CSC2556

Lecture 3

Approaches to Voting

Credit for several visuals: Ariel D. Procaccia

Announcement

- No class next week (1/30)

- Please use this time to work on the homework.
  - I'll post the full homework 1 by this weekend.

- You can also start thinking about the project idea!

Approaches to Voting

- What does an approach give us?
  - A way to compare voting rules
  - Hopefully a "uniquely optimal voting rule"

- Axiomatic Approach
- Distance Rationalizability
- Statistical Approach
- Utilitarian Approach
- …

Axiomatic Approach

- Axiom: requirement that the voting rule should behave in a certain way

- Goal: define a set of reasonable axioms, and search for voting rules that satisfy them together
  - Ultimate hope: a unique voting rule satisfies the set of axioms simultaneously!
  - What often happens: no voting rule satisfies the axioms together ☹

Axiomatic Approach

- Weak axioms, satisfied by all popular voting rules
- Unanimity: If all voters have the same top choice, that alternative is the winner.

$$(top(\succ_i) = a \; \forall i \in N) \Rightarrow f(\overrightarrow{\succ}) = a$$

  - An even weaker version requires all rankings to be identical
- Pareto optimality: If all voters prefer a to b, then b is not the winner.

$$(a \succ_i b \; \forall i \in N) \Rightarrow f(\overrightarrow{\succ}) \neq b$$

- Q: *What is the relation between these axioms?*
  - *Pareto optimality* $\Rightarrow$ *Unanimity*

Axiomatic Approach

- **Anonymity:** Permuting votes does not change the winner (i.e., voter identities don't matter).
  - E.g., these two profiles must have the same winner:
    {voter 1: $a \succ b \succ c$, voter 2: $b \succ c \succ a$}
    {voter 1: $b \succ c \succ a$, voter 2: $a \succ b \succ c$}

- **Neutrality:** Permuting alternative names just permutes the winner.
  - E.g., say a wins on {voter 1: $a \succ b \succ c$, voter 2: $b \succ c \succ a$}
  - We permute all names: $a \to b$, $b \to c$, and $c \to a$
  - New profile: {voter 1: $b \succ c \succ a$, voter 2: $c \succ a \succ b$}
  - Then, the new winner must be b.

Axiomatic Approach

- Neutrality is tricky
  - For deterministic rules, it is inconsistent with anonymity!
    - Imagine {voter 1: $a \succ b$, voter 2: $b \succ a$}
    - Without loss of generality, say a wins
    - Imagine a different profile: {voter 1: $b \succ a$, voter 2: $a \succ b$}
      - Neutrality: We just exchanged $a \leftrightarrow b$, so winner is b.
      - Anonymity: We just exchanged the votes, so winner stays a.
  - Typically, we only require neutrality for…
    - Randomized rules: E.g., a rule could satisfy both by choosing a and b as the winner with probability ½ each, on both profiles
    - Deterministic rules that return a set of tied winners: E.g., a rule could return $\{a, b\}$ as tied winners on both profiles.

Axiomatic Approach

- Stronger but more subjective axioms
- Majority consistency: If a majority of voters have the same top choice, that alternative wins.

$$\left(|\{i: top(\succ_i) = a\}| > \frac{n}{2}\right) \Rightarrow f(\overrightarrow{\succ}) = a$$

- Condorcet consistency: If a defeats every other alternative in a pairwise election, a wins.

$$\left(|\{i: a \succ_i b\}| > \frac{n}{2}, \forall b \neq a\right) \Rightarrow f(\overrightarrow{\succ}) = a$$

Axiomatic Approach

- Recall: Condorcet consistency ⇒ Majority consistency

- All positional scoring rules violate Condorcet consistency.

- Most positional scoring rules also violate majority consistency.
  - Plurality satisfies majority consistency.

Axiomatic Approach

- Consistency: If a is the winner on two profiles, it must be the winner on their union.

$$f(\overrightarrow{\succ}_1) = a \wedge f(\overrightarrow{\succ}_2) = a \Rightarrow f(\overrightarrow{\succ}_1 + \overrightarrow{\succ}_2) = a$$

  - Example: $\overrightarrow{\succ}_1 = \{a \succ b \succ c\}$, $\overrightarrow{\succ}_2 = \{a \succ c \succ b, b \succ c \succ a\}$
  - Then, $\overrightarrow{\succ}_1 + \overrightarrow{\succ}_2 = \{a \succ b \succ c, a \succ c \succ b, b \succ c \succ a\}$

- Theorem [Young '75]:
  - Subject to mild requirements, a voting rule is consistent if and only if it is a positional scoring rule!

