow in majority tournament

Examples: distance

- Define a graph by stipulating edges between some pairs of elections. Then define $d(E, E')$ to be the length of a shortest path (geodesic) in this graph. This includes:

distance	create an edge when we
Hamming d_H	change one preference order
Kemeny d_K	swap two candidates in a pref. order
insertion d_{ins}	add a preference order
deletion d_{del}	delete a preference order
tournament d_{RT}	reverse arrow in majority tournament

- Votewise distances can be formed by using any distance on $L(A)$ and combining the distances of each component with a norm on $\mathbb{R}^n$, such as the ℓ^p norm.

Distance rationalizability

- Let $\mathcal{K}$ be a consensus and d a distance on $\mathcal{E}$. The rule $R := \mathcal{R}(\mathcal{K}, d)$ is defined by

$$R(E) := \{R(C) \mid C \in \arg \min_{E' \in D(\mathcal{K})} d(E, E')\}.$$

Distance rationalizability

- Let $\mathcal{K}$ be a consensus and d a distance on $\mathcal{E}$. The rule $R := \mathcal{R}(\mathcal{K}, d)$ is defined by

$$R(E) := \{R(C) \mid C \in \arg \min_{E' \in D(\mathcal{K})} d(E, E')\}.$$

- In other words, the set of winners consists of the consensus choice from each consensus election that minimizes the distance to E. We say that R is distance rationalizable (DR) with respect to $(\mathcal{K}, d)$.

Distance rationalizability

- Let $\mathcal{K}$ be a consensus and d a distance on $\mathcal{E}$. The rule $R := \mathcal{R}(\mathcal{K}, d)$ is defined by

$$R(E) := \{R(C) \mid C \in \arg\min_{E' \in D(\mathcal{K})} d(E, E')\}.$$

- In other words, the set of winners consists of the consensus choice from each consensus election that minimizes the distance to E. We say that R is distance rationalizable (DR) with respect to $(\mathcal{K}, d)$.
- Every non-imposed rule can be represented in this way, so the point is to choose "good" $\mathcal{K}$ and d.

Examples: DR rules

$\mathcal{K}$ d	d_H	d_K	d_{RT}	d_{ins}	d_{del}
$\mathbf{S}$	MR	Kemeny	Copeland	undef	MR
$\mathbf{W}$	plurality	Borda	Copeland	undef	plurality
$\mathbf{C}$	VRR	Dodgson	Copeland	maximin	Young
$\mathbf{C}_*$			Slater		

Table: some rules in the DR framework:

Here

- VRR = "voter replacement rule" (Elkind, Faliszewski, Slinko; Soc Choice Welf 2012)
- MR = "modal ranking rule" (Caragiannis, Procaccia, Shah; AAAI 2014).

Anonymity

- A rule is anonymous if it is symmetric under permutations of the voter set. On each profile representing the same voting situation, R makes the same social choice, so R naturally yields a (quotient) rule $\overline{R}$ on $\mathcal{V}$.

Anonymity

- A rule is anonymous if it is symmetric under permutations of the voter set. On each profile representing the same voting situation, R makes the same social choice, so R naturally yields a (quotient) rule $\overline{R}$ on $\mathcal{V}$.
- Similarly a consensus is anonymous if and only if it yields a consensus on $\mathcal{V}$. How can we represent R directly on $\mathcal{V}$? We want something like $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \overline{d}) = \overline{\mathcal{R}(\mathcal{K}, d)}$.

Anonymity

- A rule is anonymous if it is symmetric under permutations of the voter set. On each profile representing the same voting situation, R makes the same social choice, so R naturally yields a (quotient) rule $\overline{R}$ on $\mathcal{V}$.
- Similarly a consensus is anonymous if and only if it yields a consensus on $\mathcal{V}$. How can we represent R directly on $\mathcal{V}$? We want something like $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \overline{d}) = \overline{\mathcal{R}(\mathcal{K}, d)}$.
- Abstractly, we want the quotient distance induced by the map from profiles to voting situations. In general quotient distances are counterintuitive and don't have a simple formula.

