How to derive the axis of candidates from approval ballots?

Théo Delemazure

Miles Seminar - July 6th 2023

Who am I?

- Second year PhD student under the supervision of **Jerome Lang** and **Dominik Peters**.
- **Formation:** ENS (2017-2021) and Master IASD (2019-2020).

Who am I?

- Second year PhD student under the supervision of **Jerome Lang** and **Dominik Peters**.
- **Formation:** ENS (2017-2021) and Master IASD (2019-2020).
- *Disclaimer:* Have not been doing much machine learning since then.

What do I do?

I chose to focus on **Computational Social Choice** (COMSOC).
And mostly **voting theory**.

My tools:

- **Axiomatic analysis:** does this voting rule satisfies this particular property?
- **Computational complexity:** how hard is it to compute the results of this problem? How hard is it to manipulate?
- **Data simulation:** if I generate voting data with some model, which rule performs the best for some metrics?
- **Data analysis:** what would be the results of this rule on this real dataset?

What do I do?

It would be interesting to mix this with **machine learning ideas:**

- Rules that use **machine learning techniques** to aggregate preferences, or for other social choice problems (e.g. matching).
- Using ML to **evaluate** the rules.
- Learning to vote: **bandit/reinforcement learning** to simulate behavior of voters.
- Using preference aggregation knowledge for **classifiers aggregation/ensemble learning**.

What do I do?

It would be interesting to mix this with **machine learning ideas:**

- Rules that use **machine learning techniques** to aggregate preferences, or for other social choice problems (e.g. matching).
- Using ML to **evaluate** the rules.
- Learning to vote: **bandit/reinforcement learning** to simulate behavior of voters.
- Using preference aggregation knowledge for **classifiers aggregation/ensemble learning**.

Let's see this on an example!

Voter Autrement

Website of the experiment *Voter Autrement 2022*

Voter Autrement

Prequel of this presentation

Approval ballot

Candidate 1	☑
Candidate 2	☐
Candidate 3	☑
Candidate 4	☑
Candidate 5	☐
Candidate 6	☐

Introducing Isabelle Lebon

Ce que le vote par approbation révèle des préférences des électeurs français by *Isabelle Lebon, Antoinette Baujard, Frédéric Gavrel, Herrade Igersheim, Jean-François Laslier*

They use the approval ballots of the experiment in 2012 to compute the most likely left-right axis and how much it fits the data.

What is the best way to obtain a left-right axis of the candidates from the approval ballots?

How to qualitatively evaluate if a set of candidates can be represented by an axis?

Example

Let's say we have the following profile:

	Sarkozy	Hollande	Joly	Melenchon	Le Pen	Bayrou
v_1	✓					✓
v_2	✓				✓	
v_3		✓	✓	✓		
v_4	✓	✓				✓
v_5		✓	✓			

Example

Let's say we have the following profile:

	Melenchon	Joly	Hollande	Bayrou	Sarkozy	Le Pen
v_1				✓	✓	
v_2					✓	✓
v_3	✓	✓	✓			
v_4			✓	✓	✓	
v_5		✓	✓			

The profile is **linear**.

Example

Let's say we have the following profile:

	Melenchon	Joly	Hollande	Bayrou	Sarkozy	Le Pen
v_1				✓	✓	
v_2					✓	✓
v_3	✓		✓			
v_4			✓		✓	
v_5		✓	✓			

The profile is **not linear anymore**.

Some rules

We select the axis that minimize some distance:

- **Voter Deletion (VD)**: how many voters do we need to delete to make the profile linear with the axis?

Example

	Melenchon	Joly	Hollande	Bayrou	Sarkozy	Le Pen
v_1				✓	✓	
v_2					✓	✓
v_3	✓		✓			
v_4			✓		✓	
v_5		✓	✓			

The profile is **not linear anymore**.