Axiomatic Approach

- Weak monotonicity: If a is the winner, and a is "pushed up" in some votes, a remains the winner.
  - $f(\overrightarrow{\succ}) = a \Rightarrow f(\overrightarrow{\succ}') = a$, where
    - $b \succ_i c \Leftrightarrow b \succ_i' c, \forall i \in N,\ b, c \in A \backslash \{a\}$ (Order of others preserved)
    - $a \succ_i b \Rightarrow a \succ_i' b, \forall i \in N,\ b \in A \backslash \{a\}$ (a only improves)

- In contrast, strong monotonicity requires $f(\overrightarrow{\succ}') = a$ even if $\overrightarrow{\succ}'$ only satisfies the 2nd condition
  - Too strong; only satisfied by dictatorial or non-onto rules [GS Theorem]

Axiomatic Approach

- Weak monotonicity: If a is the winner, and a is "pushed up" in some votes, a remains the winner.
  - $f(\overrightarrow{\succ}) = a \Rightarrow f(\overrightarrow{\succ}') = a$, where
    - $b \succ_i c \Leftrightarrow b \succ_i' c, \forall i \in N,\ b, c \in A \backslash \{a\}$ (Order of others preserved)
    - $a \succ_i b \Rightarrow a \succ_i' b, \forall i \in N,\ b \in A \backslash \{a\}$ (a only improves)

- Weak monotonicity is satisfied by most voting rules
  - Popular exceptions: STV, plurality with runoff
  - But this helps STV be hard to manipulate
    - Theorem [Conitzer-Sandholm '06]: "Every weakly monotonic voting rule is easy to manipulate on average."

Axiomatic Approach

- STV violates weak monotonicity

7 voters	5 voters	2 voters	6 voters
a	b	b	c
b	c	c	a
c	a	a	b

7 voters	5 voters	2 voters	6 voters
a	b	a	c
b	c	b	a
c	a	c	b

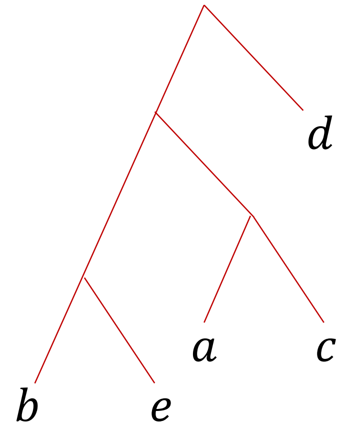
- First c, then b eliminated
- Winner: a

- First b, then a eliminated
- Winner: c

Axiomatic Approach

- Pareto optimality: If $a \succ_i b$ for all voters i, then $f(\vec{\succ}) \neq b$.

- Relatively weak requirement
  - Some rules that throw out alternatives early may violate this.
  - Example: voting trees
    - Alternatives move up by defeating opponent in pairwise election
    - d may win even if all voters prefer b to d if b loses to e early, and e loses to c



Axiomatic Approach

- Arrow's Impossibility Theorem
  - Applies to social welfare functions (profile → ranking)
  - Independence of Irrelevant Alternatives (IIA): If the preferences of all voters between a and b are unchanged, the social preference between a and b should not change
  - Pareto optimality: If all prefer a to b, then the social preference should be $a \succ b$
  - Theorem: IIA + Pareto optimality ⇒ dictatorship.

- Interestingly, automated theorem provers can also prove Arrow's and GS impossibilities!

Axiomatic Approach

- One can think of polynomial time computability as an axiom

  - Two rules that attempt to make the pairwise comparison graph acyclic are NP-hard to compute:
    - Kemeny’s rule: invert edges with minimum total weight
    - Slater’s rule: invert minimum number of edges

  - Both rules can be implemented by straightforward integer linear programs
    - For small instances (say, up to 20 alternatives), NP-hardness isn’t a practical concern.

Statistical Approach

- According to Condorcet [1785]:
  - The purpose of voting is not merely to balance subjective opinions; it is a collective quest for the truth.
  - Enlightened voters try to judge which alternative best serves society.

