Impact of Underlying Voting Distribution in Societal Tradeoff Rules

Qitong Gao

1 Introduction

Societal tradeoff rules can effectively help us to optimize the allocation of resources especially when subjective factors need to be taken into account. In other words, it allows people to vote a value on different activities and generate a general consensus which most of the people will agree with. When aggregating the numbers, we could simply choose the median value of each activities, which is known as the *Median Rule*. This rule has been proved strategy-proof, however, as Conitzer, Brill and Freeman (2015)[1] observed, the aggregate result would not be consistent if multiple values need to be evaluated. Then a *Distance Based Rule* model has been developed by Conitzer, Freeman *et al.* (2016)[2] and can output consistent results. An additive model has been defined based on this rule. It will ask people to vote a value on each pair of activities, and the value represents how much units an activity is preferred to another based on a person's own evaluation. Then the model will output the aggregate value of each pair of activities through maximum likelihood estimation of the "true" tradeoff vector, or in other words, minimizing the distance between the "true" tradeoff and the aggregate tradeoff generated from the model.

However, in real situation, it is hard to obtain a "true" voting distribution which we can test our result with. And according to Conitzer's[2] previous work, additive model could have large penalties if the input have a large number of activities and voters. So it is important to figure out a way to check the reliability of the results from additive model. *Kemeny Rule* had been proved *accurate in the limit* and could estimate the "true" ranking of activities with a high accuracy and with small number of input samples. In this project, I will compare the output from additive model with the output from *Kemeny Rule* to evaluate the reliability of additive model when the "true" distribution is unknown and analyze the sample complexity when the results from additive model consistent with results from *Kemeny Rule*.

2 Model Description

2.1 Preliminaries

Let $N = \{1, 2, ..., n\}$ be a finite set of voters. Let $A = \{a, b, c, d...\}$ be a set of m activities. Let $q_i(a)$ denotes the value voted on activity a by person i and $\mathcal{L}(A)$ be the set of votes voted by person i, where $\mathcal{L}_i(A) = \{q_i(a), q_i(b), q_i(c)...\}$. Define function f_i sorts all the elements in $\mathcal{L}_i(A)$ in a decreasing order, where $f_i : A \to \{1, 2, 3, 4, ..., m\}$. Specifically, $f_i(a)$ is the position of activity a in the set A, for example, $f_i(a) = 1$ means $q_i(a) = \max \mathcal{L}_i(A)$ and $f_i(a) = m$ means $q_i(a) = \min \mathcal{L}_i(A)$, where $i \in N$ and $a \in A$.

Then we could use a weighted directed graph G_i to represent the relationship between different activities voted by person i. For all $i \in N$, in each graph G_i, let all the activities in A be the vertices and there is an edge between each pair of activities (a, b) and $a \to b$ if and only if $q_i(a) > q_i(b)$. In addition, because we prefer G_i to be a complete graph, we expect that for all the voters, $q_i(a) = q_i(b)$ will not happens for any pairs of activities. So we would require voters to vote a slightly different value to a and b if they think a is closely as the same important as b. Let t_i^{ab} be the weight of edge (a, b), where $t_i^{ab} = q_i(a) - q_i(b)$. Let set E_i represents all the edges in graph G_i.

For example, if a CEO let three senior managers to vote on three divisions that a company should focus on in the next year. And the three divisions are marketing, manufacturing and sales. So in this case, $N = \{1, 2, 3\}$, $A = \{a, b, c\}$ and $a =$ marketing, $b =$ manufacturing, $c =$ sales. If manager 1 value a as 15 units, b as 10 units and c as 5 units, then $q_1(a) = 15, q_1(b) = 10$ and $q_1(c) = 5$ and the graph is shown as Figure 1.

