Presentation for 9 February 2021

Social Choice Theory

Berwin Gan

Framework of Problem

- Let there be a set $[n] := \{1, .., n\}$ of $n \in \mathcal{N}$ voters and a set $A := \{a_1, ..., a_m\}$ of $m \in \mathcal{N}$ candidates.
- Each voter $i \in [n]$ casts a linear ordering over A with $\succ_i$ such that $a_{b1} \succ_i a_{b2} \succ_i ... \succ_i a_{bm}$ where $a_{b1} \succ_i a_{b2} ... \succ_i a_{bm}$ is a permutation of $\{a_1, ..., a_m\}$. In each linear ordering, a_{bk} is said to have rank k.
- A combination of linear ordering is called a profile, P. An preference aggregation rule is a function that assigns to each profile a social preference relation on A.

Score Vector

- A scoring vector for m candidates is given by $(s_1, s_2, ..., s_m)$ where for each vote, the i^{th} ranked candidate is given the i^{th} score in the vector.
- Eg. The first ranked candidate in a vote is given s_1 score and the last ranked candidate is given s_m score.

Score Vector Rule

- The total score of an alternative can be obtained by summing the score the alternative obtained from every vote.
- The social preference relation is then determined based on the score of the alternatives with the alternative with the highest score being the most preferred and the alternative with the lowest score being the least preferred.

Equivalence of Two Score Vectors

- Score vectors with m entries where any two successive scores differ by the same amount is defined by $(A + (m-1)b, A + (m-2)b, ...A + 2b, A + b, A)$ where $b > 0$.
- Given a linear ordering, the alternatives with rank k will be given a score of $A + (m-k)b$.
- The social preference relation will be determined by the total scores of each candidate in descending order.

Claim: Any two score vectors with m entries where any two successive scores differ by the same amount generates equivalent social preference relation for any amount of voters.

Equivalence of Two Score Vector: Proof

Proof.

Let the two score vectors be defined by $(A + (m - 1)b, ..., A + b, A)$ and $(C + (m - 1)d, ..., C + d, C)$ where $b > 0$ and $d > 0$. As the term A is in every element of the first vector, the final score for all alternatives given n voters will be $nA + xb$ with x varying. Hence, we can simplified the vector to $((m - 1)b, ..., b, 0)$ and obtain the same social preference relation as every alternative score will be reduced by the same amount,nA. Similarly, the second vector can be simplified to $((m - 1)d, ..., d, 0)$

□

Equivalence of Two Score Vector: Proof

Proof.

Given that both scoring vectors differ only by their multiples of b and d. They will generate the same social preference relation as $b > 0$ and $d > 0$. Furthermore, the scores obtained from the first vector is $\frac{d}{b}$ the scores obtained from the second vector. □

Symmetric Borda Rule

The symmetric Borda score of an alternative a_p is calculated by:

$$Borda_P^{sym}(a_p) := \sum_{a_q \in A/\{a_p\}} Net_P(a_p > a_q)$$

The net preference for a_p over a_q is defined as:

$$Net_P(a_p, a_q) := |\{i \in [n] | a_p \succ_i a_q\}| - |\{i \in [n] | a_p \succ_i a_q\}|$$

Proposition 1

Claim: The social preference relation obtain from the symmetric Borda rule is equivalent to the social preference relation obtain from the scoring vector of $(m-1, m-3, ..., -(m-3), -(m-1))$

Poposition 1 (1)

Using the definition of net preference we can show that:

$$Borda_P^{sym}(a_p) = \sum_{a_q \in A/\{a_p\}} Net_P(a_p > a_q) \tag{1}$$

$$= \sum_{q=1}^{m} \left(|\{i \in [n] | a_p \succ_i a_q\}| - |\{i \in [n] | a_p \succ_i a_q\}| \right) \tag{2}$$

By expanding the summation over both sides of the equation we get:

$$Borda_P^{sym}(a_p) = \sum_{q=1}^{m} \left(|\{i \in [n] | a_p \succ_i a_q\}| - |\{i \in [n] | a_p \succ_i a_q\}| \right) \quad (3)$$

$$= \sum_{q=1}^{m} (|\{i \in [n] | a_p \succ_i a_q\}|) - \sum_{q=1}^{m} (|\{i \in [n] | a_p \succ_i a_q\}|) \quad (4)$$

Proposition 1 (3)

$$Borda_P^{sym}(a_p) = \sum_{q=1}^{m}(|\{i \in [n] | a_p \succ_i a_q\}|) - \sum_{q=1}^{m}(|\{i \in [n] | a_p \succ_i a_q\}|) \tag{5}$$

$$= \sum_{i=1}^{n}(|\{q \in A | a_p \succ_i a_q\}|) - \sum_{i=1}^{n}(|\{q \in A | a_p \succ_i a_q\}|) \tag{6}$$

To get from (5) to (6) we must represent the the left summation of both equations of the preference of alternative a_p towards other alternatives a_q in a matrix.

Proposition 1 (4)

$$S = \begin{bmatrix} s_{1,1} & s_{1,2} & ... & s_{1,m} \\ s_{2,1} & s_{2,2} & ... & s_{2,m} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ s_{n,1} & s_{n,2} & ... & s_{n,m} \end{bmatrix}$$

- We can represent the left summation with a matrix S with elements $s_{i,j}$ where $i = 1, ..., n$ representing the voters and $j = 1, ..., m$ representing the alternatives $a_1, a_2, ..., a_m$.
- Each element is 1 if the corresponding voter i votes $a_p > a_q$ for the corresponding candidate a_q where $q = j$ and $q \neq p$ and 0 otherwise.

Proposition 1(5)

- The left summation of equation (5) sums the ones over the rows while going over alternatives from $j = 1$ to $j = m$.
- On the other hand, the left summation of equation (6) sums the ones over the columns while going over each voter from $i = 1$ to $i = n$.
- As the sum value will be similar for both cases the two equations are equivalent.

Proposition 1 (6)

- Similarly, the right summation can be thought of an equivalent matrix with 1 in the matrix element corresponding to $a_q > a_p$

Proposition 1 (7)

When a candidate is ranked in the k^{th} position where $k = 1, ..., m$, where are $k-1$ candidate above and $m-k$ candidate below, hence:

$$Borda_P^{sym}(a_p) = \sum_{i=1}^{n}(|\{q \in A | a_p \succ_i a_q\}|) - \sum_{i=1}^{n}(|\{q \in A | a_p \succ_i a_q\}|) \quad (7)$$

$$= \sum_{i=1}^{n}(m - k_i) - \sum_{i=1}^{n}(k_i - 1) \quad (8)$$

Proposition 1 (8)

We can further simplify the equation by combining to summation to:

$$Borda_P^{sym}(a_p) = \sum_{i=1}^{n}(m - k_i) - \sum_{i=1}^{n}(k_i - 1) \tag{9}$$

$$= \sum_{i=1}^{n}(m - (2k_i - 1)) \tag{10}$$

Proposition 1 (9)

According go the symmetric Borda rule defintion of:

$$Borda_P^{sym}(a_p) = \sum_{i=1}^{n}(m - (2k_i - 1))$$

The alternative in rank k will have a score of $m - 2k + 1$ which is equal to the scoring vector of

$$(m - 1, m - 3, ..., -(m - 3), -(m - 1))$$

when $k = 1, ..., m$.