12-21-2017

Reconstructing Results From Voting Theory Using Linear Algebra

Brian Camara

Reconstructing Results From Voting Theory Using Linear Algebra

Brian Camara

Submitted in Partial Completion of the
Requirements for Departmental Honors in Mathematics

Bridgewater State University

December 21, 2017

Dr. John Pike, Thesis Director
Dr. Ward Heilman, Committee Member
Dr. Jacqueline Anderson, Committee Member

RECONSTRUCTING RESULTS FROM VOTING THEORY USING LINEAR ALGEBRA

BRIAN CAMARA

For many undergraduate students, achieving an understanding of upper-level mathematics can be extremely challenging. For us, it helps to connect these new concepts with material we are familiar with. This will be the central theme of this thesis. We will introduce some basic components of algebraic voting theory, along with briefly discussing how (Daugherty, Eustis, Minton, & Orrison, 2009) used representation theory to achieve their results. We will then provide an alternative proof to the main result of the (Daugherty et al., 2009) article using linear algebra, which should be much more familiar to my peers. We will carry out a similar process with a result from (Sarri, 1992), and attempt to reach his result with linear algebra techniques as well.

To begin, we will define some concepts that I will be using throughout this paper. We will define $V_0 = \{x \in \mathbb{Q}^n : \sum_{i=1}^{n} x_i = 0\}$ to be the sum-zero subspace. The *braid arrangement* $\{H_{ij}\}_{1 \leq i \leq j \leq n}$ is defined by $H_{ij} = \{x \in \mathbb{Q}^n : x_i = x_j\}$, and for $\sigma \in S_n$, we define a *chamber* of the braid arrangement $C_\sigma = \left\{x \in \mathbb{Q}^n : x_{\sigma(1)} > x_{\sigma(2)} > \ldots > x_{\sigma(n)}\right\}$. Also, note that the *fundamental chamber* of the braid arrangement is $C_{id} = \{x \in \mathbb{Q}^n : x_1 > x_2 > \ldots > x_n\}$. We will also define the following, $H_{ij}^0 = H_{ij} = \{x \in \mathbb{Q}^n : x_i = x_j\}$, $H_{ij}^+ = \{x \in \mathbb{Q}^n : x_i > x_j\}$, and$H_{ij}^- = \{x \in \mathbb{Q}^n : x_i < x_j\}$. We define a *face* of the braid arrangement as any non-empty intersection of hyperplanes and half spaces ranging over all pairs (i, j).

Taking a step back, consider a possible election with n candidates. We define a ranking to be any possible ordering of the candidates from first to last. Note that the set of all rankings forms a natural bijection with the symmetric group S_n, which is a group that contains all of the possible orderings of n objects. For the purpose of this paper, we will use permutations when referring to rankings. For any candidates $n \geq 3$ in an election, we define a *profile* to represent how many members of society prefer each ranking. Considering all possible rankings, we can determine the winner of this election by tallying up the total votes for each ranking, all while applying whichever voting method preferred by us. I will also define R_ℓ to be the permutation matrix corresponding with the permutation σ_ℓ, which we are labeling based on the lexicographic order of their one-line notations. I will also let $Q_p = \sum_{\ell=1}^{n!} p_\ell R_\ell$ for $p \in \mathbb{Q}^{n!}$.

Our next task will be to visualize these profiles. The most advantageous way would be to look at these profiles as column vectors in $\mathbb{Q}^{n!}$, which would look like the following: $p = \begin{bmatrix} 3 & 2 & 0 & 2 & 0 & 4 \end{bmatrix}^T$. Here the k^{th} entry of p corresponds to the permutation σ_k.

Profiles are a key aspect of voting theory along with *weighting vectors*, which we will now discuss. We will let $w = \begin{bmatrix} w_1 & \dots & w_n \end{bmatrix}^T \in C_{id}$, and suppose we are given a specific profile $p \in \mathbb{Q}^{n!}$, where p_k corresponds to the number of voters that prefer the ranking σ_k. The weighting vector, w, can be used to assign points to each of our candidates. Lets look at an example. Consider the following full rankings:

$$3 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} + 0 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} + 4 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Note that this give us an easy way to visualize the permutations of the candidates, with the 1st place candidate being at the top, 2nd place below 1st, and so on. Now, lets introduce a weighting vector $w = \begin{bmatrix} 1 & s & 0 \end{bmatrix}$. This means that first place gets 1, second place gets s, third place gets 0. If we were to calculate the points for each candidate, candidate 1 would get 5 points, candidate 2 would get $2 + 7s$ points, and the third candidate would get $4 + 4s$ points. Note that choosing different values for s will yield different results. If $s = 0$, which in election terms would mean vote for your favorite candidate only, then candidate 1 would win the election. If $s = 1$, meaning vote for everyone but your least favorite, then candidate 2 wins. If $s = \frac{1}{2}$, which would just represent a simple ordinal ranking, then candidate 3 would win. This is an example of a voting paradox. We can compute such the outcome of such a *positional voting procedure* in terms of the profile vector being acted upon by the matrix $T_w = \begin{bmatrix} \sigma_1 w & \cdots \sigma_{n!} w \end{bmatrix}$ having ℓ^{th} column $\sigma_\ell w = \begin{bmatrix} w_{\sigma_\ell^{-1}(1)} & \cdots & w_{\sigma_\ell^{-1}(n)} \end{bmatrix}^T = R_\ell w$. For our running example,

