A Recommendation Game

Isaac Swift

Hong Kong Baptist University

6 December 2021

Overview

Background

- The Secretary Problem
- The Marriage Problem
- The Sultan's Dowry Problem

A Recommendation Game

- Setup
- Results
- Zero-Sum Game
- All-or-Nothing Game
- I'm Feeling Lucky Game

Setup

You would like to hire a new secretary

- There are N candidates

Setup

You would like to hire a new secretary

- There are N candidates
- You interview the candidates sequentially

Setup

You would like to hire a new secretary

- There are N candidates
- You interview the candidates sequentially
- After each interview, you choose to accept or keep searching

Setup

You would like to hire a new secretary

- There are N candidates
- You interview the candidates sequentially
- After each interview, you choose to accept or keep searching
- You can't observe quality of a candidate, only the ordinal ranking among those seen so far

Setup

You would like to hire a new secretary

- There are N candidates
- You interview the candidates sequentially
- After each interview, you choose to accept or keep searching
- You can't observe quality of a candidate, only the ordinal ranking among those seen so far
- The goal is to hire the best candidate

Literature

Who solved the Secretary Problem?

- Genesis: Gardner (1960), Bissinger and Siegel (1963), Cayley (1875), Kepler (1613)
- Solution: Lindley (1961), Chow, Moriguti, Robbins, and Samuels (1964), Gilbert and Mosteller (1966)
- Extensions: Infinite number of possible citations even through 2021
    - Unknown N
    - Recall
    - Multiple selections
    - Different objective
    - Multiple searchers competing
- Literature Reviews: Freeman (1983), Ferguson (1989)

Solution

Lemma

The optimal strategy is a cutoff rule.

Cutoff Rule

- You reject the first k candidates automatically.
- You accept any candidate that is the best so far

Solution

What is your probability of getting the best candidate?

- Each candidate has a $\frac{1}{N}$ chance of being the best
- If the you observe the best candidate after k, you will accept
- You will get to candidate $n > k$ iff the best of the first $n - 1$ candidates was in position k or earlier

$$prob = \sum_{n=k+1}^{N} \frac{k}{n-1} \frac{1}{N} \tag{1}$$

Solution

Simple find the value of k that maximizes the probability

$$prob = \frac{1}{N}\sum_{n=k+1}^{N}\frac{k}{n-1} = \frac{k}{N}\sum_{n=k+1}^{N}\frac{N}{n-1}\frac{1}{N} \qquad (2)$$

Call $x = \frac{k}{N}$ the fraction of candidates automatically rejected. As N gets large, this probability approaches a Riemann integral.

$$prob \to x\int_{x}^{1}\frac{1}{t}\,dt \qquad (3)$$

Solution

We can evaluate the integral and maximize.

$$x \int_x^1 \frac{1}{t}\, dt = -x \log(x) \tag{4}$$

$$x^* = \frac{1}{e} = .367879 \tag{5}$$

The simple optimal rule is to reject the first 37 percent of candidates, then accept any that is the best so far.

This rule finds the best candidate about 37 percent of the time.

Example

If $N = 3$, the optimal strategy is to reject the first candidate then accept the second if their better than the first, otherwise you get the third candidate.

Possible orderings

1 2 3	1 3 2
2 1 3	2 3 1
3 1 2	3 2 1

This strategy picks the best candidate 50 percent of the time.

Other Secretary Problem

One other version of the problem needs discussing.

