Representation Theory of Symmetric Groups

Due on June 11, 2020

Professor Mei Si

Juncheng Wan

1 Introduction

In this paper, I study the use of the representation theory of the symmetric group in voting theory [Crisman and Orrison, 2015]. The reason I choose this paper is that the representation on symmetric groups naturally applies to voting theory. Besides, as I major in Computer Science, I will deal with many ranking problems closely related to voting theory. For example, I will deal with realistic problems like the recommended ordering of goods in the recommendation system, the word prediction in machine translation. In these problems, the candidate set is usually much large, from 10,000 to 100,000 and I do not just select the top one. Because designing a model to rank the data can usually make better use of the information. I hope that the research on representation theory of symmetric group in voting theory can help me understand the ranking process.

2 Definition

In this section, I introduce some definitions used in this paper. The paper formulates the problem of voting mathematically.
First, there are some basic definitions.

- **Composition**: A composition of a positive number n is a list $\lambda = (\lambda_1, \ldots, \lambda_m)$, where $\sum_{i=1}^{m} \lambda_i = n$ and λ_i is positive.
- **Partition**: A partition of n is a list $\lambda = (\lambda_1, \ldots, \lambda_m)$, where $\sum_{i=1}^{m} \lambda_i = n$ and $\lambda_1 \geq \cdots, \geq \lambda_m \geq 1$.
- **Diagram**: A Diagram of composition λ is the left-justified array of boxes that has λ_i boxes in its ith row.
- **Tableau**: A Tableau of composition λ is a Diagram which filled with the numbers $1, 2, \ldots, n$ without repetition.
- **Tabloid**: A Tabloid is an equivalence class of Tableau. Two tableaux of shape $\lambda = (\lambda_1, \ldots, \lambda_m)$ are said to be row equivalent if they have the same set of λ_i numbers for each row i.
- X^λ: The set of tabloids of shape λ.
- M^λ: The vector space of real-valued functions defined on X^λ.
- $\mathbf{f}_x$: The indicator function of M^λ with the property that $\mathbf{f}_x(x) = 1$ and $\mathbf{f}_x(y) = 0$ for all $y \neq x$. Note that the set $\{\mathbf{f}_x \in M^\lambda : x \in X^\lambda\}$ is an orthonormal basis of M^λ with respect to the usual inner product $\langle \cdot, \cdot \rangle$ on M^λ, which is defined by setting $\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{x \in X^\lambda} \mathbf{f}(x)\mathbf{g}(x)$.
- **An action of S_n on M^λ**: If $\sigma \in S_n, \mathbf{f} \in M^\lambda$ and x^λ, then $(\sigma \cdot f)(x) = \mathbf{f}(\sigma^{-1} \cdot x)$ is an action of S_n on M^λ. From this definition, M^λ may be viewed as a module over the group algebra $\mathbb{R}S_n$.
- **Voting Method T**: T is a linear transformation from $M^{(1,\ldots,1)}$ to M^λ. When $\lambda = (1, n-1)$, T uses the information of voting result in $M^{(1,\ldots,1)}$ to assign points to each of the individual candidates. When $\lambda = (1, \ldots, 1)$, T uses the information of voting result in $M^{(1,\ldots,1)}$ to assign points to full rankings.
- **Neutral Voting Method**: A voting method when the outcome does not depend on the labels of candidates. For example, when T is an $\mathbb{R}S_n$-module homomorphism, T is a neural voting method.
- **Winning Ranking**: A ranking that receives at least as many points as all of the other rankings. This is opposed to simply a winner.
- **Specht Module S^μ**: The irreducible module of $\mathbb{R}S_n$-modules, corresponding to the partition μ.

- **Kostka Number**: If λ is a partition of n, then M^λ is isomorphic to a direct sum of Specht moduls, i.e. $M^\lambda \cong \bigoplus_\mu \kappa_{\mu\lambda} S^\mu$. Each $\kappa_{\mu\lambda}$ is a Kostka number and records the multiplicity of each Specht module in M^λ.

