Ranking-Based Voting Revisited: Maximum Entropy Approach Leads to Borda Count (and Its Versions)

Olga Kosheleva[1], Vladik Kreinovich[1],
and Guo Wei[2]
[1]University of Texas at El Paso
El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu
[2]Department of Mathematics and Computer Science
University of North Carolina at Pembroke
Pembroke, North Carolina 28372 USA, guo.wei@uncp.edu

1. Need for Voting and Group Decision Making

- In many real-life situations, we need to make a decision that affects many people.
- Ideally, when making this decision, we should take into account the preferences of all the affected people.
- This group decision making situation is also known as *voting*.

2. What Information Can Be Used for Voting

- The simplest – and most widely used – type of voting is when each person selects one of the alternatives.
- After this selection, all we know is how many people voted for each alternative; clearly:
  - the more people vote for a certain alternative,
  - the better is this alternative for the community as a whole.
- Thus, if this is all the information we have, then:
  - a natural idea is
  - to select the alternative that gathered the largest number of votes.
- (Another idea is to keep only the alternatives with the largest number of votes and vote again.)

3. What Information Can Be Used (cont-d)

- In this scheme, for each person,
  - we only take into account one piece of information:
  - which alternative is preferable to this person.
- To make more adequate decision, it is desirable to use more information about people's preferences.
- An ideal case is when we use full information about people's preferences.
- This is ideal but this requires too much elicitation and is, thus, not used in practice.
- An intermediate stage – when we use more information than in the simple majority voting – is when:
  - we ask the participants to rank all the alternatives, and
  - we use these rankings to make a decision.

4. Ranking-Based Voting: A Brief Reminder

- The famous result by a Nobelist Kenneth Arrow shows:
  - that it is not possible to have a ranking-based voting scheme
  - which would satisfy all reasonable fairness-related properties.
- So what can we do? One of the ideas is *Borda count*, when:
  - for each participant i and for each alternative A_j,
  - we count the number b_{ij} of alternatives that the i-th participant ranked lower than A_j.
- Then, for each alternative A_j, we add up the numbers corresponding to different participants.
- We select the alternative with the largest sum $\sum_{i=1}^{n} b_{ij}$.

5. Why Borda Count?

- Borda count is often successfully used in practice.
- However, there are several other alternative schemes.
- This prompts a natural question: why namely Borda count and why not one of these other schemes?
- In this talk, we provide an explanation for the success of Borda count; namely, we show that:
  - if we use the maximum entropy approach – a known way for making decisions under uncertainty,
  - then the Borda count (and its versions) naturally follows.

6. How to Describe Individual Preferences

- We want to describe what we should do when only know the rankings.
- Let us first recall what decision we should make when we have full information about the preferences.
- To describe this, we need to recall how to describe these preferences.
- In decision theory, a user's preferences are described by using the notion of *utility*.
- To define this notion, we need to select two extreme alternatives:
  - a very bad alternative A_- which is worse than anything that we will actually encounter, and
  - a very good alternative A_+ which is better than anything that we will actually encounter.

7. How to Describe Preferences (cont-d)

- For each number p from the interval $[0, 1]$, we can then form a lottery $L(p)$ in which:
  - we get A_+ with probability p and
  - we get A_- with the remaining probability $1 - p$.
- For $p = 0$, the lottery $L(0)$ coincides with the very bad alternative A_-.
- Thus, $L(0)$ is worse than any of the alternatives A that we encounter: $L(0) = A_- < A$.
- For $p = 1$, the lottery $L(1)$ coincides with the very good alternative A_+.
- Thus, $L(1)$ is better than any of the alternatives A that we encounter: $A < L(1) = A_+$.
- Clearly, the larger p, the better the lottery.

8. How to Describe Preferences (cont-d)

- Thus, there exists a threshold p_0 such that:
  - for $p < p_0$, we have $A(p) < A$, and
  - for $p > p_0$, we have $A < A(p)$.
- This threshold is known as the *utility* of the alternative A; it is usually denoted by $u(A)$.
- In particular, according to this definition:
  - the very bad alternative A_- has utility 0, while
  - the very good alternative A_+ has utility 1.
- To fully describe people's preferences, we need to elicit,
  - from each person i,
  - this person's utility $u_i(A_j)$ of all possible alternatives A_j.

9. Utility Is Defined Modulo Linear Transformations

- The numerical value of utility depends on the selection of values A_- and A_+.
- One can show that, if we use a different pair of alternatives (A'_-, A'_+), then:
  - the resulting new utility values $u'(A)$ are related to the original values $u(A)$
  - by a linear dependence: $u'(A) = k + \ell \cdot u(A)$ for some k and $\ell > 0$.

10. Utility-Based Decision Making under Probabilistic Uncertainty

- In many practical situations, we do not know the exact consequences of different actions.
- For each action, we may have different consequences $c_1, \ldots, c_m$, with different utilities $u(c_1), \ldots, u(c_m)$.
- We can also usually estimate the probabilities $p_1, \ldots, p_m$ of different consequences.
- What is the utility of this action?
- This action is equivalent to selecting c_i with probability p_i.
- By definition of utility, each consequence c_i is, its turn, equivalent to a lottery in which:
  - we get A_+ with probability $u(c_i)$ and
  - we get A_- with the remaining probability $1 - u(c_i)$.