- You don't care solely about the top candidate
- Each candidate has a value v_n drawn from some known distribution F
- You still only observe ordinal rankings
- You maximize the expected value of your choice

Solution

Despite a more complicated object, this problem is actually simpler

- Can be solved with dynamic programming
- Working from the end, you can find your expected value of continuing
- You can compute the expected value of the current option given the information you have
- You accept if the expected value is better than the continuation value

Example: $F = U[0,1]$

The expected value of your current option is simple to compute

$$E[v_n|m] = \frac{n-m+1}{n+1} \tag{6}$$

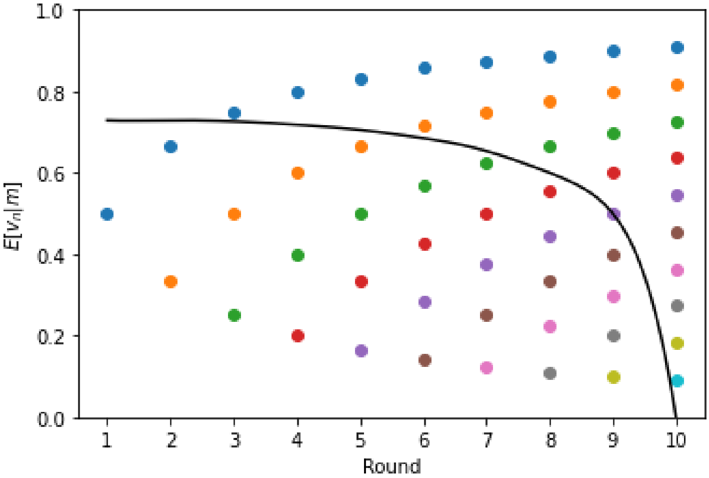
if you are in round n and the current option is the m^{th} best one you've seen.

The dynamic programming is easy now

- In the last round, you are stuck with the candidate no matter what
- $E[v_N] = \frac{1}{2}$
- In the second to last round, you will then accept if they are better than average
- Conditional on being above average, their expected value is $\frac{3}{4}$
- This gives $E[v_{N-1}] = \frac{1}{2}\frac{1}{2} + \frac{1}{2}\frac{3}{4} = \frac{5}{8}$

Solution

You can continue working backward through all N rounds.



Let's make it a game

Player 1: Recommender (R)

- Observes N options
- Knows own preference and DM's preference over options
- Observes only ordinal rank (not actually important)
- Chooses an order for the options

Player 2: Decision Maker (DM)

- Plays the Secretary Problem as before
- Knows that the order is chosen by DM

Motivation

Recommender Systems

- Amazon, Netflix, etc.
- Chooses the order to present recommendations
- Wants to maximize profits while customer wants a good deal

Bargaining

- One party is making offers the second party can accept or reject
- Should with your best offer or start with a low-ball and work your way up?

Anything where you choose an ordering

- Location problems
- Giving a good presentation

Decision Maker Notation (if interested)

- r_i^n is the ranking of option i among the n options seen so far
- $r_1^1 = 1$, r_1^2 and $r_2^2 \in \{1, 2\}$, and so forth
- A partial history is $h^n = (r_1^n r_2^n \ldots r_n^n)$
- A total history, $h \in \mathcal{H}$, is one $N!$ permutations of $\{1, 2, \ldots, N\}$
- DM's strategy is to accept or reject at each stage, $\sigma : h^n \to \{0, 1\}$ with $\sum_{n=1}^{N} \sigma(h^n) = 1$ for all h
- $\mu(h)$ is the probability of each ordering h
  - In one-player problem $\mu(h) = \frac{1}{N!}$ for all h
  - Now $\mu(h)$ is chosen by the other player
- v_i is the value of choosing the i^{th} best option
- DM their maximizes expected payoff

$$\max_{\sigma} \sum_{h \in \mathcal{H}} \sum_{n=1}^{N} \mu(h)\sigma(h^n) v_{r_n^N} \tag{7}$$

Recommender Notation (if interested)

