Distance rationalizability of voting rules

Mark C. Wilson
www.cs.auckland.ac.nz/~mcw/

Department of Computer Science
University of Auckland

CMSS Seminar, 2013-08-20

Overview

- This talk gives background for understanding Benjamin Hadjibeyli's talk next week.

Overview

- This talk gives background for understanding Benjamin Hadjibeyli's talk next week.
- We concentrate on definitions and examples, and skip proofs.

Overview

- This talk gives background for understanding Benjamin Hadjibeyli's talk next week.
- We concentrate on definitions and examples, and skip proofs.
- The subject of distance rationalization has been studied recently in several papers by Elkind, Faliszewski, and Slinko.

Overview

- This talk gives background for understanding Benjamin Hadjibeyli's talk next week.
- We concentrate on definitions and examples, and skip proofs.
- The subject of distance rationalization has been studied recently in several papers by Elkind, Faliszewski, and Slinko.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.
  - Rationalizations of Condorcet-consistent rules via distances of hamming type. SCW 2012.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.
  - Rationalizations of Condorcet-consistent rules via distances of hamming type. SCW 2012.
  - Homogeneity and monotonicity of distance-rationalizable voting rules. AAMAS 2011.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.
  - Rationalizations of Condorcet-consistent rules via distances of hamming type. SCW 2012.
  - Homogeneity and monotonicity of distance-rationalizable voting rules. AAMAS 2011.
  - Good Rationalizations of Voting Rules. AAAI 2010.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.
  - Rationalizations of Condorcet-consistent rules via distances of hamming type. SCW 2012.
  - Homogeneity and monotonicity of distance-rationalizable voting rules. AAMAS 2011.
  - Good Rationalizations of Voting Rules. AAAI 2010.
- I. Caragiannis, A. Procaccia, N. Shah. When do noisy votes reveal the truth? EC 2013.

A short reference list

- T. Meskanen, H. Nurmi. Closeness counts in social choice. 2008.
- E. Elkind, P. Faliszewski, A. Slinko.
  - (with S. Obraztsova) On swap-distance geometry of voting rules. AAMAS 2013.
  - Rationalizations of Condorcet-consistent rules via distances of hamming type. SCW 2012.
  - Homogeneity and monotonicity of distance-rationalizable voting rules. AAMAS 2011.
  - Good Rationalizations of Voting Rules. AAAI 2010.
- I. Caragiannis, A. Procaccia, N. Shah. When do noisy votes reveal the truth? EC 2013.

Basic setup

- As usual in social choice theory, we have a finite set C of candidates, of size m, and a finite set V of voters, of size n.

Basic setup

- As usual in social choice theory, we have a finite set C of candidates, of size m, and a finite set V of voters, of size n.
- Each voter submits a strict preference order over the candidates. The set of possible preference orders is the set $L(C)$ of permutations of C.

Basic setup

- As usual in social choice theory, we have a finite set C of candidates, of size m, and a finite set V of voters, of size n.
- Each voter submits a strict preference order over the candidates. The set of possible preference orders is the set $L(C)$ of permutations of C.
- A profile is a map from V to $L(C)$, stating each voter's preference order. There are $(m!)^n$ possible profiles, and the set of them we denote $\mathcal{P}$.

Symmetry assumptions

- A function is anonymous if it commutes with all permutations of the voters. In other words, voter identities are irrelevant.

Symmetry assumptions

- A function is anonymous if it commutes with all permutations of the voters. In other words, voter identities are irrelevant.
- A function is neutral if it commutes with all permutations of the candidates. In other words, each candidate is treated equally.

Symmetry assumptions

- A function is anonymous if it commutes with all permutations of the voters. In other words, voter identities are irrelevant.
- A function is neutral if it commutes with all permutations of the candidates. In other words, each candidate is treated equally.
- Most voting rules are both anonymous and neutral.

Compressing the data

- By anonymity, the set of profiles can be reduced to the set of equivalence classes, which are multisets of size $m!$ and weight n. If we order the candidates, then such a multiset is represented by a point of $\mathbb{N}^{m!}$. There are $\binom{n+m!-1}{n}$ of them of weight n.