- U_0: An irreducible module of $M^{(1,n-1)}$ which is isomorphic to $S^{(n)}$. $U_0 = \{\mathbf{f} \in M^{(1,n-1)} | \ \mathbf{f}(x) = \mathbf{f}(y) \text{ for all } x, y \in X^{(1,n-1)}\}$.

- U_1: An irreducible module of $M^{(1,n-1)}$ which is isomorphic to $S^{(n-1,1)}$. $U_1 = \{\mathbf{f} \in M^{(1,n-1)} | \ \sum_{x \in X^{(1,n-1)}} \mathbf{f}(x) = 0\}$. Note that U_1 is also the orthogonal complement of U_0.

- $\hat{\mathbf{f}}$: The projection of $\mathbf{f}$ into U_1.

- x_0: The tabloid in some X^λ that appears first when the tabloids in X^λ are listed lexicographically. This tabloid contains the tableau whose entries, when read from left to right and top to bottom, are the numbers $1, \ldots, n$ in that order.

- **Effective Space** $E(T)$: The submodule $(\ker T)^\perp$. Note that $M^\lambda = \ker T \oplus (\ker T)^\perp$ and $E(T)$ is isomorphic to the image of T.

Note that I do not fully understand Specht module. I think these two papers Specht [1935]; Peel [1975] are good materials. Besides, there are some definitions about voting theory.

- **Profile**: A function $\mathbf{p} \in M^{(1,\ldots,1)}$ where $\mathbf{p}(x)$ is the number of voters who chose the tabloid x.

- **Weighting Vector**: A column vector $\mathbf{w} = [w_1, \ldots, w_n]^t$ in $\mathbb{R}^n$ such that $w_1 \geq \cdots \geq w_n$.

- **Equivalent Weighting Vectors**: Two weighting vectors $\mathbf{w}$ and $\mathbf{w}'$ are equivalent if there exist $\alpha, \alpha' \in \mathbb{R}$ such that $\alpha > 0$ and $\mathbf{w}' = \alpha\mathbf{w} + \alpha'\mathbf{1}$. We write $\mathbf{w} \sim \mathbf{w}'$.

- **Positional Voting**: This a voting procedure. For a given weighting vector $\mathbf{w}$, we can calculate the points of candidate i as the sum over all the tabloids in $X^{(1,\ldots,1)}$, i.e. $\sum_{x \in X^{(1,\ldots,1)}} \mathbf{p}(x)\mathbf{w}^{(i)}(x)$, where $\mathbf{w}^{(i)}(x)$ is the weight of the row the candidate i in x.

- **Plurality Voting**: A special case of positional voting where $\mathbf{w} = [1, 0, \ldots, 0]^t$.

- **Anti-plurality Voting**: A special case of positional voting where $\mathbf{w} = [1, 1, \ldots, 1, 0]^t$.

- **Borda Count**: A special case of positional voting where $\mathbf{w} = [n-1, n-2, \ldots, 2, 1, 0]^t$.

- **Kendall Tau Distance**: A distance function $X^{(1,\ldots,1)} \times X^{(1,\ldots,1)} \to \mathbb{R}$ such that $d(x, y) = \binom{n}{2} - \sum_{i \neq j} \mathbf{a}_{ij}(x)\mathbf{a}_{ij}(y)$, where $\mathbf{a}_{ij} \in M^{(1,\ldots,1)}$ is defined by setting $\mathbf{a}_{ij} = 1$ whenever candidate i is ranked above j in x and $\mathbf{a}_{ij} = 0$ otherwise.