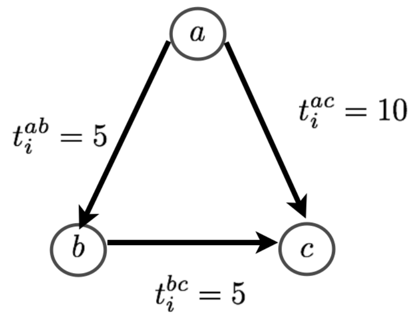

Figure 1: Weighted directed graph shows the tradeoffs

2.2 Noise Model

Differ from the way Conitzer *et al.*[2] defining additive model, where the tradeoff values between pairs of activities, or the value of t_i^{ab}s, are drawn from a "true" distribution P_{true}. In our model, for all $i \in N$, the quality of activity a, $q_i(a)$, voted by person i is drawn i.i.d. from a normal distribution $\mathcal{N}(\mu_a, \sigma)$, where μ_a represents the true quality of activity a. However, the tradeoff values between pairs of activities (a, b) based on person i's vote obey a normal distribution $\mathcal{N}(\mu_a - \mu_b, \sigma)$ which could be used to define an additive model. After all the quantities has been determined, we could sort the elements in $\mathcal{L}_i(A)$ in a decreasing order to obtain f_i. So our model could be simply transferred into additive model by calculating the t_i^{ab}s. And the optimal qualities of activities can be solved by linear programming in polynomial time. In addition, function f_is could sort the optimal qualities into a ranking and enable us to evaluate the consistency between additive model and Kemeny rule.

In our model, the probability of drawing a tradeoff value specific t_i^{ab} is proportional to $e^{-d(t^{ab}, t_i^{ab})^2}$, where $d(t^{ab}, t_i^{ab}) = t_i^{ab} - \mu^{ab}$, $t_i^{ab} = q_i(a) - q_i(b)$ and $\mu^{ab} = \mu_a - \mu_b$.

We could also sort all the $q_i(a), q_i(b), q_i(c), \cdots \in \mathcal{L}_i(A)$ in a decreasing order and obtain a ranking f_i. So there is a bijection between f_i and $\mathcal{L}_i(A)$

and the probability of drawing a ranking f given that the true order is f^* is:

$$\Pr[f|f^*] = \frac{e^{-d(t,t_i)^2}}{Z^m}$$

where $d(t,t_i) = \sum_{(a,b)\in E_i} d(t^{ab}, t_i^{ab})$.

Proposition 1. The MLE estimator of the true ranking is argmin $\sum_{i=1}^{n} d(t,t_i)$, where $d(t,t_i) = \sum_{(a,b)\in E_i} d(t^{ab}, t_i^{ab})$.

Proof.

$$\begin{aligned}
&\max \prod_i \prod_{ab} e^{-(t_i^{ab}-\mu^{ab})^2} \\
\iff &\max \sum_i \sum_{ab} -(t_i^{ab} - \mu^{ab})^2 \\
\iff &\min \sum_i \sum_{ab} |t_i^{ab} - \mu^{ab}| \\
\iff &\min \sum_i d(t,t_i)
\end{aligned}$$

□

3 Additive Model & Kemeny Rule

3.1 Kemeny Rule

Definition 1. *Kemeny Rule.* Given a ranking profile voted by n person $f_1, f_2, \ldots, f_n$, Kemeny Rule will select a ranking $f \in \mathcal{L}(A)$ that minimizes $\sum_{i=1}^{n} d_{KT}(f, f_i)$, where $d_{KT}(f, f_i)$, the $Kendall\ tau$ (KT) distance, could be defined as:

$$d_{KT}(f_1, f_2) = |\{(a,b)|((a \succ_{f_1} b) \wedge (b \succ_{f_2} a)) \vee (a \succ_{f_2} b) \wedge (b \succ_{f_1} a)\}|$$

Generally, the KT distance denotes the number of pairs of activities whose relative position in two votes are not consistent. I. Caragiannis, A. Procaccia and N. Shah[3] had already proved that Kemeny rule is *accurate in the limit* if the noise model is d-monotonic, which means that the output from Kemeny rule based on our noise model would return the correct ranking given an infinite number of samples. More detailed proof could be found in section 3.3.

3.2 Additive Model

The linear program introduced by Conitzer[2] in additive model could be slightly modified to fit our model. It contains variables $q(a)$, $q(b)$ and t_i^{ab}, where $q(a)$ and $q(b)$ denotes the optimal quality for each activity, so $t^{ab} = q(a) - q(b)$, and d_i^{ab} denotes the distance $|t_i^{ab} - t^{ab}|$. The linear program is shown as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} \sum_{(a,b) \in E_i} d_i^{ab} \\ \text{subject to} \quad & d_i^{ab} \geqslant q(a) - q(b) - t_i^{ab} \quad (\forall i, a, b) \\ & d_i^{ab} \leqslant q(a) - q(b) - t_i^{ab} \quad (\forall i, a, b) \end{aligned}$$

3.3 Accurate in the Limit

Definition 2. *Pairwise-Majority Consistent (PM-c) Rules.* If graph G_i is complete and acyclic, then we could say the rule used to define G_i is pairwise-majority consistent.