Compressing the data

- By anonymity, the set of profiles can be reduced to the set of equivalence classes, which are multisets of size $m!$ and weight n. If we order the candidates, then such a multiset is represented by a point of $\mathbb{N}^{m!}$. There are $\binom{n+m!-1}{n}$ of them of weight n.
- By neutrality, it can be further compressed, by considering only roots, equivalence classes under candidate permutation. This divides by another factor of $m!$.

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:
  - (strong unanimity $\mathcal{S}$) if all voters submit π, elect the top element;

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:
  - (strong unanimity $\mathcal{S}$) if all voters submit π, elect the top element;
  - (unanimity $\mathcal{U}$) if all voters agree that a is top-ranked, elect a;

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:
  - (strong unanimity $\mathcal{S}$) if all voters submit π, elect the top element;
  - (unanimity $\mathcal{U}$) if all voters agree that a is top-ranked, elect a;
  - (majority principle) if a majority of voters rank a top, elect a;

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:
  - (strong unanimity $\mathcal{S}$) if all voters submit π, elect the top element;
  - (unanimity $\mathcal{U}$) if all voters agree that a is top-ranked, elect a;
  - (majority principle) if a majority of voters rank a top, elect a;
  - (Condorcet principle) if a wins the majority tournament, elect a.

Consensus

- Many profiles have an obvious winner, for example if all voters submit the same preference order.
- Various consensus axioms have been proposed. For example:
  - (strong unanimity $\mathcal{S}$) if all voters submit π, elect the top element;
  - (unanimity $\mathcal{U}$) if all voters agree that a is top-ranked, elect a;
  - (majority principle) if a majority of voters rank a top, elect a;
  - (Condorcet principle) if a wins the majority tournament, elect a.
- Some of these are less controversial than others. I find Condorcet's principle far from compelling.

Distance

- A distance on a set X is a map $d : X \times X \to \mathbb{R}_+ \cup \{+\infty\}$ that satisfies

Together with these axioms, symmetry and positive definiteness make a metric. Distances that are not metrics arise often in this subject.

Distance

- A distance on a set X is a map $d : X \times X \to \mathbb{R}_+ \cup \{+\infty\}$ that satisfies
  - nonnegativity: $d(x, x) \geq 0$

Together with these axioms, symmetry and positive definiteness make a metric. Distances that are not metrics arise often in this subject.

Distance

- A distance on a set X is a map $d : X \times X \to \mathbb{R}_+ \cup \{+\infty\}$ that satisfies
  - nonnegativity: $d(x, x) \geq 0$
  - triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

Together with these axioms, symmetry and positive definiteness make a metric. Distances that are not metrics arise often in this subject.

Distances on profiles

- Distances on $L(C)$:

Distances on profiles

- Distances on $L(C)$:
  - discrete metric

Distances on profiles

- Distances on $L(C)$:
  - discrete metric
  - swap metric

Distances on profiles

- Distances on $L(C)$:
  - discrete metric
  - swap metric
  - Spearman's footrule

Distances on profiles

- Distances on $L(C)$:
  - discrete metric
  - swap metric
  - Spearman's footrule
- Given a distance on $L(C)$, we can extend to a ℓ^1-votewise distance on $\mathcal{P}$ by defining $d(\pi, \sigma) = \sum_{i=1}^{n} d(\pi_i, \sigma_i)$ when $|\pi| = |\sigma|$, and $+\infty$ otherwise.

Distances on profiles

- Distances on $L(C)$:
  - discrete metric
  - swap metric
  - Spearman's footrule
- Given a distance on $L(C)$, we can extend to a ℓ^1-votewise distance on $\mathcal{P}$ by defining $d(\pi, \sigma) = \sum_{i=1}^{n} d(\pi_i, \sigma_i)$ when $|\pi| = |\sigma|$, and $+\infty$ otherwise.
- If we start with the discrete metric, we get the Hamming metric: how many changes of voters are needed. If we start with the swap metric we get the Kemeny metric.