$$T_w p = \begin{bmatrix} 1 & 1 & s & 0 & s & 0 \\ s & 0 & 1 & 1 & 0 & s \\ 0 & s & 0 & s & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 5 \\ 2+7s \\ 4+4s \end{bmatrix}.$$

Note than we can write

$$T_w p = \begin{bmatrix} R_1 w & \cdots & R_{n!} w \end{bmatrix} p = p_1 R_1 w + ... + p_{n!} R_{n!} w =: Q_p w$$

where R_ℓ represents the $n \times n$ permutation matrix associated with σ_ℓ (so that the ij entry of Q_ℓ is 1 if $\sigma_\ell(j) = i$ and 0 otherwise). So far we have thought of positional voting procedures as weighting vectors acting on profiles, but this shows that we can also look at it in terms of profiles acting on weighting vectors. The importance of this observation cannot be overstated. From our previous example,

$$T_w p = \begin{bmatrix} 1 & 1 & s & 0 & s & 0 \\ s & 0 & 1 & 1 & 0 & s \\ 0 & s & 0 & s & 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix}$$

$$= 3 \begin{bmatrix} 1 \\ s \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ 0 \\ s \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 1 \\ s \end{bmatrix} + 4 \begin{bmatrix} 0 \\ s \\ 1 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 6 \\ 2 & 7 & 2 \\ 4 & 4 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ s \\ 0 \end{bmatrix}$$

.

A representation theoretic version of this realization led the authors of (Daugherty et al., 2009) to their main theorem:

Theorem 1 (Daugherty, Eustis, Minton, Orrison)**.** *Let $n \geq 2$, and let $\lambda = (\lambda_1, ..., \lambda_m)$ be a partition of n. Suppose that $w_1, ..., w_k$ form a linearly independent set of sum-zero weighting vectors in $\mathbb{Q}^m$. If $r_1, ..., r_k$ are any sum-zero results vectors in $\mathbb{Q}^n$, then there exist infinitely many profiles $p \in \mathbb{Q}^{n!}$ such that $T_{w_i}p = r_i$ for all $1 \leq i \leq k$.*

In other terms, this theorem is stating that given linearly independent sum-zero weighting vectors and any sum-zero results vectors, there are infinitely many profiles where each of the results vectors can be obtained by using one of the weighting vectors. The proof given for this theorem relies on results from representation theory that are beyond the scope of what undergraduate math majors typically encounter. Before we provide our proof, we will introduce some propositions that will be key to our proof.

Remark 1- There is no loss of generality in working in the sum-zero subspace V_0. Note that we can write each weighting vector $w = \overline{w} + w_0$ with $\overline{w} \in C_{id} \cap V_0$ and w_0 a constant vector. Therefore, $Q_p w = Q_p\overline{w} + Q_p w_0$, with $Q_p w_0$ a constant vector, so $Q_p\overline{w}$ and $Q_p w$ always lie in the same face.

Proposition 1 *Every matrix with equal row and column sums is a linear combination of permutation matrices*

Proof of Proposition 1. Given a matrix with $\sum_{k=1}^{n} M_{ik} = \sum_{k=1}^{n} M_{kj} = C$ for some constant C and all $i, j \in [n]$, denoting the all-ones matrix as O, we can find constants a, b so that $P = a(M + bO)$ is a doubly stochastic matrix. Then, P and $\frac{1}{n}O$ are doubly stochastic and thus convex combination of permutation matrices by the Birkhoff-von Neuman theorem, so $M = a^{-1}P - nb \cdot \frac{1}{n}O$ is a linear combination of permutation matrices.

□

For our proof of the main theorem to work, we need the existence of $n - 1$ sum-zero weighting vectors, which we will now prove.