- Modern motivation due to human computation systems
  - EteRNA: Select 8 RNA designs to synthesize so that the truly most stable design is likely one of them

Statistical Approach

- Traditionally well-explored for choosing a ranking

- For $m = 2$, the majority choice is most likely the true choice under any reasonable model.

- For $m \geq 3$: Condorcet suggested an approach, but the writing was too ambiguous to derive a well-defined voting rule.

Statistical Approach

- Young's interpretation of Condorcet's approach:
  - Assume there is a ground truth ranking σ^*
  - Each voter i makes a noisy observation σ_i
  - The observations are i.i.d. given the ground truth
    - $\Pr[\sigma|\sigma^*] \propto \varphi^{d(\sigma,\sigma^*)}$
    - d = Kendall-tau distance = #pairwise disagreements
    - Interesting tidbit: Normalization constant is independent of σ^*

      $\Sigma_\sigma \varphi^{d(\sigma,\sigma^*)} = 1 \cdot (1+\varphi) \cdot \ldots \cdot (1+\varphi+\cdots+\varphi^{m-1})$

  - Which ranking is most likely to be the ground truth (maximum likelihood estimate – MLE)?
    - The ranking that Kemeny's rule returns!

Statistical Approach

- The approach yields a uniquely optimal voting rule, but relies on a very specific distribution
  - Other distributions will lead to different MLE rankings.
  - Reasonable if sufficient data is available to estimate the distribution well
  - Else, we may want robustness to a wide family of possible underlying distributions [Caragiannis et al. ’13, ’14]
- A connection to the axiomatic approach
  - A voting rule can be MLE for *some* distribution only if it satisfies consistency. (Why?)
    - Maximin violates consistency, and therefore can never be MLE!

Implicit Utilitarian Approach

- Utilities: Voters have underlying numerical utilities
  - Utility of voter i for alternative a = $u_i(a)$
    - Normalization: $\sum_a u_i(a) = 1$ for all voters i
  - Given utility vector $\vec{u}$, $sw(a, \vec{u}) = \sum_i u_i(a)$
  - Goal: choose $a^* \in \text{argmax}_a\, sw(a, \vec{u})$

- Preferences: Voters only report ranked preferences consistent with their utilities
  - $u_i(a) > u_i(b) \Rightarrow a \succ_i b$
  - Preference profile: $\vec{\succ}$
  - Cannot maximize welfare given only partial information

Implicit Utilitarian Approach

- Modified goal: Achieve the best worst-case approximation to social welfare

- Distortion of voting rule f

$$\max_{\vec{u}} \frac{\max_a \text{sw}(a, \vec{u})}{\text{sw}\left(f\left(\vec{\succ}\right), \vec{u}\right)}$$

  - Here, $\vec{\succ}$ are the preferences cast by voters when their utilities are $\vec{u}$
  - If f is randomized, we need $E\left[\text{sw}\left(f\left(\vec{\succ}\right), \vec{u}\right)\right]$

Utilitarian Approach

- Pros:
  - Uses minimal subjective assumptions
  - Yields a uniquely optimal voting rule
    - One can define the distortion of f on a given input $\vec{\succ}$ by taking the worst case over all $\vec{u}$ which would generate $\vec{\succ}$
    - Optimal voting rule minimizes the distortion on every $\vec{\succ}$ individually

- Cons:
  - The optimal rule does not have an intuitive formula that humans can comprehend
  - In some scenarios, the optimal rule is difficult to compute

Choosing One Alternative

- Theorem [Caragiannis et al. '16]:
  Given ranked preferences, the optimal deterministic voting rule has $\Theta(m^2)$ distortion.

- Proof:
  - Lower bound: Construct a profile on which every deterministic voting rule has $\Omega(m^2)$ distortion.
  - Upper bound: Show *some* deterministic voting rule that has $O(m^2)$ distortion on every profile.