Example: anonymous distance

- Let x be a voting situation with 2 abc voters and 3 bac voters, while y has 2 bac voters and 3 cba voters.

Example: anonymous distance

- Let x be a voting situation with 2 abc voters and 3 bac voters, while y has 2 bac voters and 3 cba voters.
- Consider the Hamming distance $d := d_H$, which is anonymous.

Example: anonymous distance

- Let x be a voting situation with 2 abc voters and 3 bac voters, while y has 2 bac voters and 3 cba voters.
- Consider the Hamming distance $d := d_H$, which is anonymous.
- Different choices of representing profiles E, E' for x and y yield different answers for $d(E, E')$. For example, choosing profiles $(abc, abc, bac, bac, bac)$ and $(bac, bac, cba, cba, cba)$ yields Hamming distance 5, while choosing $(abc, abc, bac, bac, bac)$ and $(cba, cba, cba, bac, bac)$ yields Hamming distance 3.

Example: anonymous distance

- Let x be a voting situation with 2 abc voters and 3 bac voters, while y has 2 bac voters and 3 cba voters.
- Consider the Hamming distance $d := d_H$, which is anonymous.
- Different choices of representing profiles E, E' for x and y yield different answers for $d(E, E')$. For example, choosing profiles $(abc, abc, bac, bac, bac)$ and $(bac, bac, cba, cba, cba)$ yields Hamming distance 5, while choosing $(abc, abc, bac, bac, bac)$ and $(cba, cba, cba, bac, bac)$ yields Hamming distance 3.
- In this example we can check that 3 is the minimum possible value over all representing profiles.

DR for anonymous rules

Theorem

Suppose that $\mathcal{K}$ is anonymous and d is standard and anonymous. Then

- $R := \mathcal{R}(\mathcal{K}, d)$ *is anonymous.*

DR for anonymous rules

Theorem

Suppose that $\mathcal{K}$ is anonymous and d is standard and anonymous. Then

- $R := \mathcal{R}(\mathcal{K}, d)$ *is anonymous.*
- *The quotient distance is given by*

$$\overline{d}(x, y) := \min_{E \mapsto x, E' \mapsto y} d(E, E').$$

DR for anonymous rules

Theorem

Suppose that $\mathcal{K}$ is anonymous and d is standard and anonymous. Then

- $R := \mathcal{R}(\mathcal{K}, d)$ *is anonymous.*
- *The quotient distance is given by*

$$\overline{d}(x, y) := \min_{E \mapsto x, E' \mapsto y} d(E, E').$$

- $\overline{\mathcal{R}(\mathcal{K}, d)} = \overline{\mathcal{R}}(\overline{\mathcal{K}}, \overline{d})$.

Connection with optimal transportation

- The problem of optimal transportation originated with Monge in the 19th century and was generalized by Kantorovich.

Connection with optimal transportation

- The problem of optimal transportation originated with Monge in the 19th century and was generalized by Kantorovich.
- In the discrete case, it amounts to minimizing the cost of transferring mass from one histogram to another while incurring the minimal cost.

Connection with optimal transportation

- The problem of optimal transportation originated with Monge in the 19th century and was generalized by Kantorovich.
- In the discrete case, it amounts to minimizing the cost of transferring mass from one histogram to another while incurring the minimal cost.
- The minimum can be computed via a linear program.

Connection with optimal transportation

- The problem of optimal transportation originated with Monge in the 19th century and was generalized by Kantorovich.
- In the discrete case, it amounts to minimizing the cost of transferring mass from one histogram to another while incurring the minimal cost.
- The minimum can be computed via a linear program.
- In the anonymous and standard case, the distance $\overline{d}$ is the solution of an optimal transportation problem, because we must move voter mass between types of voters while incurring the minimal cost (distance). The assumption that d is standard is equivalent to saying that conservation of mass (number of voters) holds.