Example

	Melenchon	Joly	Hollande	Bayrou	Sarkozy	Le Pen
v_1				✓	✓	
v_2					✓	✓
v_5		✓	✓			

The profile is **linear!**

Some rules

We select the axis that minimize some distance:

- **Voter Deletion (VD)**: how many voter do we need to delete to make the profile linear with the axis?
- **Ballot Completion (BC)**: how many candidates do we need to add to approval ballots to make the profile linear with the axis?

Example

	Melenchon	Joly	Hollande	Bayrou	Sarkozy	Le Pen
v_1				✓	✓	
v_2					✓	✓
v_3	✓		✓			
v_4			✓		✓	
v_5		✓	✓			

We need to add **2 candidates** to the approval ballots.

Some rules

We select the axis that minimize some distance:

- **Voter Deletion (VD)**: how many voter do we need to delete to make the profile linear with the axis?
- **Ballot Completion (BC)**: how many candidates do we need to add to approval ballots to make the profile linear with the axis?
- **Minimal Flips (MF)**, **Minimal Swaps (MS)** and **Forbidden Triplets (FT)**.

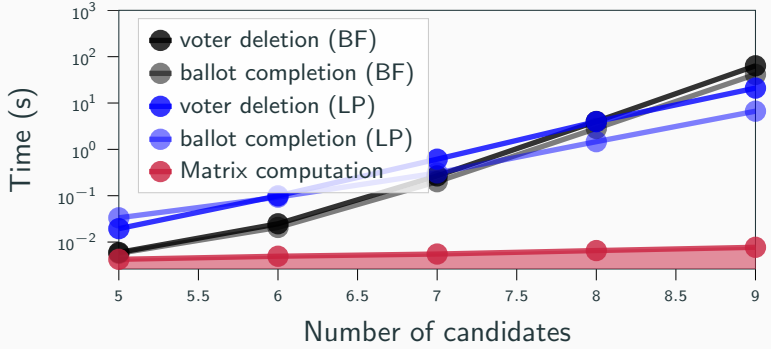
Which one to chose?

1. **Computational complexity:** how hard is it to compute the best axis?
2. **Data analysis on real dataset:** what do we observe on real data?
3. **Experiments and simulations:** which rule perform the best on simulated data?
4. **Axiomatic analysis:** what are the theorethical properties of the rules?

1. Computational complexity

- Direct reductions to already known **NP-Hard** problems.
- **Brute-force algorithm**: how to optimize it? (1 month $\rightarrow$ 1 hour).
- **Integer Linear Program**

Time to compute the axis for different number of candidates (n=1000)



2. Analysis on real datasets

French presidential elections from 2002 to 2022 , and other political elections (below: *France 2017*, $\sim 10,000$ voters).

Voter Deletion: FA $\prec$ MLP $\prec$ NDA$\prec$ FF $\prec$ EM $\prec$ BH $\prec$ JLM $\prec$ PP $\prec$ NA $\prec$ JL $\prec$ JC

Ballot Completion: MLP $\prec$ NDA $\prec$ FF $\prec$ JL $\prec$ EM $\prec$ BH $\prec$ JLM $\prec$ PP $\prec$ NA $\prec$ FA $\prec$ JC

73% of votes are intervals (69% if we exclude votes with 1 candidate), and ~ 0.7 candidates to add per voter for *Ballot Completion* (excluding votes with 1 candidate).

