Lecture 12: Social Choice, Condorcet, and Majority

Feb. 22, 2005

Lecturer: Ryan O'Donnell *Scribe: Aaron Roth*

1 Social Choice Theory

Social choice theory studies the aggregation of many individual preferences into one collective preference – how one may define "the will of the people". It is a topic without a home, studied in economics, political science, and mathematics. We will take a mathematical approach. Throughout this discussion we will use the convention that an election has n voters and $k \geq 2$ candidates.

Definition 1.1 *A* **Voter Preference Profile** *is a total ordering on the candidates.*

Definition 1.2 $\mathbf{P_k}$ *is the set of all $k!$ voter preference profiles.*

Definition 1.3 *A* **Social Choice Function** *is a function $f : \mathbf{P_k}^n \to [k]$, mapping n voter preferences into a winner from the set of candidates.*

Remark 1.4 *We will most often study the case of elections between two candidates ($k = 2$). In this case we label the candidates $-1, 1$ and identify $\mathbf{P_2}$ with $\{-1, 1\}$ (A preference profile is identified with the favored candidate). Social choice functions are then boolean functions $f : \{-1, 1\}^n \to \{-1, 1\}$.*

Examples of social choice functions include:

1. **Plurality:** The winner is the candidate that was ranked first most often.
2. **Borda Count:** Each voter gives the i'th ranked candidate $(k + 1 - i)$ points. The winner is the candidate with the most points.

Remark 1.5 *In practice, social choice functions would have to specify behavior in case of ties, but we will ignore this issue which tends to be unenlightening.*

1.1 Properties of Social Choice Functions

There are certain criteria that we would like a 'good' social choice function f to meet:

1. **Relevant:** Every coordinate of the function f should be relevant – There should be no voters whose votes are simply discarded.

2. **Positively Responsive:** Suppose that $f(q_1, \ldots, q_n) = j$. Then for all i, if the preference profile q_i is modified by swapping candidate j with a candidate of higher rank, to give profile q_i', then $f(q_1, \ldots, q_i', \ldots, q_n) = j$. A voter increasing his ranking of candidate j should not prevent j from winning the election.

3. **Faithful** (Also called **Paretian**): If all voter profiles rank j as the #1 candidate, then $f(q_1, \ldots, q_n) = j$. (For $k = 2$: $f(-1, \ldots, -1) = -1$ and $f(1, \ldots, 1) = 1$.)

4. **Neutral:** If σ is a permutation on $[k]$ then:

$$f(\sigma q_1, \ldots, \sigma q_n) = \sigma(f(q_i, \ldots, q_n))$$

(For $k = 2$: f(-x) = -f(x), i.e. f is odd.)

5. **Anonymous:** For all permutations π on $[n]$:

$$f(q_1, \ldots, q_n) = f(q_{\pi(1)}, \ldots, q_{\pi(n)})$$

Equivalently, f depends only on the $k!$ quantities $|\{q_i : q_i = q\}|$ for each $q \in \mathbf{P_k}$. (For $k = 2$: $f(x)$ is a **Totally Symmetric** function, depending only on $\sum_{i=1}^n x_i$.)

Remark 1.6 *For $k = 2$, anonymous monotone functions are* **Unweighted Threshold Functions** *$f(x) = \operatorname{sgn}(\sum_{i=1}^n x_i - \theta)$ for $\theta \notin \mathbb{Z}$ (to avoid ties). Note that if f is also neutral, then for odd n, f must be the majority function, and for even n, there is no such f. Already we see an impossibility result for social choice functions with properties that we would like to have.*

If strong conditions prove to be impossible to reconcile, we can define weaker conditions. For example, the "majority of majorities" function (if we let all second level majority functions be on an equal number of variables) is **Weakly Symmetric** or **Transitive** since it satisfies: $\forall i \neq i'$ there exists a permutation π on $[n]$ such that:

$$f(q_{\pi(1)}, \ldots, q_{\pi(n)}) = f(q_1, \ldots, q_n) \text{ and } \pi(i) = i$$

The majority of majorities function roughly models the electoral college system used for presidential elections in the United States, with each inner majority function representing a state. Notice that if i and i' in the above definition are within the same 'state', then we may use any permutation on the 'residents' of that state that map i to i'. If i and i' are in different states, we may use permutations on their states if we additionally swap their states, again mapping i to i'.

Another weaker condition (a relaxed version of neutrality) is to insist that f be **Unbiased**: If voter profiles are chosen i.i.d. from the uniform distribution, then each candidate has a $1/k$ probability of winning.