Proposition 2. Kemeny rule based on the model defined in section 2 is pairwise-majority consistent.

Proof. According to the definition, there is always an edge between each pair of activities. So the graph of additive model should be complete. Now prove the graph is acyclic by contradiction: Assume there exist a circle in graph G_i, say the path is $a \to b \to c \to d \to a$. But according to the definition of G_i, $a \to b$ exists if and only if $q_i(a) > q_i(b)$ so apparently $q_i(a) > q_i(b) > q_i(c) > q_i(d) > q_i(a)$ cannot happens. Consequently, path $a \to b \to c \to d \to a$ should not exist in graph G_i. □

Definition 3. *d-Monotonic Noise Models.* Let f^* denote the "true" ranking. A noise model is called d-monotonic with respect to distance d if for any $f, f' \in \mathcal{L}(A)$, $d(f, f^*) < d(f', f^*)$ implies $\Pr[f|f^*] > \Pr[f'|f^*]$ and $d(f, f^*) = d(f', f^*)$ implies $\Pr[f|f^*] = \Pr[f'|f^*]$.

Proposition 3. The noise model defined in section 2 is d-monotonic noise model.

Proof. According to the definition in section 2.2:

$$\Pr[f|f^*] = \frac{e^{-d(f,f^*)^2}}{Z^m}$$
$$\Pr[f'|f^*] = \frac{e^{-d(f',f^*)^2}}{Z^m}$$

Then apparently:

$$d(f, f^*) < d(f', f^*) \iff \Pr[f|f^*] > \Pr[f'|f^*]$$

and

$$d(f, f^*) = d(f', f^*) \iff \Pr[f|f^*] = \Pr[f'|f^*]$$

□

Definition 4. *Accurate in the limit.* If a voting rule could return the correct ranking given an infinite number of samples, then we say this rule is *accurate in the limit.*

Lemma 1. All the PM-c rules are accurate in the limit with respect to any noise model that is d-monotonic with respect to a distance function d.

Proof. Proved by Caragiannis *et al.*[3] □

Proposition 4. Kemeny rule is *accurate in the limit* based on the noise model defined in section 2.

Proof. According to Proposition 2, Proposition 3 and Lemma 1, our noise model is d-monotonic and Kemeny rule is PM-c based on our noise model. Then if the input quality values generated from the model defined in section 2, the output from Kemeny rule is *accurate in the limit.* □

Based on Proposition 4, we could conclude that the output from Kemeny rule could represent the true ranking with a high accuracy, under the condition that the true distribution remains unknown. In the following experiment section (section 5), the output from Kemeny rule and from additive model will be compared with the true ranking. And the result shows that we could evaluate the accuracy of the ranking calculated by additive model by assuming that the true ranking is the output from Kemeny rule.

4 Sample Complexity

Propositon 5. For any given $\epsilon > 0$, Kemeny rule could determine the true ranking with probability at least $1 - \epsilon$ given $O(ln(m/\epsilon))$ samples from the noise model defined in section 2.

Proof. We need to prove that

$$\Pr[\forall a, b \in A, a \succ_{f^*} b \Longrightarrow n_{ab} - n_{ba} \geqslant 1] \geqslant 1 - \epsilon$$

where $a \succ_f b$ means that in ranking f, activity a is more important than b ($f(a) < f(b)$). n_{ab} denotes the number of rankings f such that $a \succ_f b$.

By using Hoeffding's inequality and the union bound, Caragiannis[3] already proved that

$$\Pr[\exists a, b \in A, \{(a \succ_{f^*})b \wedge (n_{ab} - n_{ba} \leqslant 0)\}] \leqslant m^2 \cdot e^{2 \cdot \delta_{min}^2 \cdot n}$$

where $\delta_{ab} = p_{a \succ b} - p_{b \succ a}$ and $p_{a \succ b}$ denotes that the probability of $a \succ b$ in a random ranking f.

Now we only need to prove that $\delta_{\min} = \Omega(1)$. The detailed proof will be shown in Appendix.

□

5 Experiment

Because of proposition 4 and proposition 5, when the true distribution is remained unknown, we could evaluate the accuracy of the optimal solution from additive model by assuming the underlying true ranking is the output from Kemeny rule. And the difference between these two output could be used to evaluate the additive model.