- K: Define $T_\mathbf{z} : M^{(1,\ldots,1)} \to M^{(1,\ldots,1)}$ by setting $T_\mathbf{z}(f) = \sum_{\sigma \in S_n} \mathbf{f}(\sigma \cdot x_0)(\sigma \cdot \mathbf{z}) = \tilde{\mathbf{f}} \cdot \mathbf{z}$ for all $\mathbf{f} \in M^{(1,\ldots,1)}$, where $z \in M^{(1,\ldots,1)}$. If $\mathbf{z}(x) = \sum_{i \neq j} \mathbf{a}_{ij}(x)\mathbf{a}_{ij}(x_0)$, we denote $K = T_\mathbf{z}$ for this special case.

3 Theorem

In this section, I've written down some useful theorems. First, there are three theorems revealing the structures of $M^{(1,\ldots,1)}$, $M^{(k,n-k)}$, and $M^{(n)}$.

Theorem 3.1. *$M^{(1,\ldots,1)}$ is isomorphic to the regular $\mathbb{R}S_n$-module, and thus $M^\lambda \cong \bigoplus_\mu (\dim S^\mu) S^\mu$.*

Theorem 3.2. *If $1 \leq k \leq n/2$ (so that $(n-k, k)$ is a partition of n), then $M^{(k,n-k)} \cong M^{(n-k,k)} \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus \cdots \oplus S^{(n-k,k)}$.*

Theorem 3.3. *$M^{(n)}$ corresponds to the trivial $\mathbb{R}S_n$-module, thus $M^{(n)} \cong S^{(n)}$.*

The following two theorems say that the positional voting procedure can be viewed as the group algebra $\mathbb{R}S_n$ acting on $M^{(1,n-1)}$. From my perspective, the weighting vector $\mathbf{w}$ itself is a voting result in $M^{(1,n-1)}$ and the profile $\mathbf{f}$ can be viewed as an element in $\mathbb{R}S_n$.

Theorem 3.4. *Given the weighting vector $\mathbf{w}$, the positional voting procedure is the module homomorphism $T_{\mathbf{w}} : M^{(1,\dots,1)} \to M^{(1,n-1)}$ by setting $T_{\mathbf{w}}(\mathbf{f})) = \langle \mathbf{f}, \mathbf{w}^{(1)} \rangle \mathbf{c}_1 + \cdots + \langle \mathbf{f}, \mathbf{w}^{(n)} \rangle \mathbf{c}_n$, where $\mathbf{c}_i \in M^{(1,n-1)}$ denotes the indicator function of the tabloid that candidate i in the top row.*

Theorem 3.5. *The positional voting procedure result $T_{\mathbf{w}}(\mathbf{f})$ is also the result of the group algebra element $\tilde{\mathbf{f}} \in \mathbb{R}S_n$ acting on the function $\mathbf{w} \in M^{(1,n-1)}$. Specifically, the weighting vector is identified with the function $\sum_{i=1}^{n} w_i \mathbf{c}_n \in M^{(1,n-1)}$ and $\tilde{\mathbf{f}} = \mathbf{f}(\sigma \cdot x_0)$.*

Proof.

$$\begin{aligned}
T_{\mathbf{w}}(\mathbf{f}) &= \sum_{i=1}^{n} \langle \mathbf{f}, \mathbf{w}^{(i)} \rangle \mathbf{c}_i \\
&= \sum_{i=1}^{n} \Big(\sum_{x \in X^{(1,\dots,1)}} \mathbf{f}(x) \mathbf{w}^{(i)}(x) \Big) \mathbf{c}_i \\
&= \sum_{x \in X^{(1,\dots,1)}} \mathbf{f}(x) \sum_{i=1}^{n} \mathbf{w}^{(i)}(x) \mathbf{c}_i \\
&= \sum_{\sigma \in S_n} \mathbf{f}(\sigma \cdot x_0) \sum_{i=1}^{n} \mathbf{w}^{(i)}(\sigma \cdot x_0) \mathbf{c}_i \\
&= \sum_{\sigma \in S_n} \mathbf{f}(\sigma \cdot x_0) \sum_{i=1}^{n} w_i \mathbf{c}_{\sigma(i)} \\
&= \sum_{\sigma \in S_n} \mathbf{f}(\sigma \cdot x_0) \sum_{i=1}^{n} w_i \mathbf{c}_{\sigma(i)} \\
&= \sum_{\sigma \in S_n} \tilde{\mathbf{f}}(\sigma)(\sigma \cdot \mathbf{w}) \\
&= \tilde{\mathbf{f}} \mathbf{w}
\end{aligned} \tag{1}$$

