Representative Democracy Dream: the Missing Link

PhD Brown Bag - HEC Paris

Mateus Hiro Nagata

June 20, 2024

Motivation

Next Steps

Pitch

"The moment a people allows itself to be represented, it is no longer free: it no longer exists. The day you elect representatives is the day you lose your freedom."

– Jean-Jacques Rousseau

Direct (Pure) Democracy is considered to be the ideal "rule of the people".

Representative Dream Problem

Looking at the past (in-sample data), can we predict the decision of the candidate (out-of-sample)?

Primitives

- $P \subset \mathcal{P}$ Problems (m past problems), distribution μ_P
- $A \subset \mathcal{A}$ Actions
- $R \in \mathcal{R}$ Results
- $\succsim \subset \mathcal{A} \times \mathcal{A}$ Preference relation of the society
- $\succsim_1, \ldots, \succsim_n \subset \mathcal{A} \times \mathcal{A}$ Preference relation of each candidate, distribution $\mu_{\succsim}$

Primitives

- $P \subset \mathcal{P}$ Problems (m past problems), distribution μ_P
- $A \subset \mathcal{A}$ Actions
- $R \in \mathcal{R}$ Results
- $\succsim \subset \mathcal{A} \times \mathcal{A}$ Preference relation of the society
- $\succsim_1, \ldots, \succsim_n \subset \mathcal{A} \times \mathcal{A}$ Preference relation of each candidate, distribution $\mu_{\succsim}$
- $\succsim^*$ "The closest" candidate's preference

Presidents as Prediction Machines

Representative Dream Problem: if we have enough candidates n and problems m and a new problem p_{m+1} arises, can we guarantee that $\succsim$ and $\succsim^*$ are close?

Short Answer

Yes!

Short Answer

Yes![1]

[1] If it does not overfit

CBDT Case

$$
\begin{array}{ccccc}
\succsim^* & \Leftrightarrow & \sum & s^*(p,q) & x^*(q) \\
 & & & \downarrow & \downarrow \\
\succsim & \Leftrightarrow & \sum & s(p,q) & x(q)
\end{array}
$$

CBDT Case

$$
\begin{tikzcd}
\succsim^* \arrow[d] & \Leftrightarrow & \sum & s^*(p,q) \arrow[d] & x^*(q) \arrow[d] \\
\succsim & \Leftrightarrow & \sum & s(p,q) & x(q)
\end{tikzcd}
$$

General

$$
\begin{tikzcd}
\succsim^* \arrow[r, phantom, "\Leftrightarrow"] \arrow[d] & f^* \arrow[d] \\
\succsim \arrow[r, phantom, "\Leftrightarrow"] & f
\end{tikzcd}
$$

VC Dimension

Vapnik-Chervonenkis (VC) dimension: greatest number of points that can be shattered

VC Dimension

Representative Dream Problem is true $\Leftrightarrow$ VC dimension is finite

Illustration

$$H(M) = \begin{matrix} 10 & 0 & 0 & 0 \\ \underbrace{0}_{\text{Problem } p_1} & \underbrace{100}_{\text{Problem } p_2} & \underbrace{-2}_{\text{Problem } p_3} & \underbrace{0}_{\text{Problem } p_4} \end{matrix} \begin{matrix} \}\text{ Act } a_1 \\ \}\text{ Act } a_2 \end{matrix}$$

$$U(a_1) = U_{p_5,M}(a_1) = \sum_{(p,a,r)\in M} s(p_5, p)u(r)$$

- Case 1: p_5 is very similar to p_1
- Case 2: p_5 is very similar to p_3
- Case 3: p_5 resemble all of the previous years