Distances on profiles

- Distances on $L(C)$:
  - discrete metric
  - swap metric
  - Spearman's footrule
- Given a distance on $L(C)$, we can extend to a ℓ^1-votewise distance on $\mathcal{P}$ by defining $d(\pi, \sigma) = \sum_{i=1}^{n} d(\pi_i, \sigma_i)$ when $|\pi| = |\sigma|$, and $+\infty$ otherwise.
- If we start with the discrete metric, we get the Hamming metric: how many changes of voters are needed. If we start with the swap metric we get the Kemeny metric.
- Tournament distances: for example the ℓ^1-norm of the reduced weighted adjacency matrix of the majority tournament.

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:
  - Kemeny metric, strong unanimity yields Kemeny's rule

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:
  - Kemeny metric, strong unanimity yields Kemeny's rule
  - Kemeny metric, weak unanimity yields Borda's rule

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:
  - Kemeny metric, strong unanimity yields Kemeny's rule
  - Kemeny metric, weak unanimity yields Borda's rule
  - Kemeny metric, Condorcet set yields Dodgson's rule

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:
  - Kemeny metric, strong unanimity yields Kemeny's rule
  - Kemeny metric, weak unanimity yields Borda's rule
  - Kemeny metric, Condorcet set yields Dodgson's rule
  - Hamming metric, weak unanimity yields plurality rule

Distance rationalization

- Define the rule $R(\mathcal{K}, d)$ by electing the winner at π to be the winner in the closest element of $\mathcal{K}$ to π when measured by d.
- Every rule can be represented in this way, so the important thing is to choose "good" $\mathcal{K}$ and d.
- Some examples:
  - Kemeny metric, strong unanimity yields Kemeny's rule
  - Kemeny metric, weak unanimity yields Borda's rule
  - Kemeny metric, Condorcet set yields Dodgson's rule
  - Hamming metric, weak unanimity yields plurality rule
  - ℓ^1-tournament, Condorcet set yields Copeland's rule

New rules

- Using the above framework we can derive new rules.

New rules

- Using the above framework we can derive new rules.
  - Hamming metric, Condorcet set yields "voter replacement rule" (EFS)

New rules

- Using the above framework we can derive new rules.
  - Hamming metric, Condorcet set yields "voter replacement rule" (EFS)
  - Hamming metric, strong unanimity yields "plurality ranking rule"

New rules

- Using the above framework we can derive new rules.
  - Hamming metric, Condorcet set yields "voter replacement rule" (EFS)
  - Hamming metric, strong unanimity yields "plurality ranking rule"
  - There must be many more that are worth studying.

Properties of rules

- If distance and consensus satisfy anonymity (neutrality) then so does the rule. Proof is straightforward.

Properties of rules

- If distance and consensus satisfy anonymity (neutrality) then so does the rule. Proof is straightforward.
- Homogeneity does not follow so easily. Many rules are in fact homogeneous, but proving it in this framework still seems hard.

Properties of rules

- If distance and consensus satisfy anonymity (neutrality) then so does the rule. Proof is straightforward.
- Homogeneity does not follow so easily. Many rules are in fact homogeneous, but proving it in this framework still seems hard.
- Monotonicity is also tricky, but some positive results exist.

Properties of rules

- If distance and consensus satisfy anonymity (neutrality) then so does the rule. Proof is straightforward.
- Homogeneity does not follow so easily. Many rules are in fact homogeneous, but proving it in this framework still seems hard.
- Monotonicity is also tricky, but some positive results exist.
- Complexity: Note that the Kemeny rule has an NP-hard winner determination problem, but scoring rules can be dealt with in polynomial time.

Errors

- Condorcet considered the model where there is a true answer and each voter may err in estimating it. A voting rule is an aggregator that chooses the correct decision.

Errors

- Condorcet considered the model where there is a true answer and each voter may err in estimating it. A voting rule is an aggregator that chooses the correct decision.
- This viewpoint was common in mediaeval church applications of voting.

Errors

- Condorcet considered the model where there is a true answer and each voter may err in estimating it. A voting rule is an aggregator that chooses the correct decision.
- This viewpoint was common in mediaeval church applications of voting.
- Condorcet essentially used the Mallows model of noise (error): each voter ranks each pair of alternatives correctly with probability $p > 1/2$, and all choices are independent.