Proposition 2 *There exist $n - 1$ linearly independent vectors in $C_{id} \cap V_0$*

Proof of Proposition 2. C_{id} is n dimensional, so there exist linearly independent vectors $v_1, \ldots, v_n \in C_{id}$. There exists a unique set of scalars $\alpha_1, \ldots, \alpha_n$, not all 0, such that $\sum_{k=1}^{n} \alpha_k v_k = \mathbf{1}$. By relabeling the v_k if necessary, we can assume that $\alpha_n \neq 0$. Now, for $k = 1, \ldots, n-1$, define $\beta_k = \frac{1}{n} \langle v_k, \mathbf{1} \rangle$ and $w_k = v_k - \beta_k \mathbf{1}$. By construction, each w_k lies in $C_{id} \cap V_0$. Now suppose that there exist $\gamma_1, \ldots, \gamma_{n-1}$ such that $\sum_{k=1}^{n-1} \gamma_k w_k = \mathbf{0}$. Writing $\eta = \sum_{k=1}^{n-1} \gamma_k \beta_k$, we have

$$
\begin{aligned}
0 &= \sum_{k=1}^{n-1} \gamma_k w_k = \sum_{k=1}^{n-1} \gamma_k v_k - \sum_{k=1}^{n-1} \gamma_k \beta_k \mathbf{1} = \sum_{k=1}^{n-1} \gamma_k v_k - \eta \sum_{k=1}^{n} \alpha_k v_k \\
&= \sum_{k=1}^{n-1} \left(\gamma_k - \eta \alpha_k\right) v_k - \eta \alpha_n v_n.
\end{aligned}
$$

As $v_1, \ldots, v_n$ are linearly independent, we must have that the coefficients are all 0. In particular, since $\alpha_n \neq 0$, $\eta = 0$, so for each $k = 1, \ldots, n-1$, $\gamma_k = \gamma_k - \eta \alpha_k = 0$. This shows that $w_1, \ldots, w_{n-1}$ are linearly independent. □

Now, we have all of the tools to produce our proof of the main theorem from (Daugherty et al., 2009).

Our proof of Theorem 1. We have already discussed that we will only be working with full rankings. If needed, as (Daugherty et al., 2009) stated, we could always translate partial rankings to full rankings. we will work with full rankings and take $k = n - 1$. We let $w_1, \ldots, w_k$ be a linearly independent set of sum-zero weighting vectors, with $w_0 = [1, \ldots, 1]^t$. Similarly, let $r_1, ..., r_k$ be sum-zero results vectors with $r_0 = [1, \ldots, 1]^t$.Writing $F = [w_0, \ldots, w_{n-1}]$ and $R = [r_0, \ldots, r_{n-1}]$, we define $Q = RF^{-1}$. Note that F is invertible because our vectors used to create F are linearly independent. By assumption, Q satisfies $Qw_k = r_k, k = 0, 1, \ldots, n-1$. Our goal is to show that Q is a linear combination of permutation matrices, which we can accomplish by showing that the rows and columns of Q sum to 1. We know that $Qw_0 = r_0$ by construction, so the rows of

Q must sum to 1:

$$\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix} = Q\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix} = \begin{bmatrix}Q_{11}+...+Q_{1n}\\ Q_{21}+...+Q_{2n}\\ \vdots\\ Q_{n1}+...+Q_{nn}\end{bmatrix}$$

Now, for the columns to sum to 1, we need ${w_0}^t RF^{-1} = {w_0}^t$. Multiplying the column vector ${w_0}^t$ by the rows of R, we get

$${w_0}^t R = [\langle w_0, r_0\rangle, \langle w_0, r_1\rangle \dots \langle w_0, r_{n-1}\rangle] = [n, 0, ..., 0] =: x^t.$$

This is true since w_0 is orthogonal to the sum-zero vectors. Next, we have that

$${w_0}^t RF^{-1} = x^t F^{-1} = [nF_{11}^{-1}, nF_{12}^{-1}, ..., nF_{1n}^{-1}] =: n{f_1}^t.$$

Keep in mind that f_1^t is the first row of F^{-1}.Given the identity property $F^{-1}F = I$, we can conclude that

$$[\langle f_1, w_0\rangle, \langle f_1, w_1\rangle \dots \langle f_1, w_{n-1}\rangle] = [1, 0, ..., 0].$$

From this result, we see that f_1 is orthogonal to the sum-zero subspace, and the entries of f_1 sum to 1, hence

$$f_1 = \begin{bmatrix}\frac{1}{n}\\ \frac{1}{n}\\ \vdots\\ \frac{1}{n}\end{bmatrix}$$

Therefore, ${w_0}^t RF^{-1} = nf_1^t = [1, ..., 1] = {w_0}^t$. We have shown that the rows and columns sum to 1. By proposition 1, Q is a linear combination of permutation matrices. Since the dimension of the space of doubly stochastic $n \times n$ matrices is $(n-1)^2$ and there are $n!$ permutation matrices, there are infinitely many profiles $p \in \mathbb{Q}^{n!}$ with

$$Q = Q_p := p_1 R_1 + \ldots + p_{n!} R_{n!}. \qquad \square$$

Theorem 2 (Sarri). *There exists infinitely many profiles $p \in \mathbb{Q}^{n!}$ such that for every $\pi \in S_n$ satisfying $\pi(1) > \pi(n)$, there is some $w \in C_{id} \cap V_0$ with $T_{w(\pi)}p \in C_\pi$.*