Optimal transportation picture

earthmover.png

Homogeneity

- A rule is homogeneous if it gives the same result when we replicate each voter the same number of times.

Homogeneity

- A rule is homogeneous if it gives the same result when we replicate each voter the same number of times.
- An anonymous homogeneous rule can therefore by described by how it acts on the simplex Δ.

Homogeneity

- A rule is homogeneous if it gives the same result when we replicate each voter the same number of times.
- An anonymous homogeneous rule can therefore by described by how it acts on the simplex Δ.
- There is an obvious quotient map from $\mathcal{V}$ to Δ ("divide by n") and we want results analogous to the ones for the anonymous case above. However it is a bit trickier because we don't have such nice sufficient conditions for homogeneity of $\mathcal{R}(\mathcal{K}, d)$.

Homogeneity

- A rule is homogeneous if it gives the same result when we replicate each voter the same number of times.
- An anonymous homogeneous rule can therefore by described by how it acts on the simplex Δ.
- There is an obvious quotient map from $\mathcal{V}$ to Δ ("divide by n") and we want results analogous to the ones for the anonymous case above. However it is a bit trickier because we don't have such nice sufficient conditions for homogeneity of $\mathcal{R}(\mathcal{K}, d)$.
- If we *assume* that $\mathcal{R}(\mathcal{K}, d)$ is homogeneous, everything works as above and we can operate only on Δ.

Connection with Wasserstein distance

The simplex Δ can be identified with the space of probability distributions on the finite set $L(A)$. There is a famous family of distances on such spaces (for not necessarily finite underlying sets), called the Wasserstein distances, defined by

$$d_W^p(x, y) = \inf E[d(X, Y)^p]$$

where the infimum is over all couplings (pairs of random variables X, Y with marginal distributions x, y).

Theorem

Let d be an l^p-votewise distance. Then $\overline{d} = d_W^p$, the p-Wasserstein distance based on d.

Connection with Minkowski geometry

- When $p = 1$ the Wasserstein distance is called the earth mover distance. It is widely used in image processing.

Connection with Minkowski geometry

- When $p = 1$ the Wasserstein distance is called the earth mover distance. It is widely used in image processing.
- The earth mover distance is induced by a norm. Thus in this case we are looking at a Minkowski space.

Connection with Minkowski geometry

- When $p = 1$ the Wasserstein distance is called the earth mover distance. It is widely used in image processing.
- The earth mover distance is induced by a norm. Thus in this case we are looking at a Minkowski space.
- The simplest special case is when $d = d_H$, when d_W^1 is half the ℓ^1 distance on Δ.

Connection with Minkowski geometry

- When $p = 1$ the Wasserstein distance is called the earth mover distance. It is widely used in image processing.
- The earth mover distance is induced by a norm. Thus in this case we are looking at a Minkowski space.
- The simplest special case is when $d = d_H$, when d_W^1 is half the ℓ^1 distance on Δ.
- For $p > 1$ it is not induced by a norm, but by something that is probably geometrically fairly nice.

Decisiveness of DR rules

- Ideally a voting rule has few profiles where it does not give a unique result (otherwise it is just “passing the buck” to the tiebreaking rule).

Decisiveness of DR rules

- Ideally a voting rule has few profiles where it does not give a unique result (otherwise it is just "passing the buck" to the tiebreaking rule).
- The definition shows that rules defined in the DR framework can have many tied profiles. This is because the minimum distance may not be uniquely attained.

Decisiveness of DR rules

- Ideally a voting rule has few profiles where it does not give a unique result (otherwise it is just "passing the buck" to the tiebreaking rule).
- The definition shows that rules defined in the DR framework can have many tied profiles. This is because the minimum distance may not be uniquely attained.
- Sometimes this happens just because the distance does not distinguish points well. For example, Copeland's rule has a large tied region in Δ.

