CSC2556

Lectures 4-5

Distortion

Distortion Approach

Distortion Approach

- A quantitative approach to voting

- Assumptions
  1. Voters' *ranked* preferences are induced by their *underlying numerical utilities / costs*
  2. The goal is to maximize the social welfare (sum of voter utilities) / minimize the social cost (sum of voter costs)
  3. Select an alternative that *approximately optimizes* the goal even in the worst case (the approximation ratio is called *distortion*)

- Increasingly popular in recent years
  - Yields an optimal voting rule with minimal assumptions, but the optimal rule can be hard to understand or compute

Utilitarian Framework

- Underlying utility profile $\vec{u} = (u_1, \ldots, u_n)$
  - $u_i(a)$ = utility of voter i for alternative a
  - Normalization: $\sum_a u_i(a) = 1$ for all voters i
  - Each voter i submits a consistent ranking $\succ_i$
    - $\forall a, b : a \succ_i b \Rightarrow u_i(a) \geq u_i(b)$

- Goal: social welfare $sw(a, \vec{u}) = \sum_i u_i(a)$
  - Ideally, we would like to choose $a^* \in \text{argmax}_a\, sw(a, \vec{u})$
  - But voting rule f only gets access to the ranked profile $\vec{\succ}$

Utilitarian Framework

- Distortion of f

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

  - Supremum is over consistent pairs of $\vec{u}$ and $\vec{\succ}$
  - If f is randomized, we use $E\left[\text{sw}\left(f\left(\vec{\succ}\right), \vec{u}\right)\right]$

- Example on the board!

Deterministic Rules

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

- Proof (lower bound):
  - High-level approach:
    - Take an arbitrary voting rule f
    - Construct a preference profile $\vec{\succ}$
    - Let f choose a winner a on $\vec{\succ}$
    - Reveal a bad utility profile $\vec{u}$ consistent with $\vec{\succ}$ in which a is $\Omega(m^2)$ factor worse than the optimal alternative

Deterministic Rules

- Proof (lower bound):
  - Let f be any voting rule
  - Consider $\overrightarrow{\succ}$ on the right
  - Case 1: $f(\overrightarrow{\succ}) = a_m$
    - Infinite distortion. WHY?
  - Case 2: $f(\overrightarrow{\succ}) = a_i$ for some $i < m$
    - Bad utility profile $\vec{u}$ consistent with $\overrightarrow{\succ}$
      - 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 = $\Omega(m^2)$

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

Deterministic Rules

- Proof (upper bound):
  - Actually, the simple plurality rule achieves $O(m^2)$ distortion
  - Suppose a is a plurality winner
    - At least n/m voters have a as their top choice
    - Every voter has utility at least $1/m$ for their top choice
  - $sw(a, \vec{u}) \geq n/m^2$
  - $sw(a^*, \vec{u}) \leq n$ for every alternative a^*
  - $O(m^2)$ distortion

Randomized Rules

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

- Proof (lower bound):
  - Same high-level approach:
    - Take an arbitrary *randomized* voting rule f
    - Construct a preference profile $\vec{\succ}$
    - Let f choose a distribution over alternatives p
    - Reveal a bad utility profile $\vec{u}$ consistent with $\vec{\succ}$ in which the expected social welfare under p is $\Omega(\sqrt{m})$ factor worse than the optimal social welfare

Randomized Rules

- Proof (lower bound):
  - Let f be an arbitrary rule
  - Consider $\vec{\succ}$ on the right with $\sqrt{m}$ special alternatives
  - f must choose at least one special alternative (say a^*) w.p. at most $1/\sqrt{m}$
  - Bad utility profile $\vec{u}$ consistent with $\vec{\succ}$:
    - All voters ranking a^* first give utility 1 to a^*
    - All other voters give utility $1/m$ to each alternative
    - $\frac{n}{\sqrt{m}} \leq \text{sw}(a^*, \vec{u}) \leq \frac{2n}{\sqrt{m}}$
    - $sw(a, \vec{u}) \leq n/m$ for every other a
    - Distortion lower bound: $\Omega(\sqrt{m})$ (proof on the board!)

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

Randomized Rules

- Proof (upper bound):
  - Given preference profile $\vec{\succ}$, define harmonic scores $\text{sc}(a, \vec{\succ})$:
    - Each voter gives $1/k$ points to her k^{th} most preferred alternative
    - Take the sum of points across voters
  - How does the harmonic score relate to social welfare?
    - It is an upper bound on social welfare
      - $\text{sw}(a, \vec{u}) \leq \text{sc}(a, \vec{\succ})$ (WHY?)
    - On average, it is a relatively tight upper bound
      - $\sum_a sc(a, \vec{\succ}) = n \cdot \sum_{k=1}^{m} 1/k = n\, H_m \leq n \cdot (\ln m + 1)$
      - $\sum_a sw(a, \vec{\succ}) = n$

Randomized Rules

- Proof (upper bound):
  - Golden rule f:
    - With probability ½:
      - Choose every a with probability proportional to $sc(a, \vec{\succ})$
    - With the remaining probability ½:
      - Choose every a with probability $1/m$ (uniformly at random)
  - $\text{dist}(f) \le 2\sqrt{m \cdot (\ln m + 1)}$ (proof on the board!)

