Power Round Solutions

DMM 2022

1 Power Round

The theme is *Borda Score and Elections*. There are a total of 60 points for this round. Throughout the problem, ties are broken arbitrarily (you cannot break ties to your favor).

1.1 Borda Score in Single-Winner Elections

The Duke University Math Union (DUMU) is running an election for officers! There are three voters: Alice, Bob, and Cady, and three candidates: Xavier, Yisa, and Zack. We want to select a single winner. Each voter ranks the three candidates as follows:

$$\begin{aligned} \text{Alice} &: \text{Xavier} > \text{Yisa} > \text{Zack}, \\ \text{Bob} &: \text{Zack} > \text{Xavier} > \text{Yisa}, \\ \text{Cady} &: \text{Yisa} > \text{Xavier} > \text{Zack}, \end{aligned}$$

This means, for instance, Alice prefers Xavier the most and Zack the least. In this election, one might intuitively conclude that Xavier, who has the highest average rank, should win. The DUMU executive board wants to formalize this intuition, so they decide to select the candidate with the smallest *Borda score*.

The definition for Borda score is straightforward: the Borda score of a candidate c for a voter v is simply the rank of c in v's ranking, and the Borda score of c is simply her average rank. In this example, the Borda score of Zack for both Alice and Cady are 3, and 1 for Bob. Hence, the Borda score of Zack is $(1 + 3 + 3)/3 = \frac{7}{3}$.

Problem 1: (4 points total)

(a) (2 points) Similarly compute the Borda score for Xavier and Yisa, and explain why Xavier wins under Borda score.

Solution. Borda score for Xavier is $(1+2+2)/3 = \frac{5}{3}$, and the Borda score for Yisa is $(1+2+3)/3 = 2$. Xavier wins under Borda score because he has the smallest Borda score.

(b) (2 points) If we add one more voter, is it possible for Yisa to win? Prove your answer.

Solution. Yes. If the preference of the fourth voter is Yisa > Zack > Xavier, then Yisa wins.

Going beyond this example, we explore some properties of Borda score.

Problem 2: (6 points total)

(a) (2 points) If a candidate c is ranked first by more than half of the votes in an election, does c necessarily win under Borda score? Prove your answer.

Solution. No. Consider the following election with five voters v_1, v_2, v_3, v_4, v_5 and four candidates c_1, c_2, c_3, c_4, where the preferences of the voters for the candidates are

$$\begin{aligned} v_1 &: c_1 > c_2 > c_3 > c_4, \\ v_2 &: c_1 > c_2 > c_3 > c_4, \\ v_3 &: c_1 > c_2 > c_3 > c_4, \\ v_4 &: c_4 > c_2 > c_3 > c_1, \\ v_5 &: c_4 > c_2 > c_3 > c_1. \end{aligned}$$

Then, even though c_1 is ranked first by more than half of the votes, c_2 wins the election.

(b) (2 points) Suppose c wins under Borda score in an election. If we improve the position of c in some votes and leave everything else the same (*i.e.* if we exclude c, the rankings remain the same after the change), does c still win? Prove your answer.

Solution. Yes, because the Borda score of c strictly decreases while the Borda score of other candidates won't decrease. Thus, c still wins.

(c) (2 points) Suppose c wins under Borda score in an election. We then change votes in such a way that for each vote, if a candidate w was ranked below c originally, w is still ranked below c in the new vote. Does c still win under the new votes? Prove your answer.

Solution. No. Consider the same election as in (a). If we change the votes of v_4 and v_5 such that the preferences of both voters become $c_4 > c_2 > c_1 > c_3$, then c_1 would win the election.

1.2 Borda Score in Multi-Winner Elections

More generally, let $\mathcal{V}$ denote the set of voters and $\mathcal{C}$ denote the set of candidates. Suppose there are n voters and m candidates, *i.e.* $|\mathcal{V}| = n$ and $|\mathcal{C}| = m$. Let $r_v(c)$ denote the Borda score of candidate c for voter v.

In multi-winner elections, we select a set of candidates T, which we call a *committee*, instead of a single candidate. The Borda score of T for a voter v is $r_v(T) = \min_{c \in T} r_v(c)$, and the Borda score of T is $r_{\mathcal{V}}(T) = \frac{1}{n} \sum_{v \in V} r_v(T)$. To interpret this score, for each voter, we consider the candidate with the smallest Borda score; then, we take the sum of these scores, and average it over all voters.