11. Utility-Based Decision Making (cont-d)

- Thus, the original action is equivalent to a 2-stage lottery as a result of which we get either A_+ or A_-.
- One can easily conclude that the probability of getting A_+ in this 2-stage lottery is equal to the sum

$$p_1 \cdot u(c_1) + \ldots + p_m \cdot u(c_m).$$

- Thus, by definition of utility, this sum is the utility of the corresponding action.
- It should be mentioned that this sum happens to be the expected value of utility.

12. How to Make a Group Decision: Simplest Choice Situation

- Suppose that we know the utility $u_i(A_j)$ of each alternative A_j for each participant i.
- Now, we need to decide which alternative to select.
- Each alternative is thus characterized by the tuple of the corresponding utility values $(u_1(A_j), \ldots, u_n(A_j))$.
- Based on the tuples corresponding to different alternatives, we need to select the best one.
- In other words, we need to be able:
  - given $(u_1(A_j), \ldots, u_n(A_j))$ and $(u_1(A_k), \ldots, u_n(A_k))$,
  - to decide which of the two alternatives is better, i.e., whether

$$(u_1(A_j), \ldots, u_n(A_j)) < (u_1(A_k), \ldots, u_n(A_k)) \text{ or}$$

$$(u_1(A_k), \ldots, u_n(A_k)) < (u_1(A_j), \ldots, u_n(A_j)).$$

13. How to Make a Group Decision (cont-d)

- In the voting situation, there is usually a status quo state:
  - the state that exists right now and
  - that will remain if we do not make any decision.
- For example:
  - if we are voting on different plans to decrease the traffic congestion in a city,
  - the status quo situation is not to do anything and to continue suffering traffic delays.
- The status quo situation is worse than any of the alternatives.
- So, we can take this status quo situation as the value A_-.
- In this case, for all participants, the utility of the status quo situation is 0.

14. How to Make a Group Decision (cont-d)

- The only remaining freedom is selecting A_+.
- If we replace the original very good alternative A_+ with a new alternative A'_+, then:
  - the corresponding linear transformation
  - should transform 0 into 0.
- Thus, it should have the form $u'_i(A) = \ell_i \cdot u_i(A)$.
- In principle, each participant can select his/her own scale.
- It is reasonable to require that:
  - if one of the participants selects a different A_+,
  - then the resulting group choice should not change.

15. How to Make a Group Decision (cont-d)

- For the order on the set of all the tuples:
  - if $(u_1, \ldots, u_n) < (u'_1, \ldots, u'_n)$
  - then $(\ell_1 \cdot u_1, \ldots, \ell_n \cdot u_n) < (\ell_1 \cdot u'_1, \ldots, \ell_n \cdot u'_n)$.
- Other requirements include:
  - monotonicity: if an alternative is better for everyone it should be preferred, and
  - fairness: the order should not change is we simply rename the participants.
- It turns out that the only order with this property is:

$$(u_1, \ldots, u_n) < (u'_1, \ldots, u'_n) \Leftrightarrow \prod_{i=1}^{n} u_i < \prod_{i=1}^{n} u'_i.$$

- This comparison is known as *Nash's bargaining solution* after the Nobelist John Nash.

16. How to Make a Group Decision: Case of Transferable Utility

- The above analysis refers to the case when we make a simple decision: e.g., when we simply elect an official.
- In many other group decision situations, however, the situation is more complicated.
- For example, some people may be opposed a road construction plan, since:
  - during this construction,
  - their access to their homes and businesses will be limited.
- In such situations, if this particular alternative seems to be overall the best, a reasonable idea is:
  - to use some of its benefits
  - to compensate those who will experience temporary inconveniences.

17. Case of Transferable Utility (cont-d)

- The possibility of such a compensation is known as *transferable utility:*
  - in contrast to the above simple choice situation,
  - we can transfer utility from one participant to another.
- That we can move utility from person to person means that we have a common unit for utility.
- So, when some utility is transferred, the sum of all utilities remains constant.
- Suppose that without the transfers, the utilities corresponding to some alternative are $u_1, \ldots, u_n$.

18. Case of Transferable Utility (cont-d)

- The possibility of transfer means that:
  - we can have different values $u'_1, \ldots, u'_n$,
  - as long as the sum of all the utilities remains the same: $\sum\limits_{i=1}^{n} u_i = \sum\limits_{i=1}^{n} u'_i$.
- The optimal transfer is when the product of the individual utilities attains the largest possible value.
- To find the resulting utility values, we need:
  - given the values $u_1, \ldots, u_n$,
  - to find the values $u'_1, \ldots, u'_n$ for which $\prod\limits_{i=1}^{n} u'_i$ is the largest under the constraint

$$\sum_{i=1}^{n} u_i = \sum_{i=1}^{n} u'_i.$$

19. Case of Transferable Utility (cont-d)

- By applying the Lagrange multiplier method, we can:
  - reduce this constraint optimization problem
  - to the unconstraint problem of optimizing the following objective function:

$$\prod_{i=1}^{n} u'_i + \lambda \cdot \left(\sum_{i=1}^{n} u_i - \sum_{i=1}^{n} u'_i \right).$$