2. Analysis on real datasets

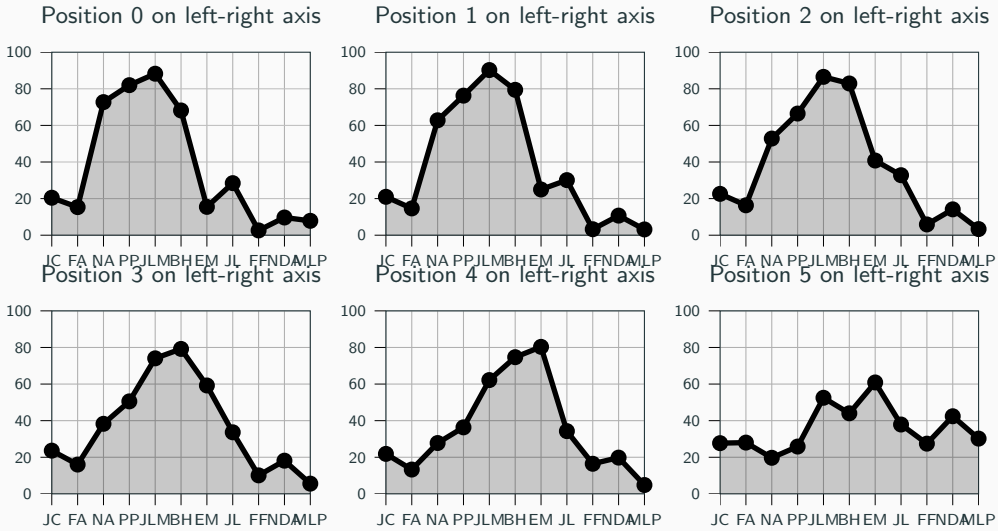
French presidential elections from 2002 to 2022 , and other political elections (below: *France 2017*, $\sim 10,000$ voters).

Voter Deletion: FA $\prec$ MLP $\prec$ NDA$\prec$ FF $\prec$ EM $\prec$ BH $\prec$ JLM $\prec$ PP $\prec$ NA $\prec$ JL $\prec$ JC

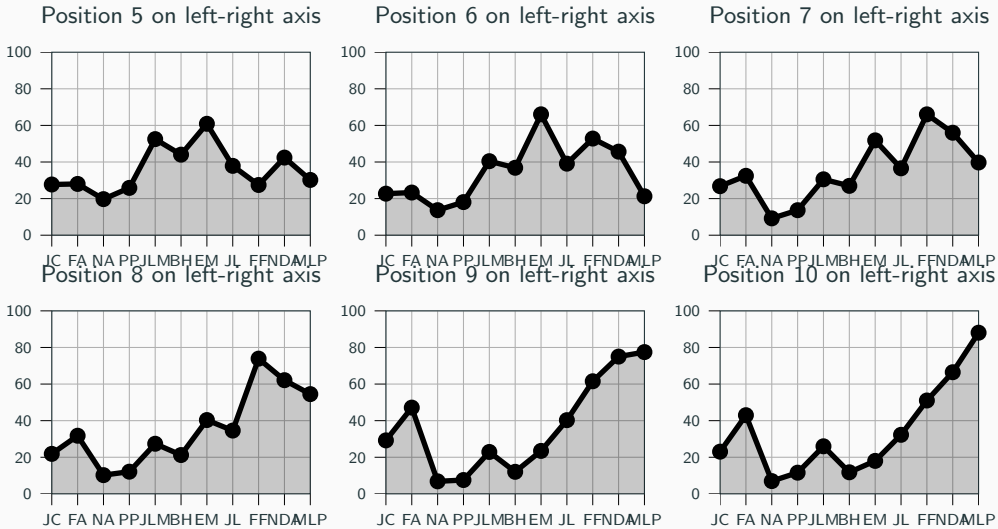
Ballot Completion: MLP $\prec$ NDA $\prec$ FF $\prec$ JL $\prec$ EM $\prec$ BH $\prec$ JLM $\prec$ PP $\prec$ NA $\prec$ FA $\prec$ JC

Small candidates seem pushed towards the extremes.