□

The following three theorems reveal that if two weighting vectors are not equivalent, there always exist some profiles that their corresponding voting results by the two weighting vectors can be arbitrarily different. The proof of these theorems should not be difficult by construction.
From these three theorems, I realize that when I rank the goods in recommendation system, when the weighting vector of model is not equivalent to "the best weighting vector", the ranking results can be arbitrarily bad in the worst case.

Theorem 3.6. *Let $n \geq 2$, and suppose $\mathbf{w}_1, \dots, \mathbf{w}_k \in U_1 \subset M^{(1,n-1)}$ is linearly independent. For arbitrary $\mathbf{r}_1, \dots, \mathbf{r}_k \in U_1 \subset M^{(1,n-1)}$, there exist infinitely many functions $\mathbf{f} \in M^{(1,\dots,1)}$ such that $T_{\mathbf{w}_i}(\mathbf{f}) = \mathbf{r}_i$ for all i such that $1 \leq i \leq k$.*

Theorem 3.7. *Let $\mathbf{w}, \mathbf{w}' \in M^{(1,n-1)}$ be weighting vectors. The ordinal rankings of $T_{\mathbf{w}}(\mathbf{p})$ and $T_{\mathbf{w}'}(\mathbf{p})$ will be the same for all profiles $\mathbf{p} \in M^{(1,\dots,1)}$ if and only if $\mathbf{w} \sim \mathbf{w}'$.*

Theorem 3.8. *Let $\mathbf{w}, \mathbf{w}' \in U_1 \subset M^{(1,n-1)}$ be nonzero weighting vectors. Then two effective spaces $E(T_{\mathbf{w}}) = E(T_{\mathbf{w}'})$ if and only if $\mathbf{w} \sim \mathbf{w}'$. Furthermore, if $E(T_{\mathbf{w}}) \neq E(T_{\mathbf{w}'})$, then $E(T_{\mathbf{w}}) \cap E(T_{\mathbf{w}'}) = \{\mathbf{0}\}$.*

Kendall tau distance can be used to do full ranking.

Theorem 3.9. *Kendall tau distance is invariant under the action of S_n.*

Proof. Let $\sigma \in S_n$. Notice that if $i \neq j$, we have:

$$\mathbf{a}_{ij}(x)\mathbf{a}_{ij}(y) + \mathbf{a}_{ji}(x)\mathbf{a}_{ji}(y) = \mathbf{a}_{\sigma(i)\sigma(j)}(x)\mathbf{a}_{\sigma(i)\sigma(j)}(y) + \mathbf{a}_{\sigma(j)\sigma(i)}(x)\mathbf{a}_{\sigma(j)\sigma(i)}(y) \tag{2}$$

□

Theorem 3.10. *For a profile $\mathbf{p} \in M^{(1,\ldots,1)}$, we have $K(\mathbf{p}) = \sum_{i\neq j}\langle \mathbf{p}, \mathbf{a}_{ij}\rangle \mathbf{a}_{ij}$.*

Proof. Notice that Kendall tau distance is invariant under the action of S_n. Using the similar proof as above. □

The following constructive proof says that any positional voting procedure have the same result of a voting procedure of full rankings.