Decisiveness of DR rules

- Ideally a voting rule has few profiles where it does not give a unique result (otherwise it is just "passing the buck" to the tiebreaking rule).
- The definition shows that rules defined in the DR framework can have many tied profiles. This is because the minimum distance may not be uniquely attained.
- Sometimes this happens just because the distance does not distinguish points well. For example, Copeland's rule has a large tied region in Δ.
- More interestingly, sometimes it happens because of geometric properties of the unit ball of the distance.

Geometric background

- Choose finitely many sites (subsets of the ambient space).

Geometric background

- Choose finitely many sites (subsets of the ambient space).
- The bisector of sites a and b is the set of points equidistant from a and b: $d(x, a) = d(x, b)$. Here we use the usual definition of distance to a set: $d(x, S) = \inf_{s \in S} d(x, s)$.

Geometric background

- Choose finitely many sites (subsets of the ambient space).
- The bisector of sites a and b is the set of points equidistant from a and b: $d(x,a) = d(x,b)$. Here we use the usual definition of distance to a set: $d(x,S) = \inf_{s \in S} d(x,s)$.
- The Voronoi region associated to a site a is the set of points for whom a is the closest site: $d(x,a) \leq d(x,b)$ for all $b \in S$. The boundary of a Voronoi region is contained in a union of bisectors.

Geometric background

- Choose finitely many sites (subsets of the ambient space).
- The bisector of sites a and b is the set of points equidistant from a and b: $d(x,a) = d(x,b)$. Here we use the usual definition of distance to a set: $d(x,S) = \inf_{s \in S} d(x,s)$.
- The Voronoi region associated to a site a is the set of points for whom a is the closest site: $d(x,a) \leq d(x,b)$ for all $b \in S$. The boundary of a Voronoi region is contained in a union of bisectors.
- In our situation, the sets $\overline{\mathcal{K}}_a$ are the sites. A point $x \in \Delta$ lies on a bisector if and only if $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \bar{d})$ does not have a unique winner at x. The interiors of the Voronoi regions are those places where a unique winner is defined.

Geometric background

- Choose finitely many sites (subsets of the ambient space).
- The bisector of sites a and b is the set of points equidistant from a and b: $d(x,a) = d(x,b)$. Here we use the usual definition of distance to a set: $d(x,S) = \inf_{s\in S} d(x,s)$.
- The Voronoi region associated to a site a is the set of points for whom a is the closest site: $d(x,a) \leq d(x,b)$ for all $b \in S$. The boundary of a Voronoi region is contained in a union of bisectors.
- In our situation, the sets $\overline{\mathcal{K}}_a$ are the sites. A point $x \in \Delta$ lies on a bisector if and only if $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \bar{d})$ does not have a unique winner at x. The interiors of the Voronoi regions are those places where a unique winner is defined.
- Voronoi theory has hugely many applications in science.

Minkowski geometry: ℓ^2 vs ℓ^1

Euclidean_Voronoi_Diagram.png | Manhattan_Voronoi_Diagram.png

Figure: ℓ^2 (left) vs ℓ^1 (right) Voronoi diagram. Source: Wikipedia.

Minkowski geometry: ℓ^1 vs ℓ^2

- The unit ball of ℓ^2 is strictly convex, which implies that (if sites are isolated points) bisectors are hyperplanes and Voronoi regions are convex polyhedra.

Minkowski geometry: ℓ^1 vs ℓ^2

- The unit ball of ℓ^2 is strictly convex, which implies that (if sites are isolated points) bisectors are hyperplanes and Voronoi regions are convex polyhedra.
- The unit ball of ℓ^1 is not strictly convex. The bisector of two points can have nonempty interior, and Voronoi regions can be nonconvex.

Minkowski geometry: ℓ^1 vs ℓ^2

- The unit ball of ℓ^2 is strictly convex, which implies that (if sites are isolated points) bisectors are hyperplanes and Voronoi regions are convex polyhedra.
- The unit ball of ℓ^1 is not strictly convex. The bisector of two points can have nonempty interior, and Voronoi regions can be nonconvex.
- It is NP-hard to determine whether the bisector under ℓ^1 of two points of Δ contains an open ball. We suspect this is also the case for every ℓ^1-votewise metric.