Problem 3: (10 points total)

(a) (2 points) Consider the following election, where we have 5 voters $\mathcal{V} = \{v_1, \ldots, v_5\}$ and 5 candidates $\mathcal{C} = \{c_1, \ldots, c_5\}$, where the preferences of the voters for the candidates are

$$v_1 : c_1 > c_2 > c_3 > c_4 > c_5,$$
$$v_2 : c_2 > c_1 > c_4 > c_3 > c_5,$$
$$v_3 : c_5 > c_2 > c_1 > c_3 > c_4,$$
$$v_4 : c_3 > c_4 > c_2 > c_5 > c_1,$$
$$v_5 : c_4 > c_1 > c_2 > c_3 > c_5.$$

Find the committee of size 2 with the smallest Borda score, and compute its Borda score.

Solution. The committee with the smallest Borda score is $\{c_2, c_4\}$, whose Borda score is $(2 + 1 + 2 + 2 + 1)/5 = \frac{8}{5}$.

(b) (3 points) Given an election, let T_k^* denote the committee with the smallest Borda score of size k. Is it necessarily true that $T_k^* \subset T_{k+1}^*$? Prove your answer.

Solution. No. Consider the election given in the solution for Problem 2(a). The committee of size 1 with the smallest Borda score is $\{c_2\}$. However, the committee of size 2 with the smallest Borda score is $\{c_1, c_4\}$. Hence, we do not necessarily have $T_k^* \subset T_{k+1}^*$.

(c) (5 points) If we select k candidates uniformly at random from $\mathcal{V}$ to form a committee T, what is $\mathbf{E}[r_{\mathcal{V}}(T)]$, *i.e.* the expected value of the Borda score of T? Express your answer in terms of n, m, k, and prove your answer.

Solution. We show that $\mathbf{E}[r_{\mathcal{V}}(T)] = \frac{m+1}{k+1}$. Mark $m + 1$ points on a circle. Pick a subset of $k + 1$ points uniformly at random, and then choose one point P of these $k + 1$ as the cut-off point uniformly at random. Starting from P and going clockwise, mark the next point as the candidate with rank 1, and the point after that as the candidate with rank 2, and so on, until the last point which is marked as the candidate with rank m. The picked subset comprises P and a uniformly random size-k subset of $\mathcal{C}$. By symmetry, the expected clockwise distance going from the t^{th}-smallest ranked chosen candidate to the $(t + 1)^{\text{st}}$ is the same for every $t \in \{0, 1, \ldots, k\}$, if we view P as simultaneously the 0^{th} and the $(k + 1)^{\text{st}}$ smallest. Since these $k + 1$ distances sum to $m + 1$, all of them should be $\frac{m+1}{k+1}$. In particular, we have $\mathbf{E}[r_{\mathcal{V}}(T)] = \frac{m+1}{k+1}$.

1.3 Finding a Good Committee

In practice, we often find a good committee with the following procedure: pick candidates in k rounds, during which we build sets $\emptyset = T_0 \subsetneq T_1 \subsetneq \cdots \subsetneq T_k$, and declare T_k as the selected committee. In the j^{th} round, we pick candidate $c_j \in \mathcal{C} \setminus T_{j-1}$ that minimizes $r_{\mathcal{V}}(T_{j-1} \cup \{c_j\})$. In other words, we greedily pick the candidate that minimizes the Borda score in each round. We denote this procedure by Greedy.

In this section, we explore some properties of Greedy.

Problem 4: (10 points total)

(a) (2 points) For $k = 3$, compute the committee that GREEDY produces in the election given in Problem 3(a).

Solution. The committee that GREEDY produces is $\{c_1, c_2, c_4\}$.

(b) (3 points) Does GREEDY always produce the optimal committee, *i.e.* the committee with the smallest Borda score? Prove your answer.

Solution. No. Consider the election given in the solution for Problem 2(a). The committee of size 2 with the smallest Borda score is $\{c_1, c_4\}$. However, GREEDY produces $\{c_1, c_2\}$.