- Sees both players' preference rankings for each option
  - This can be represented as a permutation, $p \in \mathcal{H}$
  - $p = (p_1 p_2 \ldots p_N)$, with p_i being the preference ranking for DM of R's i^{th} best choice
  - $p_1 = 3$ means R's number 1 choice is DM's number 3 choice
  - $f(p)$ is the distribution over these permutations
  - $f(p) = \frac{1}{N!}$ represents R and DM having independent preferences
- Chooses the order in which the options will be presented, h
- w_i is R's payoff from their i^{th} best option being chosen
- Maximizes the expected payoff

$$\max_{h(p)\in\mathcal{H}} \sum_{p\in\mathcal{H}} \sum_{n=1}^{N} \sigma(h^n) f(p) w_{p_{h_n}} \tag{8}$$

Bound equilibrium payoffs

Lemma

Against any strategy of R, DM's best response gets a payoff of at least $\bar{v} = \frac{1}{N}\sum_{i=1}^{N} v_i$.

Lemma

Against any pure strategy of DM, R's best response gets a payoff of w_1.

Proof Sketch

The first lemma is pretty simple to prove.

- Consider randomly choosing a number from 1 to N with equal probability, then selecting option N.
- This gives DM a payoff of $\bar{v}$ regardless of R's strategy

It can be done with a pure strategy too

- Given R's strategy, there must be some round that has expected payoff of at least $\bar{v}$
- DM can just take the option in that round ignoring the history

Proof Sketch

The second lemma is more complicated.

- For any selection strategy of DM, there is an order where DM gets their worst option
  - If DM's strategy ever selects one that is the worst so far, replace that option with the worst overall
  - If DM's strategy never selects one that is the worst so far, place the options in decreasing order
- For any selection strategy of DM, there is an order where DM gets their second worst option
- For any selection strategy of DM, there is an order where DM gets any particular option
- R can always force their top choice to be chosen.

Correlation

Call ρ the expected value to DM of R's top choice

$$\rho = \sum_{p \in \mathcal{H}} v_{p_1} f(p) \tag{9}$$

ρ captures the relevant notion of correlation in this game.

- Positive correlation: $\rho > \bar{v}$
- No correlation: $\rho = \bar{v}$
- Negative correlation: $\rho < \bar{v}$

Equilibrium Payoffs

Corollary

Any pure strategy equilibrium gives R and DM payoffs of (w_1, ρ).

Corollary

If R and DM have negatively correlated preferences, no pure strategy equilibrium exists.

Simplification

Consider a zero-sum version of the game

- The two players have opposite preferences
  - $f(p)$ is degenerate on $p = (N, N-1, N-2, \ldots, 2, 1)$
- The payoffs are equal, $w_i = v_{N-i+1}$

Proposition

The unique equilibrium payoffs are $(-\bar{v}, \bar{v})$.

Proof

We already have $\bar{v}$ as a lower bound on DM's payoff.

- We found a strategy for DM that guarantees $\bar{v}$ regardless of R's strategy
- Randomly accepting on of the options

We can now establish $\bar{v}$ as an upper bound on DM's payoff

- We will find a strategy for R that guarantees a payoff of $-\bar{v}$ regardless of DM's strategy
- Maximal Risk Ordering

The two strategies will be best responses to each other

Maximal Risk Ordering

In each round, randomize between presenting the best and the worst of the remaining options. Play the the best option with probability q.

$$q = \frac{\tilde{v} - v_{min}}{v_{max} - v_{min}} \tag{10}$$

where v_{max} and v_{min} are the best and worst options remaining and $\tilde{v}$ is the average of remaining options.

- In the first round v_{max}, v_{min}, and $\tilde{v}$ are v_1, v_N, and $\bar{v}$
- The expected value of accepting is $v_{max}q + v_{min}(1 - q) = \tilde{v}$
- If the v_i's are evenly spaced, q is always $\frac{1}{2}$

Maximal Risk Ordering

DM can do no better against this ordering than $\bar{v}$ in expectation.

- All remaining options are indistinguishable
- Proof by induction
- DM is always indifferent between accepting and rejecting
- Any strategy of DM give payoff of $\bar{v}$

Start with a low-ball offer and work your way up? Start with your best offer right out of the gate? A mix of the two.