ℓ^1 large bisector

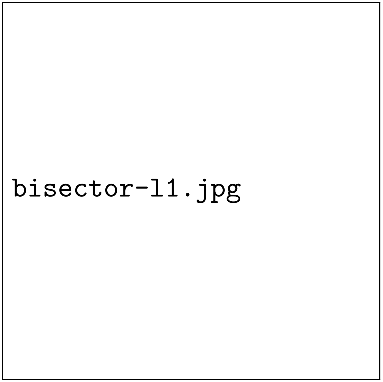

Figure: Large bisector in ℓ^1. Source: http://www.ams.org/samplings/feature-column/fcarc-taxi.

Dichotomy: ℓ^1-votewise distance

- Let d be an l^1-votewise metric and let $\mathcal{K}$ be an anonymous and homogeneous consensus (alternatively $\overline{\mathcal{K}}$ is a consensus in Δ).

Dichotomy: ℓ^1-votewise distance

- Let d be an l^1-votewise metric and let $\mathcal{K}$ be an anonymous and homogeneous consensus (alternatively $\overline{\mathcal{K}}$ is a consensus in Δ).
- Thm: There exists such a $\overline{\mathcal{K}}$, consisting of finitely many points, such that $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \overline{d})$ has a tied region with nonempty interior.

Dichotomy: ℓ^1-votewise distance

- Let d be an l^1-votewise metric and let $\mathcal{K}$ be an anonymous and homogeneous consensus (alternatively $\overline{\mathcal{K}}$ is a consensus in Δ).
- Thm: There exists such a $\overline{\mathcal{K}}$, consisting of finitely many points, such that $\overline{\mathcal{R}}(\overline{\mathcal{K}}, \overline{d})$ has a tied region with nonempty interior.
- Thm: However, if $\overline{\mathcal{K}}$ lies on the boundary of Δ, then $\mathcal{R}(\mathcal{K}, d)$ is a hyperplane rule.

Hyperplane rules

- Mossel-Procaccia-Racz (JAIR 2013) defined hyperplane rules as rules defined on the voting simplex (hence all are anonymous and homogeneous).

Hyperplane rules

- Mossel-Procaccia-Racz (JAIR 2013) defined hyperplane rules as rules defined on the voting simplex (hence all are anonymous and homogeneous).
- The winner is constant on each chamber of the preference simplex, the chambers being defined by (removal of) a finite set of hyperplanes.

Hyperplane rules

- Mossel-Procaccia-Racz (JAIR 2013) defined hyperplane rules as rules defined on the voting simplex (hence all are anonymous and homogeneous).
- The winner is constant on each chamber of the preference simplex, the chambers being defined by (removal of) a finite set of hyperplanes.
- These are the same as the generalized scoring rules of Xia and Conitzer, and can be characterized axiomatically by finite local consistency.

Hyperplane rules

- Mossel-Procaccia-Racz (JAIR 2013) defined hyperplane rules as rules defined on the voting simplex (hence all are anonymous and homogeneous).
- The winner is constant on each chamber of the preference simplex, the chambers being defined by (removal of) a finite set of hyperplanes.
- These are the same as the generalized scoring rules of Xia and Conitzer, and can be characterized axiomatically by finite local consistency.
- Most known rules are hyperplane rules. MPR define Copeland's rule to be one, but this is not convincing. In fact Copeland has a very large tied region.

Hyperplane rules

- Mossel-Procaccia-Racz (JAIR 2013) defined hyperplane rules as rules defined on the voting simplex (hence all are anonymous and homogeneous).
- The winner is constant on each chamber of the preference simplex, the chambers being defined by (removal of) a finite set of hyperplanes.
- These are the same as the generalized scoring rules of Xia and Conitzer, and can be characterized axiomatically by finite local consistency.
- Most known rules are hyperplane rules. MPR define Copeland's rule to be one, but this is not convincing. In fact Copeland has a very large tied region.
- The connection between hyperplane rules and distance rationalization is not yet clear.