2. Analysis on real datasets

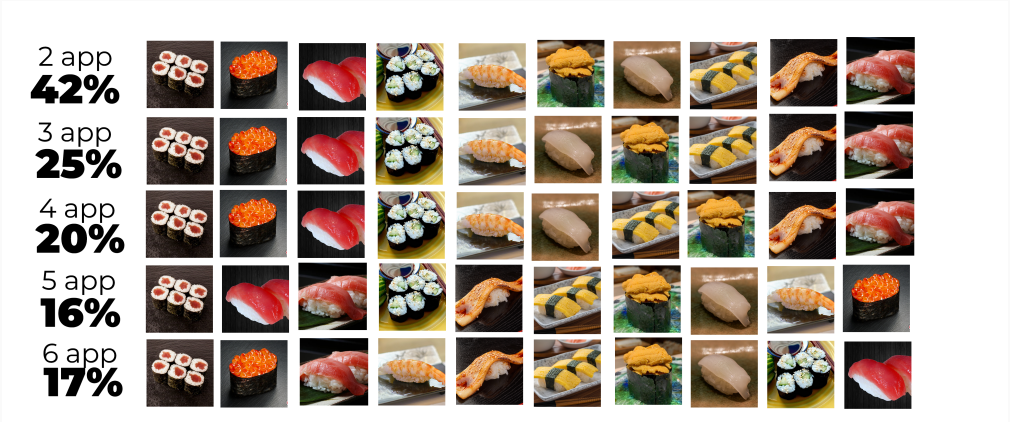


2. Analysis on real datasets

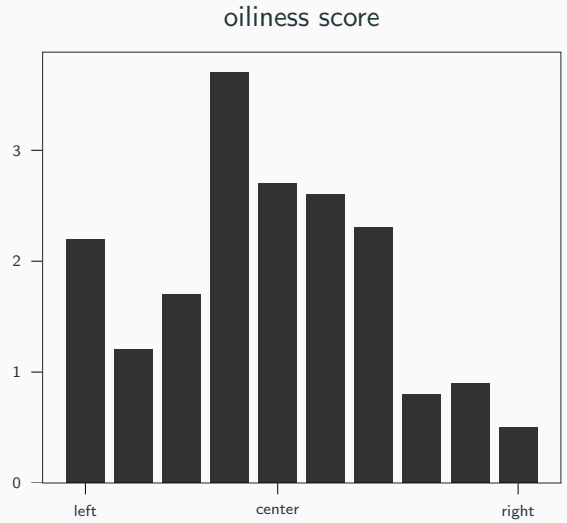


2. Analysis on real datasets

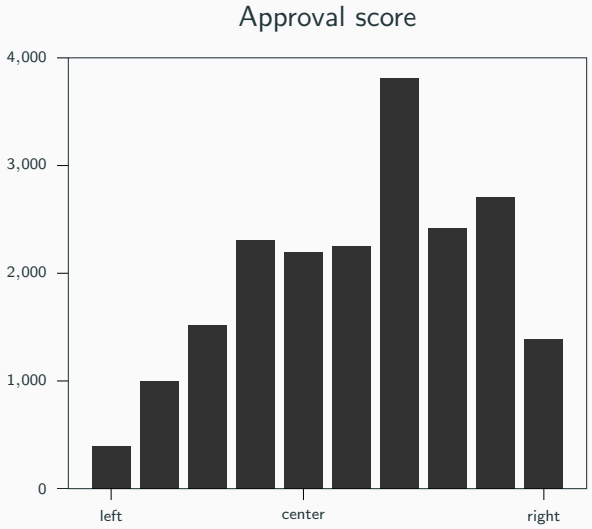
Sushi dataset of rankings, we can vary the number of approved candidates.



2. Analysis on real datasets



2. Analysis on real datasets



3. Simulations on synthetic data

- We need a model to **generate synthetic approval data** with an underlying axis.
- We also want to have *"big"* and *"small"* candidates.
- **Proposal:**
  1. Voters and candidates have a positions $x \in [0, 1]$ and candidates also have a *fame score* $s_c \in [0, 1]$.
  2. Voters have a higher chance to approve candidates that are close to them, and that are famous.

Experiments

- Which rule find the correct axis most of the time? How does it depends on the parameters of the model?
- Which rule is the less sensitive to the eccentricity issue?
- Which rule is the most robust to slight variations of the profile?
- How do the rules compare to MLE?

Which properties are satisfied by our rules?