Definition 1.7 *A* **Social Welfare Function** *is a function $f : \mathbf{P_k}^n \to \mathbf{P_k}$, that is, a method of vote aggregation that maps voter preferences to an ordering on the candidates, rather than only a single winner.*

Remark 1.8 *Condorcet and Borda (A contemporary of Condorcet) suggested building social welfare functions from social choice functions by doing all $\binom{k}{2}$ pairwise comparisons. Unfortunately, this method may lead to cycles (recall the Condorcet paradox).*

Condorcet was not completely discouraged by his 'paradox', however, and was concerned with how likely it was that cycles would be generated in practice. He wrote in his 1788 *On the form of decisions made by plurality vote*: "But after considering the facts, the average values or the results, we still need to determine their probability." In order to address such issues, we need randomized models of voting, which were popularized in the 1970s.

Definition 1.9 *The* **Impartial Culture (IC) Assumption:** *The voters vote i.i.d. from the uniform distribution over* $\mathbf{P_k}$*. (For* $k = 2$ *this is the uniform distribution over* $\{-1, 1\}^n$*).*

Remark 1.10 *This is of course an unrealistic assumption – votes are not at all independent. But perhaps it is a reasonable model. We may eliminate staunch Republicans and Democrats by encoding them as fixed properties of the social choice function, and then model the undecided voters as random.*

Definition 1.11 *For* $k = 3$ *and* $f : \{-1, 1\}^n \to \{-1, 1\}$, $\textbf{Rationality}(f)$ *is the probability that under the IC assumption,* f *produces no cycles.*

Recall:

$$\text{Rationality}(f) = \frac{3}{4} - \frac{3}{4} \sum_S (-\frac{1}{3})^{|S|} \hat{f}(S)^2 = \frac{3}{4} - \frac{3}{4} \mathbb{S}_{-\frac{3}{4}}(f)$$

Remark 1.12 *Majority is the best boolean function ever!*

2 An Ode to Majority

Imagine that we model voting as a game, in which each individual voter tries to guess the outcome of the election. Each voter derives positive utility if he guesses correctly, and negative utility otherwise. Then majority is the function that maximizes the social good by maximizing the number of winners of this game. If most people guessed correctly, perhaps we can interpret their guesses as "The Will of the People"

Proposition 2.1 *Let* n *be odd and* $k = 2$*. Assume IC. Then majority is the unique social choice function maximizing* $\mathbf{E}_x[|\{x_i = f(x)\}|]$

Proposition 2.2

$$\mathbf{Pr}[x_i = f(x)] = \frac{1}{2} + \frac{1}{2} \hat{f}(\{x_i\})$$

Proof:

$$\mathbf{Pr}[x_i = f(x)] = \mathbf{E}[\frac{1}{2} + \frac{1}{2}x_i f(x)] = \frac{1}{2} + \frac{1}{2}\mathbf{E}[x_i f(x)] = \frac{1}{2} + \frac{1}{2}\hat{f}(\{x_i\})$$

$\square$

Proposition 2.3 *Equivalently to Proposition 2.2,* $\sum_{i=1}^{n} \hat{f}(\{i\})$ *is maximized by majority.*

Proof:

$$\begin{aligned}\sum_{i=1}^{n} \hat{f}(\{i\}) &= \sum_{i=1}^{n} \mathbf{E}_x[f(x)x_i] \\ &= \mathbf{E}_x[f(x)\sum_{i=1}^{n} x_i] \\ &\leq \mathbf{E}[|\sum_{i=1}^{n} x_i|]\end{aligned}$$

with equality if and only if $f(x) = \mathrm{sgn}(\sum_{i=1}^{n} x_i)$, which is majority. $\square$

Proposition 2.4 *If* $f : \{-1,1\}^n \to \{-1,1\}$ *is monotone, then* $\hat{f}(\{i\}) = \mathrm{Inf}_i(f)$.

We present two proofs:
Proof: Consider an experiment in which we write down all 2^n strings $x \in \{-1,1\}$ and then pair all strings that are identical except on their i'th coordinate (Getting 2^{n-1} pairs). Then $\mathrm{Inf}_i(f)$ represents a count of the number of pairs such that $f(x) = -f(x^{(i)})$. Similarly, $\hat{f}(\{i\})$ represents a count of the number of pairs such that $f(x^{(i=-1)}) = -1$ and $f(x^{(i=1)}) = 1$ minus the number of pairs such that $f(x^{(i=-1)}) = 1$ and $f(x^{(i=1)}) = -1$ (Note that in any other case, the pair contributes 0 to the expectation $\hat{f}(\{i\}) = \mathbf{E}_x[x_i f(x)]$). However, the second case can never occur if f is monotone, since by definition flipping a single bit from -1 to 1 can only increase the value of $f(x)$. Therefore, both $\hat{f}(\{i\})$ and $\mathrm{Inf}_i(f)$ are counting the same quantity. $\square$

Before the second proof we need a definition:

Definition 2.5 D_i *is the operator on functions* $f : \{-1,1\}^n \to \mathbb{R}$ *such that* $D_i f : \{-1,1\} \to \mathbb{R}$ *is defined as:*

$$D_i f(x) = \frac{f(x^{(i=1)}) - f(x^{(i=-1)})}{2}$$

Proposition 2.6

$$D_i f(x) = \sum_{S \ni i} \hat{f}(S) x_{S \setminus \{i\}}$$

so D_i *acts as the differentiation operator with respect to* i.