Case 1: p_5 is very similar to q_1

$$
\begin{array}{cccccl}
 & \frac{91}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \text{Similar} \\
H(M) = & 10 & 0 & 0 & 0 & \} \text{ Act } a_1 \\
 & \underbrace{0}_{\text{Problem } p_1} & \underbrace{100}_{\text{Problem } p_2} & \underbrace{-2}_{\text{Problem } p_3} & \underbrace{0}_{\text{Problem } p_4} & \} \text{ Act } a_2
\end{array}
$$

$$U(a_1) = U_{p_5,M}(a_1) = \sum_{(p,a,r)\in M} s(p_5, p)u(r)$$

Case 1: p_5 is very similar to q_1

$$\begin{array}{lccccl} & \frac{91}{100} & \frac{3}{100} & \frac{3}{100} & \frac{3}{100} & \text{Similar} \\ H(M) = & 10 & 0 & 0 & 0 & \} \text{ Act } a_1 \\ & \underbrace{0}_{\text{Problem } p_1} & \underbrace{100}_{\text{Problem } p_2} & \underbrace{-2}_{\text{Problem } p_3} & \underbrace{0}_{\text{Problem } p_4} & \} \text{ Act } a_2 \end{array}$$

$$U(a_1) = U_{p_5,M}(a_1) = \sum_{(p,a,r)\in M} s(p_5, p)u(r)$$

$$U(a_1) = \frac{91}{100} * 10 = 9.1$$

$$U(a_2) = \frac{3}{100} * 100 + \frac{3}{100}(-2) + \frac{3}{100}0 = 2.94$$

The Model

- Given a new problem p_{m+1}, agent chooses act a that maximizes
- s similarity function

$$U(a) = U_{p_{m+1},M}(a) = \sum_{(p,a,r)\in M} s(p_{m+1}, p)r$$

Motivation

Next Steps

Possible Flaws of the Author

- PAC-learning (what if there is an ε noise on everything?)
- Similarity on pairs of (problem $\times$ act)
- Goal: general characterization of learnable preferences
- Pindown utility of CBDT
- Find the bounds for number of candidates n and problems m
- Similar interpretations of the problem: finding a partner, a job candidate

Generalized Representative Voting

Then I propose the following voting mechanism:

1. Each citizen responds (perhaps partially) to a questionnaire
2. Each citizen can point out who they hate
3. For each citizen, the vote goes to the closest candidate that is not hated

Characteristics:

- Imperfect information impedes strategic play
- Gibbard–Satterthwaite theorem can be avoided
- Vetoes mitigate morality problems
- Votes can be counted only in areas where people have preferences
- Problems $\approx$ characteristics
- Candidates can have nuanced preferences without needing to follow the party's agenda

Representative Democracy Dream: the Missing Link

PhD Brown Bag - HEC Paris

Mateus Hiro Nagata

June 20, 2024

For every $p, q \in \mathbb{R}^m$

$$s(p, q) = exp[-\boldsymbol{\nu}(p - q)]$$

- The charm of CBDT is that it works without any explicit functional form
- However, there is nothing we can say without an explicit functional form
- Interpretation: think of every dimension of the vector as some important aspect of policy (economic, environmental, social, ...)

Conjecture : all of the important aspects can be incapsulated in those finite variables but not sure how to actually compute distance between them

Generalization Problem: Minimization of Empirical Risk

$$q_1, \ldots, q_m \text{ } \textit{i.i.d.} \text{ questions}$$

Let p be a new question

$$R_{cases}(i) = \int \left(\sum_{q \in H} s(p,q)u(x(q)) - \sum_{q \in H} s_i(p,q)u_i(x(q)) \right)^2 dF_q$$

$$R_{emp,cases}(i) = \frac{1}{m} \sum_{j=1}^{m} \left(\sum_{q \in H} s(q_j,q)u(x(q)) - \sum_{q \in H} s_i(q_j,q)u_i(x(q)) \right)^2$$

Generalization Problem: Minimization of Empirical Risk

$$q_1, \ldots, q_m \textit{ i.i.d.} \text{ questions}$$

Let p be a new question

$$R_{cases}(i) = \text{(Population) average distance of weighted utilities}$$

$$R_{emp,cases}(i) = \text{(Sample) average distance of weighted utilities}$$

Generalization Problem

$$R_{sim}(i) = \int (s(p,q) - s_i(p,q))^2 \, dF_q$$

$$R_{uti}(i) = \int (u(x(q)) - u_i(x(q)))^2 \, dF_q$$

Generalization Problem: Minimization of Empirical Risk

Under what conditions does the convergence

$$\inf_{i \in I} R_{emp,cases}(i) \xrightarrow[m \to \infty]{P} \inf_{i \in I} R_{cases}(i)$$