Votewise minimizer property

- Sometimes the minimum distance is attained in a nice way. For example, the Kemeny distance of E to $\mathbf{W}_a$ is found by pushing up a in each preference order in the profile until it reaches the top.

Votewise minimizer property

- Sometimes the minimum distance is attained in a nice way. For example, the Kemeny distance of E to $\mathbf{W}_a$ is found by pushing up a in each preference order in the profile until it reaches the top.
- However, this doesn't always work. For example with $\mathbf{C}$, the Kemeny distance is not minimized in this way.

Votewise minimizer property

- Sometimes the minimum distance is attained in a nice way. For example, the Kemeny distance of E to $\mathbf{W}_a$ is found by pushing up a in each preference order in the profile until it reaches the top.
- However, this doesn't always work. For example with $\mathbf{C}$, the Kemeny distance is not minimized in this way.
- Positive result: if $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise then $\mathcal{R}(\mathcal{K}, d)$ is a hyperplane rule.

Votewise minimizer property

- Sometimes the minimum distance is attained in a nice way. For example, the Kemeny distance of E to $\mathbf{W}_a$ is found by pushing up a in each preference order in the profile until it reaches the top.
- However, this doesn't always work. For example with $\mathbf{C}$, the Kemeny distance is not minimized in this way.
- Positive result: if $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise then $\mathcal{R}(\mathcal{K}, d)$ is a hyperplane rule.
- This recaptures scoring rules and Kemeny's rule, for example, and it shows that Dodgson's rule cannot be distance rationalized in this way.

Conclusions

- If we want a rule that is decisive and defined in the (anonymous and homogeneous) DR framework, we should restrict to quasimetrics. How to restrict the consensus?

Conclusions

- If we want a rule that is decisive and defined in the (anonymous and homogeneous) DR framework, we should restrict to quasimetrics. How to restrict the consensus?
- The Condorcet consensus is less compelling than the weak unanimity consensus. This is reflected in the large tied sets obtained above.

Conclusions

- If we want a rule that is decisive and defined in the (anonymous and homogeneous) DR framework, we should restrict to quasimetrics. How to restrict the consensus?
- The Condorcet consensus is less compelling than the weak unanimity consensus. This is reflected in the large tied sets obtained above.
- If $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise, many good things happen. Is this class big enough to be interesting? Are there interesting examples with $p > 1$?

Future work

- Is the homogenized version of Dodgson's rule a hyperplane rule? If not, is its tied region small in some sense?

Future work

- Is the homogenized version of Dodgson's rule a hyperplane rule? If not, is its tied region small in some sense?
- Find better sufficient conditions for homogeneity of R. The best we have so far is essentially that $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise.

Future work

- Is the homogenized version of Dodgson's rule a hyperplane rule? If not, is its tied region small in some sense?
- Find better sufficient conditions for homogeneity of R. The best we have so far is essentially that $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise.
- Can we obtain nice results for important subclasses of the set of anonymous and homogeneous DR rules? For example, maximum likelihood estimators (simple rank scoring functions) and mean proximity rules have been recently studied in detail.

Future work

- Is the homogenized version of Dodgson's rule a hyperplane rule? If not, is its tied region small in some sense?
- Find better sufficient conditions for homogeneity of R. The best we have so far is essentially that $\mathcal{K} \in \{\mathbf{S}, \mathbf{W}\}$ and d is ℓ^p-votewise.
- Can we obtain nice results for important subclasses of the set of anonymous and homogeneous DR rules? For example, maximum likelihood estimators (simple rank scoring functions) and mean proximity rules have been recently studied in detail.
- What interesting rules are obtained by using arbitrary metrics on Δ (not necessarily derived in a natural way from a metric on $\mathcal{E}$)? There are many such statistical distances such as Kullback-Liebler.