The following experiment, shown in Figure 2, drawn a true quality from a normal distribution $\mathcal{N}(\mu_a, \sigma)$ for each activity a and shows the difference between the output from additive model and Kemeny rule, the difference between the output from additive model and true values and the difference between the output from Kemeny rule and true values. The difference is estimated by *Footrule Distance*, which is defined as follows:

$$d_{FR}(f_1, f_2) = \sum_{a \in A} |f_1(a) - f_2(a)|$$

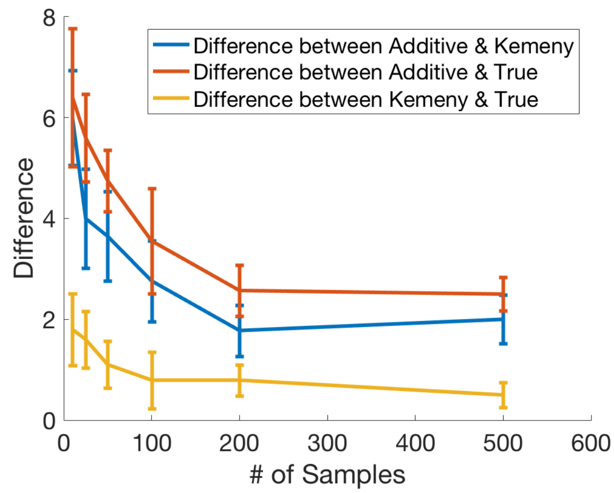

Figure 2: Experiment

The number of activities is 10 in this experiment and the distance is evaluated at the sample size (number of voters) of 10, 25, 50, 100, 200 and 500. So the maximum of $d_{FR}(f_1, f_2)$ could attain is 52, when f_1 and f_2 has completely reverse order. Therefore, the yellow line clearly shows the high accuracy of Kemeny rule and the blue line and red line shows the difference between additive & true could be well represented by the difference between additive & Kemeny.

6 Conclusion

This project build a bridge to connect Kemeny rule and additive method. And it enables us to evaluate the accuracy of the results from additive model without using the true distribution. Additive model has the advantage of optimizing the qualities quantitatively and Kemeny rule has the advantage of high accuracy. The model defined in this project combines these two rules, as well as the advantages, together. So we can not only obtain specific values on qualities of activities, not just a ranking from Kemeny rule, but also obtain the optimal tradeoff with a high accuracy.

7 References

1. V. Conitzer, B. Markus, and R. Freeman. "Crowdsourcing societal tradeoffs." *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems.*

2. V. Conitzer, R. Freeman, B. Markus and Y. Li. "Rules for Choosing Societal Tradeoffs" *Forest 2016: Volume 100, Page 200.*

3. I. Caragiannis, A. Procaccia, N. Shah. "When do noisy votes reveal the truth?" *Proceedings of the fourteenth ACM conference on Electronic commerce.*

8 Appendix

Proof of $\delta_{\min} = \Omega(1)$:

$$\begin{aligned}
\delta_{ab} &= p_{a \succ b} - p_{b \succ a} \\
&= \sum_{f \in \mathcal{L}(A) | a \succ_f b} \Pr[f | f^*] - \sum_{f \in \mathcal{L}(A) | b \succ_f a} \Pr[f | f^*] \\
&= \sum_{f \in \mathcal{L}(A) | a \succ_f b} (\Pr[f | f^*] - \Pr[f_{a \leftrightarrow b} | f^*]) \\
&= \sum_{f \in \mathcal{L}(A) | a \succ_f b} \frac{e^{-d^2(t,t^*)} - e^{-d^2(t_{a \leftrightarrow b},t^*)}}{Z^m} \\
&> \sum_{f \in \mathcal{L}(A) | a \succ_f b} \frac{e^{-d^2(t,t^*)} - e^{-d^2(t,t^*)+k}}{Z^m} \\
&= \sum_{f \in \mathcal{L}(A) | a \succ_f b} \frac{e^{-d^2(t,t^*)}(1 - e^k)}{Z^m} \\
&= (1 - e^k) p_{a \succ b} \\
&= (1 - e^k) \left(\frac{1 + \delta_{ab}}{2} \right)
\end{aligned}$$

where $f_{a \leftrightarrow b}$ means that in ranking f, the position of a and b has been swapped. The fourth transition follows $d(t_{a \leftrightarrow b}, t^*) \geqslant d(t, t^*) + k$ for any $k > 0$.

So we could obtain

$$\delta_{ab} > (1 - e^k) \left(\frac{1 + \delta_{ab}}{2} \right)$$

$$\delta_{ab} > \frac{1 - e^k}{1 + e^k}$$

which apparently shows $\delta_{\min} = \Omega(1)$.