- Differentiating this expression with respect to each unknown u'_i and equating the derivative to 0, we get

$$\prod_{i' \neq i} u'_{i'} - \lambda = 0, \text{ i.e., } \prod_{i' \neq i} u'_{i'} = \lambda.$$

- Thus, for each i, $u'_i = \dfrac{\prod_{i'=1}^{n} u'_{i'}}{\prod_{i' \neq i} u'_{i'}} = \dfrac{\prod_{i'=1}^{n} u'_{i'}}{\lambda}.$

20. Case of Transferable Utility (cont-d)

- The right-hand side of this formula does not depend on i, thus we have $u'_1 = \ldots = u'_n$.
- From the condition that $\sum\limits_{i=1}^n u_i = \sum\limits_{i=1}^n u'_i$, we conclude that $u'_1 = \ldots = u'_n = \dfrac{1}{n} \cdot \sum\limits_{i=1}^n u_i$ and thus, that

$$\prod_{i=1}^n u'_i = \left(\frac{1}{n} \cdot \sum_{i=1}^n u_i\right)^n.$$

- Among several alternatives, we should select the one for which this product is the largest.
- This is equivalent to selecting the alternative for which the sum $\sum\limits_{i=1}^n u_i$ attains its largest possible value.

21. Ranking-Based Voting: Reminder

- In the situation of ranking-based voting, we do not know the utilities.
- All we know, for each participant, is the ranking given by this participants to possible alternatives.
- Ranking $A_{i_1} < A_{i_2} < \ldots$ means that:
  - we can have different utility values $u(A_i) \in [0, 1]$
  - as long as these utility values are consistent with this ranking:

$$u(A_{i_1}) < u(A_{i_2}) < \ldots$$

22. Ranking-Based Voting (cont-d)

- In line with the above description of decision making under uncertainty:
  - to find an actual utility of each alternative for this participant,
  - we must find the expected value of the corresponding utility $u(A_j)$.
- To find this expected value, we need to select some probability distribution on the set of all possible tuples.

23. Maximum Entropy Approach: Idea

- There may be many different probability distributions on the set of all the property ordered tuples.
- We need to select one of them.
- Some of these distributions may have more uncertainty, some less.
- A reasonable idea is to keep the original uncertainty and not to add artificial certainty.
- So, we select, among all possible distributions, a distribution with the largest possible value of uncertainty.
- A natural measure of this uncertainty is the entropy.
- So, we select the distribution with the largest possible value of the entropy.

24. What Happens When We Apply the Maximum Entropy Approach

- The largest possible entropy is attained for a uniform distribution on the set of all the tuples.
- *Known:* for k alternatives $A_{i_1} < A_{i_2} < \ldots < A_{i_k}$, the resulting expected utility values $\overline{u}(A_j)$ are

$$\overline{u}_i(A_{i_q}) = \frac{q}{k+1}.$$

- For each alternative $A_j = A_{i_q}$:
  - its Borda count b_{ij} for this participant i is its number of worse-then-A_j alternatives,
  - so, it is equal to $b_{ij} = q - 1$.
- Thus, $q = b_{ij} + 1$, and in terms of this Borda count, the expected utility of each alternative A_j is equal to

$$\overline{u}_i(A_j) = \frac{b_{ij}+1}{k+1}.$$

25. This Explains the Borda Count

- We consider the case of transferable utility.
- So, we must select the alternative A_j for which the sum $\sum\limits_{i=1}^n u_i(A_j)$ of the utilities is the largest possible.
- In our case, this means that we compare the values

$$\sum_{i=1}^n \overline{u}_i(A_j) = \sum_{i=1}^n \frac{b_{ij}+1}{k+1}.$$

- This sum is, in its turn, equal to

$$\sum_{i=1}^n \frac{b_{ij}+1}{k+1} = \frac{1}{k+1} \cdot \sum_{i=1}^n b_{ij} + \frac{n}{k+1}.$$

- Thus, the largest value of this sum corresponds to the largest value of the Borda sum $\sum\limits_{i=1}^n b_{ij}$.

26. Comment: in the Simplest Selection Case, We Get a Version of the Borda Count

- What if we have a simple selection?
- In this case, we should select the alternative A_j for which the *product* is the largest:

$$\prod_{i=1}^{n} \overline{u}_i(A_j) = \prod_{i=1}^{n} \frac{b_{ij}+1}{k+1}.$$

- This, in its turn, is equivalent to maximizing the product $\prod\limits_{i=1}^{n}(b_{ij}+1)$, or, alternatively, to maximizing:

$$\sum_{i=1}^{n} \ln(b_{ij}+1).$$

27. Acknowledgments

This work was supported in part by the National Science Foundation grants:

- 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science),
- and HRD-1242122 (Cyber-ShARE Center of Excellence).