(c) (5 points) Recall that T_j is the committee produced by GREEDY after j rounds, $r_v(T_j)$ is the Borda score of T_j for voter v, and $r_{\mathcal{V}}(T_j)$ is the Borda score of T_j. Prove that

$$r_{\mathcal{V}}(T_j) - r_{\mathcal{V}}(T_{j+1}) \geq \frac{\sum_{v \in \mathcal{V}} r_v(T_j)(r_v(T_j) - 1)}{2n(m-j)}.$$

Solution. For a candidate $c \notin T_j$, define $\Delta_c := r_{\mathcal{V}}(T_j) - r_{\mathcal{V}}(T_j \cup \{c\})$, *i.e.* the current marginal contribution of c to the 1-Borda score. Taking the sum of Δ_c over $c \notin T_j$:

$$\sum_{c \in \mathcal{C} \setminus T_j} \Delta_c = \frac{1}{n} \sum_{v \in \mathcal{V}} \sum_{j=1}^{r_v(T_j)-1} j = \frac{\sum_{v \in \mathcal{V}} r_v(T_j)(r_v(T_j) - 1)}{2n}.$$

GREEDY chooses $c^* = \arg\max_c \Delta_c$ at the $(j+1)^{\text{st}}$ iteration, giving us

$$r_{\mathcal{V}}(T_j) - r_{\mathcal{V}}(T_{j+1}) = \Delta_{c^*} \geq \frac{1}{m-j} \sum_{c \in \mathcal{C} \setminus T_j} \Delta_c = \frac{\sum_{v \in \mathcal{V}} r_v(T_j)(r_v(T_j) - 1)}{2n(m-j)}.$$

In the following, we investigate theoretical guarantees on the quality of the committee produced by GREEDY. You can use the conclusion from Problem 4(c) even if you haven't solved it. Complete proofs to the Problem 5 can be hard, and partial credits will be offered to useful observations and reasonable attempts. Write down whatever you think can take you closer to the solution!

Let RAND(k) denote the answer of Problem 3(c), *i.e.* the expected Borda score of a randomly selected committee of size k. Recall that T_k is the committee produced by GREEDY after k rounds.

Problem 5: (20 points total)

(a) (15 points) Show that, for any election, we have

$$r_{\mathcal{V}}(T_k) \leq 2 \cdot \text{RAND}(k).$$

Solution. Since $\textsc{Rand}(k) = \frac{m+1}{k+1}$, we actually want to prove $r_{\mathcal{V}}(T_k) \le 2 \cdot \frac{m+1}{k+1}$. We prove by induction. The base case clearly holds. Now suppose that the claim holds for some $k-1$ and we will prove that it also holds for k. By induction hypothesis, we have:

$$r_{\mathcal{V}}(T_{k-1}) \le 2 \cdot \frac{m+1}{k}.$$

If $r_{\mathcal{V}}(T_{k-1}) \le 2 \cdot \frac{m+1}{k+1}$, then $r_{\mathcal{V}}(T_k) \le r_{\mathcal{V}}(T_{k-1}) \le 2 \cdot \frac{m+1}{k+1}$ finishes the proof. Thus, we only need to consider the following case:

$$2 \cdot \frac{m+1}{k+1} < r_{\mathcal{V}}(T_{k-1}) \le 2 \cdot \frac{m+1}{k}.$$

We now have the following, where the first inequality is by the conclusion from Problem 4(c) and second by Cauchy-Schwarz inequality:

$$\begin{aligned} r_{\mathcal{V}}(T_{k-1}) - r_{\mathcal{V}}(T_k) &\ge \frac{\sum_{v\in\mathcal{V}} r_v(T_{k-1})(r_v(T_{k-1})-1)}{2n(m-k+1)} \\ &\ge \frac{\frac{1}{n}(\sum_{v\in\mathcal{V}} r_v(T_{k-1}))^2 - \sum_{v\in\mathcal{V}} r_v(T_{k-1})}{2n(m-k+1)} \\ &= \frac{(\sum_{v\in\mathcal{V}} r_v(T_{k-1}))^2}{2n^2(m+1)} \cdot \frac{m+1}{m-k+1} \cdot \frac{\sum_{v\in\mathcal{V}}(r_v(T_{k-1})-1)}{\sum_{v\in\mathcal{V}} r_v(T_{k-1})}. \end{aligned}$$