Errors

- Condorcet considered the model where there is a true answer and each voter may err in estimating it. A voting rule is an aggregator that chooses the correct decision.
- This viewpoint was common in mediaeval church applications of voting.
- Condorcet essentially used the Mallows model of noise (error): each voter ranks each pair of alternatives correctly with probability $p > 1/2$, and all choices are independent.
- Under this model, the probability of observing π' given that π is the true answer is proportional to $q^{d(\sigma,\sigma')}$ where $q = (1-p)/p$ and d is the Kemeny distance.

Errors

- Condorcet considered the model where there is a true answer and each voter may err in estimating it. A voting rule is an aggregator that chooses the correct decision.
- This viewpoint was common in mediaeval church applications of voting.
- Condorcet essentially used the Mallows model of noise (error): each voter ranks each pair of alternatives correctly with probability $p > 1/2$, and all choices are independent.
- Under this model, the probability of observing π' given that π is the true answer is proportional to $q^{d(\sigma,\sigma')}$ where $q = (1-p)/p$ and d is the Kemeny distance.
- If we use an arbitrary distance d and the same construction, this is a generalized Mallows model.

Maximum likelihood estimators

- Finding the maximum likelihood estimator of the correct ranking under the generalized Mallows model is equivalent to computing the social welfare rule $R(\mathcal{S}, d)$.

Maximum likelihood estimators

- Finding the maximum likelihood estimator of the correct ranking under the generalized Mallows model is equivalent to computing the social welfare rule $R(\mathcal{S}, d)$.
- For example, the Kemeny rule gives the MLE under the Condorcet-Mallows model.

Maximum likelihood estimators

- Finding the maximum likelihood estimator of the correct ranking under the generalized Mallows model is equivalent to computing the social welfare rule $R(\mathcal{S}, d)$.
- For example, the Kemeny rule gives the MLE under the Condorcet-Mallows model.
- Scoring rules occur in the same way when we try to compute the correct winner.

Maximum likelihood estimators

- Finding the maximum likelihood estimator of the correct ranking under the generalized Mallows model is equivalent to computing the social welfare rule $R(\mathcal{S}, d)$.
- For example, the Kemeny rule gives the MLE under the Condorcet-Mallows model.
- Scoring rules occur in the same way when we try to compute the correct winner.
- MLE is a stringent condition. What about other statistical estimators?

Aside: other estimators

- Caragiannis, Procaccia and Shah study the sample complexity of voting rules.

Aside: other estimators

- Caragiannis, Procaccia and Shah study the sample complexity of voting rules.
- This is the number of samples required to find the correct ranking with probability at least $1 - \varepsilon$.

Aside: other estimators

- Caragiannis, Procaccia and Shah study the sample complexity of voting rules.
- This is the number of samples required to find the correct ranking with probability at least $1 - \varepsilon$.
- Under Mallows model, Kemeny rule is optimal for all ε, and many rules take only $O(\log(1/\varepsilon))$ samples. But plurality takes exponentially many samples.

Aside: other estimators

- Caragiannis, Procaccia and Shah study the sample complexity of voting rules.
- This is the number of samples required to find the correct ranking with probability at least $1 - \varepsilon$.
- Under Mallows model, Kemeny rule is optimal for all ε, and many rules take only $O(\log(1/\varepsilon))$ samples. But plurality takes exponentially many samples.
- How does this relate to the DR framework?

Aside: other estimators

- Caragiannis, Procaccia and Shah study the sample complexity of voting rules.
- This is the number of samples required to find the correct ranking with probability at least $1 - \varepsilon$.
- Under Mallows model, Kemeny rule is optimal for all ε, and many rules take only $O(\log(1/\varepsilon))$ samples. But plurality takes exponentially many samples.
- How does this relate to the DR framework?
- What about estimators based on methods other than maximum likelihood?

Next week

- How is the DR framework expressed when we compress the data using anonymity and homogeneity?

Next week

- How is the DR framework expressed when we compress the data using anonymity and homogeneity?
- What are the geometric consequences?

Next week

- How is the DR framework expressed when we compress the data using anonymity and homogeneity?
- What are the geometric consequences?
- The answers are interesting!