Idea: if

$$s^* \xrightarrow[m\to\infty, n\to\infty]{P} s$$

$$u^* \xrightarrow[m\to\infty, n\to\infty]{P} u$$

then

$$s^* u^* \xrightarrow[m\to\infty, n\to\infty]{P} su$$

$$\sum s^* u^* \xrightarrow[m\to\infty, n\to\infty]{P} \sum su$$

then finally

$$\succsim^* \xrightarrow[m\to\infty, n\to\infty]{P} \succsim$$

With enough questions m

Generalization Problem

Let, from now on, denote * to be the candidate that minimizes the empirical risk function. Our strategy is to prove:

$$\inf_{i \in I} R_{emp,s}(i) \xrightarrow[m \to \infty]{P} \inf_{i \in I} R_s(i)$$

$$\inf_{i \in I} R_{emp,u_i}(\cdot) \xrightarrow[m \to \infty]{P} \inf_{i \in I} R_{emp,u}(\cdot)$$

Finite VC Dimension

Convergence in probability of these set of functions is guaranteed if the VC dimension is finite (it also converges fast)

VC Dimension (for indicator functions)

VC Dimension h is the maximum number of vectors such that there is (at least one example) that can be completely shattered.

Example: VC Dimension

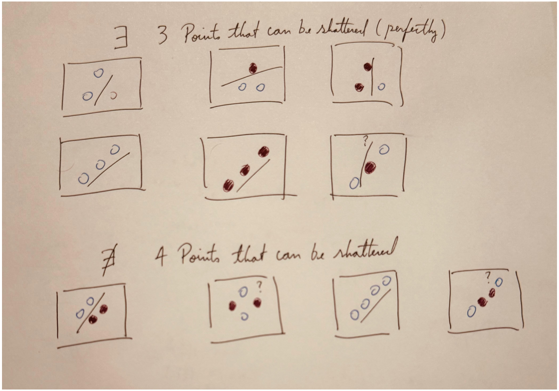

Figure: VC Dimension of a family of linear functions

Example

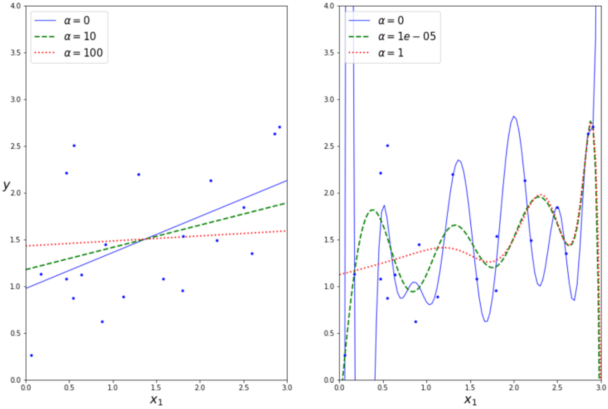

Figure: Caption

VC Dimension of similarity function

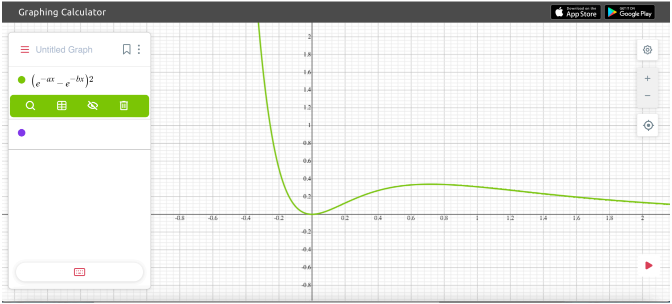

Figure: Caption

Representative Democracy Dream: the Missing Link

PhD Brown Bag - HEC Paris

Mateus Hiro Nagata

June 20, 2024