Theorem 3.11. *For any positional voting procedure $T_{\mathbf{w}}$, where $\mathbf{w} \in M^{(1,n-1)}$ there exist a voting procedure of full rankings that have the same result.*

Proof. Let $\mathbf{b} \in M^{(1,n-1)}$ be the weighting vector for the Borda Count:

$$\mathbf{b} = \sum_{i=1}^{n}(n-i)\mathbf{c}_i \tag{3}$$

Next, recall that:

$$T_{\mathbf{w}}(\mathbf{f}) = \sum_{i=1}^{n}\langle \mathbf{f}, \mathbf{w}^{(i)}\rangle \mathbf{c}_i \tag{4}$$

Let the adjoint $T^*_{\mathbf{B}} : M^{(1,n-1)} \to M^{(1,\ldots,1)}$, then we create the map $T^*_{\mathbf{B}} \circ T_{\mathbf{w}} : M^{(1,\ldots,1)} \to M^{(1,\ldots,1)}$ where:

$$T^*_{\mathbf{B}} \circ T_{\mathbf{w}}(\mathbf{f}) = \sum_{i=1}^{n}\langle \mathbf{f}, \mathbf{w}^{(i)}\rangle \mathbf{b}_i \tag{5}$$

Thus, $T^*_{\mathbf{B}} \circ T_{\mathbf{w}}$ is the corresponding result. □

Following three theorems give the structure of $M^{(1,1,n-2)}$, the effective space of a full ranking to pair-wise ranking voting procedure, and the effective space of Borda Count. Some details are in [Daugherty *et al.*, 2009].

Theorem 3.12. $M^{(1,1,n-2)} \cong S^{(n)} \oplus 2S^{(n-1,1)} \oplus S^{(n-2,2)} \oplus S^{(n-2,1,1)}$

Theorem 3.13. *Let $\mathbf{c}_{ij} \in M^{(1,1,n-2)}$ be the indicator function corresponding to the tabloid that has i in the top row and j in the second row. The map $P : M^{(1,\ldots,1)} \to M^{(1,1,n-2)}$ is given by $P(\mathbf{f}) = \sum_{i\neq j}\langle \mathbf{f}, \mathbf{a}_{ij}\rangle \mathbf{c}_{ij}$. Then the effective space $E(P) \cong S^{(n)} \oplus S^{(n-1,1)} \oplus S^{(n-2,1,1)}$.*

Theorem 3.14. *Let $\mathbf{b}$ be the weighting vector of Borda Count. Then the effective space $E(T_{\mathbf{b}}) \cong S^{(n)} \oplus S^{(n-1,1)}$.*

The following theorem is an interesting result of Borda count and Kemeny rule (full ranking procedure with Kendall tau distance). This means that when adopting easy voting procedure, using Borda count can guarantee that the top-ranked candidate and bottom-ranked candidate have the same order compared with Kemeny rule. However, if weighting vector not equivalent to Borda count, this is not guaranteed.

Theorem 3.15. *For $n \geq 3$ candidates, the Borda count always ranks the Kemeny rule top-ranked candidate strictly above the Kemeny rule bottom-ranked candidate. Conversely, the Kemeny rule ranks the Borda count top-ranked candidate strictly above the Borda count bottom-ranked candidate. For any positional voting method other than the Borda count, however, there is no relationship between the Kemeny rule ranking and the positional ranking.*

4 Conclusion

In this study, I think that the representation theory of $\mathbb{R}S_n$ is rather important for voting theory. We need to the structure of the vector space of real-valued funtions defined on X^λ, which is the direct sum of some Specht modules S^μ. In voting procedure, Borda count is closely related to Kemeny rule.

References

Karl-Dieter Crisman and Michael E. Orrison. Representation theory of the symmetric group in voting theory and game theory. *CoRR*, abs/1508.05891, 2015.

Zajj Daugherty, Alexander K Eustis, Gregory Minton, and Michael E Orrison. Voting, the symmetric group, and representation theory. *The American Mathematical Monthly*, 116(8):667–687, 2009.

MH Peel. Specht modules and symmetric groups. *Journal of Algebra*, 36(1):88–97, 1975.

Wilhelm Specht. Die irreduziblen darstellungen der symmetrischen gruppe. *Mathematische Zeitschrift*, 39(1):696–711, 1935.