Which properties are satisfied by our rules?

- **Linear-consistency:** if a profile is a linear (i.e. all votes are interval of some axis), the output should be all the axis consistent with the profile.

Which properties are satisfied by our rules?

- **Linear-consistency:** if a profile is a linear (i.e. all votes are interval of some axis), the output should be all the axis consistent with the profile.
- **Anonymity:** if we permute the voters, the result should be the same

Which properties are satisfied by our rules?

- **Linear-consistency:** if a profile is a linear (i.e. all votes are interval of some axis), the output should be all the axis consistent with the profile.
- **Anonymity:** if we permute the voters, the result should be the same
- **Reinforcement:** if two profiles result in the same axis, the union should also results in this axis.

Which properties are satisfied by our rules?

- **Linear-consistency:** if a profile is a linear (i.e. all votes are interval of some axis), the output should be all the axis consistent with the profile.
- **Anonymity:** if we permute the voters, the result should be the same
- **Reinforcement:** if two profiles result in the same axis, the union should also results in this axis.
- **Neutrality, Continuity, Clone-proofness, Heredity, Stability,...**

What can we do with this?

What can we do with this?

- **Characterize a family of rule:** $f(A, \prec) = \sum_{v \in V} \texttt{score}(A_v, \prec)$ (*Anonymity* + *Reinforcement* + *Continuity*).

What can we do with this?

- **Characterize a family of rule:** $f(A, \prec) = \sum_{v \in V} \texttt{score}(A_v, \prec)$ (*Anonymity* + *Reinforcement* + *Continuity*).
- + *Stability*: $\texttt{score}(A_v, \prec) \in \{0, 1\}$.

What can we do with this?

- **Characterize a family of rule:** $f(A, \prec) = \sum_{v \in V} \texttt{score}(A_v, \prec)$ (*Anonymity* + *Reinforcement* + *Continuity*).
- + *Stability*: $\texttt{score}(A_v, \prec) \in \{0, 1\}$.
- **Characterize specific rules** (Voter Deletion: *Anonymity* + *Neutrality* + *Reinforcement* + *Continuity* + *Stability* + *Linear-consistency*).

What can we do with this?

- **Characterize a family of rule:** $f(A, \prec) = \sum_{v \in V} \texttt{score}(A_v, \prec)$ (*Anonymity* + *Reinforcement* + *Continuity*).
- + *Stability*: $\texttt{score}(A_v, \prec) \in \{0, 1\}$.
- **Characterize specific rules** (Voter Deletion: *Anonymity* + *Neutrality* + *Reinforcement* + *Continuity* + *Stability* + *Linear-consistency*).
- **Highlights impossibilities**: *Linear-consistency* + *Indifference to unknown candidates*.

- **Linear-consistency:** if a profile is a linear (i.e. all votes are interval of some axis), the output should be all the axis consistent with the profile.
- **Indifference to unknown candidates:** if nobody approve a candidate, its position is interchangeable in the output axis.

Example: 3 candidates $\{a, b, c\}$ and 1 voter $\{a, b\}$. c is never approved but there is no reason to put it between a and b.

There are two ways out of this issue:

1. **A feature, not a bug:** if nobody like some candidate, it makes sense that it is put at an extremity.
2. **Finer model:** instead of a simple ordering, we could have something like a *fuzzy relation*, or only output a subset of the candidates. We could also output positions on a metric space.

From orders to metric space

What if, instead of a an axis, we want to output more information, for instance **the positions of all candidates on a 1D (or 2D) metric space?**

$\Rightarrow$ Dimension reduction.

I tried with the dimension reduction algorithms I know: *PCA, TSNE, Isomap, MDS*.
Results are bad: the resulting axis do not make any sense, and some of them are not deterministic.

Conclusion

- Can we use **machine learning techniques** to find interesting way to solve this problem?
- **Can this problem be useful** in machine learning ? (applied to a group of classifiers for instance?)

Thanks for your attention!