Some Thoughts

- How do we interpret the distortion number?
  - Sometimes distortion can be high for all alternatives
  - The exact distortion number may be less useful than determining *which* alternative minimizes distortion
- Optimal vs asymptotically optimal
  - Plurality and "golden rule" are (almost) asymptotically optimal
  - But one can also write an optimization program that chooses the exact alternative minimizing distortion on each input $\vec{\succ}$
  - Polytime-time computable for both deterministic (via a direct formula) and randomized (via a non-trivial LP) cases

Some Thoughts

- Extensions
  - Selecting a subset of k alternatives or a ranking of alternatives
  - Participatory budgeting
  - Graph matching
  - Resource allocation
  - …

Deployed @ ROBOVOTE

Metric Framework

- Costs instead of utilities
- Underlying metric d over voters and alternatives
  - [Triangle Inequality] $\forall x, y, z: d(x,y) + d(y,z) \geq d(x,z)$
  - Each voter i submits a consistent ranking $\succ_i$
    - $\forall a, b : a \succ_i b \Rightarrow d(i,a) \leq d(i,b)$
- Goal: social cost $sc(a,d) = \sum_i d(i,a)$
  - ➢ Ideally, we would choose $a^* \in \text{argmin}_a\, sc(a,d)$
  - ➢ But voting rule f only gets access to the ranked profile $\vec{\succ}$

Metric Distortion

- Metric distortion of f

$$\text{dist}(f) = \sup_{d, \overrightarrow{\succ}} \frac{\text{sc}\left(f\left(\overrightarrow{\succ}\right), d\right)}{\min_a sc(a, d)}$$

  - Supremum is over consistent pairs of d and $\overrightarrow{\succ}$
  - If f is randomized, we use $E\left[sc\left(f\left(\overrightarrow{\succ}\right), d\right)\right]$

- Example on the board!

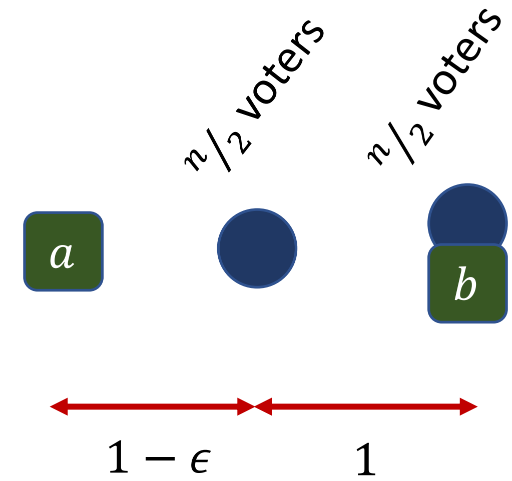
Deterministic Rules

- A simple lower bound of 3 with just two candidates

$^n/_2$ voters : $a \succ b$

$^n/_2$ voters : $b \succ a$

Rule selects wlog → **Winner** $\boldsymbol{a}$ → Bad instance



What about upper bounds?

Deterministic Rules

Distortion	Rule	Citation
Unbounded	k-approval ($k > 2$)	[Anshelevich et al., 2015]
$\Theta(m)$	Plurality, Borda count	[Anshelevich et al., 2015]
$\Theta(\sqrt{m})$	Ranked pairs, Schulze	[Kempe 2020]
$O(\log m)$, $\Omega(\sqrt{\log m})$	STV	[Skowron and Elkind, 2017]
5	Copeland's rule	[Anshelevich et al., 2015]
$2 + \sqrt{5} \approx 4.236$	A new rule	[Munagala and Wang, 2019]
3	PluralityMatching	[Gkatzelis et al., 2020]

Randomized Rules

Distortion	Rule	Citation
$3 - {}^{2}/_{n}$	Random Dictatorship	[Anshelevich and Postl, 2017]
$3 - {}^{2}/_{m}$	Smart Dictatorship	[Kempe 2020, Gkatzelis et al. 2020]
≥ 2	Lower bound	Same example as before
≥ 2.0261	Lower bound	[Charikar and Ramakrishnan, 2021]

- Major open question:
  - What is the optimal metric distortion for randomized voting rules?