Since $r_{\mathcal{V}}(T_{k-1}) \ge 2 \cdot \frac{m+1}{k+1}$ by assumption, we have:

$$\begin{aligned} \frac{m+1}{m-k+1} \cdot \frac{\sum_{v\in\mathcal{V}}(r_v(T_{k-1})-1)}{\sum_{v\in\mathcal{V}} r_v(T_{k-1})} &\ge \frac{m+1}{m-k+1} \cdot \frac{2\cdot\frac{m+1}{k+1} - 1}{2\cdot\frac{m+1}{k+1}} \\ &= \frac{2(m+1)-k-1}{2(m+1)-2k} \ge 1. \end{aligned}$$

Combining the previous two inequalities, we therefore have:

$$r_{\mathcal{V}}(T_{k-1}) - r_{\mathcal{V}}(T_k) \ge \frac{(\sum_{v\in\mathcal{V}} r_v(T_{k-1}))^2}{2n^2(m+1)} = \frac{r_{\mathcal{V}}^2(T_{k-1})}{2(m+1)},$$

which is equivalent to:

$$r_{\mathcal{V}}(T_k) \le -\frac{1}{2(m+1)} r_{\mathcal{V}}^2(T_{k-1}) + r_{\mathcal{V}}(T_{k-1}).$$

Notice that the right hand side is a quadratic function in $r_{\mathcal{V}}(T_{k-1})$, which is monotonically increasing for $r_{\mathcal{V}}(T_{k-1}) \le m+1$. Since $r_{\mathcal{V}}(T_{k-1}) \le 2 \cdot \frac{m+1}{k} \le m+1$, the right hand side reaches its maximum at $2 \cdot \frac{m+1}{k}$. Thus, we have:

$$r_{\mathcal{V}}(T_k) \le -\frac{1}{2(m+1)} \cdot \left(\frac{2(m+1)}{k}\right)^2 + \frac{2(m+1)}{k} \le \frac{2(m+1)}{k+1},$$

which concludes our induction.

(b) (5 points) Show that there exists an instance such that

$$r_{\mathcal{V}}(T_k) > \textsc{Rand}(k).$$

(Hint: You want to show that with appropriate choice of n, m, k, and rankings of voters for candidates, this inequality is possible.)

Solution. Take m to be a power of 2 that is large enough, $n = (m-1)!$, and $k = 2$. Suppose the candidates are $\{c_1, \ldots, c_m\}$. c_1 ranks at the $(\frac{m}{2})^{\text{th}}$ place for every voter, and any pair of voters have different preferences on other candidates. Greedy chooses candidate 1 in the first round, and without loss of generality, we assume that Greedy chooses candidate 2 in the second round; *i.e.* $T_k = \{c_1, c_2\}$. The Borda score of this committee is $\frac{3}{8}m$, while $\textsc{Rand}(k) = \frac{1}{3}(m+1)$. Since m is large enough, we have $r_{\mathcal{V}}(T_k) > \textsc{Rand}(k)$. in this instance.

1.4 Generalization: s-Borda Score

One commonly used generalization of Borda Score is *s-Borda score*. In this section, we use $r_{\mathcal{V}}(T)$ to denote the s-Borda score of T instead of the usual Borda score, which is defined by

$$r_{\mathcal{V}}(T) = \frac{1}{n} \sum_{v \in \mathcal{V}} \left(\min_{Q \subseteq T, |Q| = s} \sum_{c \in Q} r_v(c) \right).$$

Here, $r_v(c)$ still denotes the Borda score of c for v, which is the rank of c in v's ranking.

Problem 6: (10 points total)

(a) (3 points) Interpret this definition in plain English.

Solution. For each voter, consider the s candidates in T whose Borda score is the smallest. Now, take the sum of these scores, and average it over all the voters.

(b) (2 points) For $s = 2$, compute the committee of size 3 with the smallest s-Borda score in the election given in Problem 3(a).

Solution. The committee with the smallest s-Borda score is $\{c_1, c_2, c_4\}$.

(c) (5 points) If we select k candidates uniformly at random from $\mathcal{V}$ to form a committee T, what is $\mathbf{E}[r_{\mathcal{V}}(T)]$, *i.e.* the expected value of the s-Borda score of T? Prove your answer.

Solution. The proof for Problem 3(c) actually shows that $\mathbf{E}[r_{\mathcal{V}}(T)] = \frac{s(s+1)}{2} \cdot \frac{m+1}{k+1}$.