Proof. To begin, define $f_k = e_k - e_{k+1}$, $k = 1, \dots, n-1$ where $e_1, \dots e_n$ are the standard basis vectors in $\mathbb{Q}^n$, and let $w_1, \dots w_{n-1}$ be as in Theorem 1 with each $r_k = f_k$. We fix $\pi \in S_n$ with $\pi(1) > \pi(n)$ and write $a = \pi(1), b = \pi(n)$. The result will follow if we can find $\alpha_1, \dots, \alpha_{n-1}$ so that

$$w = w(\pi) = \sum_{k=1}^{n-1} \alpha_k w_k \in C_{id} \cap V_0$$

and

$$s = s(\pi) = T_w p = Q_p w = \sum_{k=1}^{n-1} \alpha_k f_k \in C_\pi.$$

Now, for $1 \le i < j \le n$, define $f_{ij} = e_i - e_j = \sum_{k=i}^{j-1} f_k$ and $w_{ij} = \sum_{k=i}^{j-1} w_k \in C_{id} \cap V_0$. Then, for any collection of numbers $\{\beta_{ij}\}_{i<j}$,

$$\sum_{i<j} \beta_{ij} f_{ij} = \sum_{i<j} \beta_{ij} \sum_{k=i}^{j-1} f_k = \sum_{k=1}^{n-1} \alpha_k f_k$$

and

$$\sum_{i<j} \beta_{ij} w_{ij} = \sum_{k=1}^{n-1} \alpha_k w_k$$

with $\alpha_k = \sum_{i<j} \beta_{ij} 1\{i \le k < j\}$. We construct $w = \sum_{i<j} \beta_{ij} w_{ij} \in C_{id} \cap V_0$ with $s = Q_p w = \sum_{i<j} \beta_{ij} f_{ij} \in C_\pi$. Note that each $w_{ij} \in C_{id} \cap V_0$, so w will be as well whenever the β_{ij}'s are non-negative and not all zero, since $C_{id} \cap V_0$ is closed under conical combinations.

If $b = n$ then we can take $\beta_{kn} = n - \pi^{-1}(k) > 0$ for $k = 1, \dots, n-1$ and $\beta_{ij} = 0$ for $j \neq n$ as this gives the k^{th} place candidate $s_{\pi(k)} = n - \pi^{-1}(\pi(k)) = n - k > 0$ for $k = 1, \dots n-1$ and gives -$\binom{n}{2} < 0$ points to candidate n. Since all β_{ij} are non-negative, $w(\pi) = \sum_{i<j} \beta_{ij} w_{ij} \in C_{id} \cap V_0$

If $b \neq n$ let $\widetilde{\pi}$ be the permutation formed from π by moving n to last place (so $\widetilde{\pi}(i) = \pi(i)$ for $i < \pi^{-1}(n)$, $\widetilde{\pi} = \pi(i+1)$ for $\pi^{-1}(n) \le i < n$, and $\widetilde{\pi}(n) = n$), and let $\widetilde{w} = w(\widetilde{\pi})$, $\widetilde{s} = s(\widetilde{\pi})$ be constructed as above. Now, let $\widehat{w} = \widetilde{w} - \beta'_{bn} w_{bn}$ with $\beta'_{bn} = \binom{n}{2} + n - \pi^{-1}(n) + \frac{1}{2}$. Then $\widehat{s} = Q_p \widehat{w} = \widetilde{s} - \beta'_{bn} f_{bn}$ has all candidates ordered properly, although $\widehat{w}$ may not be in C_{id}. To fix this, we can add a large multiple of w_{ab} to $\widehat{w}$. Note that this will not impact the relative order of the coordinates in the corresponding results vector since it adds points to candidate a, subtracts points from candidate b, and leaves everyone else unchanged. Specifically, set $m = \min_{1\le k\le n-1}\left[(w_{ab})_k - (w_{ab})_{k+1}\right]$, $M = \max_{i\le k\le n}|\widehat{w}_k|$, and define $w = \widehat{w} + \frac{2M}{m}w_{ab}$. By construction, $w \in C_{id} \cap V_0$ and $s = Q_p w = \widehat{s} + M f_{ab} \in C_\pi$.

□

References

Daugherty, Z., Eustis, A. K., Minton, G., & Orrison, M. E. (2009). Voting, the symmetric group, and representation theory. *American Mathematical Monthly*, *116.8*, 667-687.

Sarri, D. (1992). Millions of election outcomes from a single profile. *Social Choice and Welfare*, *9*(4), 277-306.