Setup

Consider an all-or-nothing example

- $v_1 = w_1 = 1$ and $v_i = w_i = 0$ for $i \neq 1$
- Each player has a top choice (not necessarily the same)
- Each player is maximizing the probability of getting their top choice

R can use a modified form of the previous strategy because there is so much indifference

- Each round, DM's top choice is presented with probability $q = \frac{1}{N-n+1}$
- Otherwise one of the other options is presented, but maybe not the worst one

Solution

R can us a cyclic Latin square to construct their strategy.

5	4	3	2	1
4	3	2	1	5
3	2	1	5	4
2	1	5	4	3
1	5	4	3	2

Randomizing between the rows preserves information in a way similar to the Maximal Risk Ordering.

Solution

Suppose that R's top is just as likely to be DM's number choice as any other number.

$$\sum_{p \in \mathcal{H}} f(p)\chi_{\{p_1=1\}} \geq \sum_{p \in \mathcal{H}} f(p)\chi_{\{p_1=i\}} \qquad \forall i \in \{2, 3, \ldots .N\} \qquad (11)$$

Proposition

The following strategies constitute an equilibrium.

- *R: Selects the row of the cyclic Latin square that starts with their own most preferred option*
- *DM: Accepts the first option*

R makes progressively better and better offers.

Mixed Strategy Equilibrium ($\rho < \frac{1}{N}$)

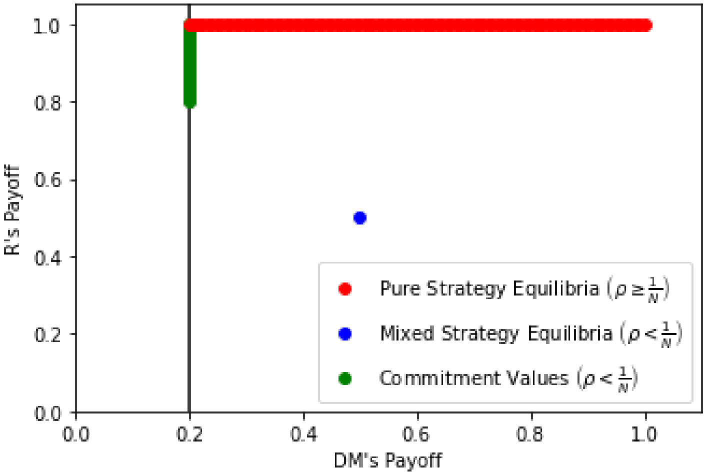
Suppose there is negative correlation

- R plays a random row of cyclic Latin square
- DM picks a random option
- Both players get $\frac{1}{N}$

Both players would be better off if there were fewer options

- They can just pretend there are only two options
- DM randomizes over choosing either of the first two options
- R randomizing putting R's top choice or DM's top choice first and the other second
- No equilibrium can give R better than $\frac{1}{2}$

Equilibrium Payoffs



DM does worse with a small positive correlation than with a negative correlation.

Google


Google
Google Search
I'm Feeling Lucky
Google offered in: 日本語 繁體中文 中文(简体)

Lucky

Let R choose one option to be the "Lucky" option.

- DM can take the "Lucky" option without seeing it
- DM can play the game as usual

Suppose $\rho \geq \bar{v}$

- R will always put their own most preferred option in the "Lucky" spot
- R puts all option in the Maximal Risk Order otherwise
- DM will always take the "Lucky" option

So Lucky

Suppose now that $\rho < \bar{v}$

- Suppose R is able to commit to their recommendation strategy
- Without loss, your can consider only equilibria where DM takes the "Lucky" option (like the Revelation Principle)
- DM's payoff is equal to their worst equilibrium payoff without the "Lucky" button
- R's payoff is weakly higher than their best equilibrium payoff without the "Lucky" button

Conclusion

Thank you for your time.