Observation 2.7 *If $f : \{-1,1\}^n \to \{-1,1\}$ is monotone then $D_i f(x) \geq 0$.*

We are now ready for the second proof:

Proof:

$$\begin{aligned} \mathrm{Inf}_i &= \mathbf{E}_x[D_i f(x)^2] \\ &= \mathbf{E}_x[D_i f(x)] \\ &= \hat{D_i} f(\emptyset) \\ &= \hat{f}(\{i\}) \end{aligned}$$

where the second equality follows from the fact that f is boolean valued and monotone. □

Proposition 2.8

$$\mathrm{Inf}_i(Maj) = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{n}} \pm O(\frac{1}{n^{3/2}})$$

As a corollary:

$$\sum_{i=1}^{n} \hat{Maj}(\{i\}) = \sqrt{\frac{2}{\pi}}\sqrt{n} \pm O(\frac{1}{\sqrt{n}})$$

Proof: Since bit i is influential in majority only if it casts a 'deciding vote' – that is, only if there would be a tie if bit i were removed, we have:

$$\mathrm{Inf}_i(Maj) = 2^{-(n-1)}\binom{n-1}{\frac{n-1}{2}}$$

Writing $m = n-1$ and applying Stirling's Approximation ($k! = \sqrt{2\pi k}(\frac{k}{e})^k(1 \pm O(1/k))$) gives:

$$\begin{aligned} \mathrm{Inf}_i(Maj) &= 2^{-m}\frac{\sqrt{2\pi m}(m/e)^m}{2\pi(m/2)(m/2e)^m}(1 \pm O(\frac{1}{m}) \\ &= \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{m}} \pm O(\frac{1}{m^{3/2}}) \end{aligned}$$

□

Corollary 2.9 *For monotone functions f:*

$$\mathbb{I}(f) = \sum_{i=1}^{n} \mathrm{Inf}_i(f)$$

is maximized by Maj_n. Therefore for all monotone f:

$$\mathbb{I}(f) \leq O(\sqrt{n})$$

Corollary 2.10 *If f is monotone, $\forall \epsilon > 0$ f is ϵ-concentrated on $\{S : |S| \leq O(\sqrt{n}/2)\}$.*

Corollary 2.11 *[Bshouty-Tamon '96] $\{f : \{-1,1\}^n \to \{-1,1\} \text{ monotone}\}$ is learnable in time $n^{O(\sqrt{n}/\epsilon)} << 2^n$.*

Proposition 2.12

$$W_1(Maj) = \sum_{|S|=1} \hat{Maj}(S)^2 \; \frac{2}{\pi}$$

Recall:

$$\begin{aligned} \text{Rationality}(f) &= \frac{3}{4} - \frac{3}{4}\mathbb{S}_{-\frac{1}{3}}(f) \\ &= \frac{3}{4} - \frac{3}{4}W_0(f) - \frac{3}{4}(-\frac{1}{3})W_1(f) - \frac{3}{4}(-\frac{1}{3})^2 W_2(f) - \ldots \end{aligned}$$

Proposition 2.13 *Let f be any social choice function. Then:*

$$\text{Rationality}(f) \leq \frac{7}{9} + \frac{2}{9}W_1(f)$$

and if f is neutral (odd) then additionally:

$$\text{Rationality}(f) \geq \frac{3}{4} + \frac{1}{4}W_1$$

Proof:

$$\text{Rationality}(f) \leq \frac{3}{4} + \frac{1}{4}W_1 + \frac{1}{36}(1 - W_1) = \frac{7}{9} + \frac{2}{9}W_1(f)$$

If f is odd, then $W_0 = W_2 = W_4 = \ldots = W_{2i} = \ldots = 0$. In this case:

$$\text{Rationality}(f) = \frac{3}{4} + \frac{1}{4}W_1 + \frac{1}{36}W_3 + \ldots \geq \frac{3}{4} + \frac{1}{4}W_1$$

□

Corollary 2.14

$$\text{Rationality}(Maj_n) \leq \frac{7}{9} + \frac{2}{9} \cdot \frac{2}{\pi} \approx .919$$

$$\text{Rationality}(Maj_n) \geq \frac{3}{4} + \frac{1}{4} \cdot \frac{2}{\pi} \approx .909$$

Note that all weakly symmetric functions have the same upper bound.

Proposition 2.15 *[Guilbaud '52]*

$$\text{Rationality}(Maj_n) \; \frac{3}{4} - \frac{3}{4} \cdot \frac{2}{\pi} \arcsin(-\frac{1}{3})$$