Choosing One Alternative

$n/(m-1)$ voters per column			
a_1	a_2	…	a_{m-1}
a_m	a_m	…	a_m
⋮	⋮	⋮	⋮

- Proof (lower bound):
  - Consider the profile on the right
  - If the rule chooses a_m:
    - Infinite distortion. WHY?
  - If the rule chooses a_i for $i < m$:
    - Construct a bad utility profile $\vec{u}$ as follows
      - Voters in column i have utility $1/m$ for every alternative
      - All other voters have utility $1/2$ for their top two alternatives
    - $\text{sw}(a_i, \vec{u}) = \frac{n}{m-1} \cdot \frac{1}{m}$, $\text{sw}(a_m, \vec{u}) \geq \frac{n - n/(m-1)}{2}$
    - Distortion = $\Omega(m^2)$

Choosing One Alternative

- Proof (upper bound):
  - Simply using plurality achieves $O(m^2)$ distortion.
    - WHY?
  - Suppose plurality winner is a.
    - At least n/m voters prefer a the most, and thus have utility at least $1/m$ for a.
  - $sw(a, \vec{u}) \geq n/m^2$
  - $sw(a^*, \vec{u}) \leq n$ for every alternative a^*
  - $O(m^2)$ distortion

Implicit Utilitarian Voting

- Plurality is as good as any other deterministic voting rule!

- Alternatively:
  - If we must choose an alternative deterministically, ranked preferences provide no more useful information than top-place votes do, in the worst case.

- There's more hope if we're allowed to randomize.

Choosing One Alternative

- Theorem [Boutilier et al. '12]:
  Given ranked preferences, the optimal randomized voting rule has distortion $O(\sqrt{m} \cdot \log^* m)$, $\Omega(\sqrt{m})$.

- Proof:
  - Lower bound: Construct a profile on which every randomized voting rule $\Omega(\sqrt{m})$ distortion.
  - Upper bound: Show *some* randomized voting rule that has $O(\sqrt{m} \cdot \log^* m)$ distortion
    - We'll do the much simpler $O(\sqrt{m \log m})$ distortion

Choosing One Alternative

- Proof (lower bound):
  - Consider a similar profile:
    - $\sqrt{m}$ special alternatives
    - Voting rule must choose one of them (say a^*) w.p. at most $1/\sqrt{m}$

${}^{n}/_{\sqrt{m}}$ voters per column			
a_1	a_2	…	$a_{\sqrt{m}}$
⋮	⋮	⋮	⋮

  - Bad utility profile $\vec{u}$:
    - All voters ranking a^* first give utility 1 to a^*
    - All other voters give utility $1/m$ to each alternative
    - $\frac{n}{\sqrt{m}} \le \text{sw}(a^*, \vec{u}) \le \frac{2n}{\sqrt{m}}$
    - $sw(a, \vec{u}) \le n/m$ for every other a.
    - Distortion lower bound: $\sqrt{m}/3$ (proof on the board!)

Choosing One Alternative

- Proof (upper bound):
  - Given profile $\vec{\succ}$, define the harmonic score $\text{sc}(a, \vec{\succ})$:
    - Each voter gives $1/k$ points to her k^{th} most preferred alternative
    - Take the sum of points across voters
    - $\text{sw}(a, \vec{u}) \le \text{sc}(a, \vec{\succ})$ (WHY?)
    - $\sum_a sc(a, \vec{\succ}) = n \cdot \sum_{k=1}^{m} 1/k = n\, H_m \le n \cdot (\ln m + 1)$
  - Golden rule:
    - W.p. ½: Choose every a w.p. proportional to $sc(a, \vec{\succ})$
    - W.p. ½: Choose every a w.p. $1/m$ (uniformly at random)
  - Distortion $\le 2\sqrt{m \cdot (\ln m + 1)}$ (proof on the board!)

Optimal vs Near-Optimal Rules

- The distortion is often bad for large m
  - E.g., $\Theta(m^2)$ for deterministic rules.
  - But one can argue that the optimal alternative which minimizes distortion represents *some* meaningful aggregation of information.

- How difficult is it to find the *optimal* alternative?
  - Polynomial time computable for both deterministic (via a direct formula) and randomized (via a non-trivial LP) cases

Input Format

- What if we ask about underlying numerical utilities in a format other than ranking?

- Threshold approval votes
  - Voting rule selects a threshold τ, asks each voter i, for each alternative a, whether $u_i(a) \geq \tau$
  - $O(\log m)$ distortion!

- Food for thought
  - What is the tradeoff between the number of bits of information elicited and the distortion achieved?
  - What is the best input format for a given number of bits?

Implicit Utilitarian Approach

- Extensions
  - Selecting a subset of alternatives or a ranking
    - Lack of an obvious objective function
    - Has been studied for some natural objective functions [Caragiannis et al. ’16, ongoing work]
  - Participatory budgeting [Benade et al. ’17]
  - Graph problems
  - Project idea: Replace numbers with rankings in any problem!

- Deployed

ROBOVOTE