Learning and Optimizing from Pairwise Preferences

Nir Ailon
Technion

Technion—Cornell/AOL Mini Workshop Oct. 2016

What Are *Preferences*?

- A central token of information obtained from human decision making
  - Given a set of (at least two) objects, which is *preferred*?
    - Search results (information, product, restaurant…)
    - Advertisements
    - Recommendations
    - …

How Can We Get Preferences?

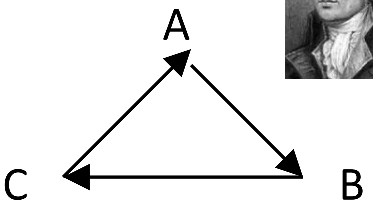
- Experiment #1:
  What is the value of item A?
  Individual judment
  …
  What is the value of item B?
  $\Rightarrow$ Infer A>B if val(A) > val(B)
- Experiment #2:
  Which is preferred, A or B?
  Pairwise judgment
- **In certain settings, the results differ**
  - **Can be verified experimentally**
- **So which information is “better”?**

Why Do We Care About Preferences?

- Comparative judgments easier for most people than absolute judgments
  - No need to calibrate scale
- In economic theory
  - Theory *of revealed preference (Samuelson)*
    - Infer utilities from preference information
  - Theory of *discrete choice (McFadden)*
    - *Explain choices from noisy utility evaluations*
- Social choice theory
  - Impossibility theorems (*Arrow*)
- Observed in the billions
  - Clicks on search results
  - Clicks on ads
  - Online purchases
  - …

What Are the Challenges Underlying Preference Analysis?

- Statistics/ML
  - Modeling, predicting preferences
    "What would X choose if their income was doubled?"
  - Inferring utility from preferences *(utility may be abstract)*
  - Given past preference observations (of a person), what would future ones look like?
- Information theory
  - Which preferences are more informative given past observations?
  - Could we get users to reveal preferences that are most informative to us? At what cost?
- Computational (Combinatorial)
  - Dealing with inconsistency
    - Irrationality
    - Idiosyncrasy
      - When combining preferences from a population



Problem I

Unsupervised Learning: Rank Aggregation

- Given a set V of n alternatives (millions…)
- A population of *voters*
- Each voter provides permutation (full ordering) over V
- How to reconcile them into a single (or few) orderings?
  - Meta-Search: Combining search results from multiple search engines
  - M ta-algor ec ev eri ith one

- Geo int ati d p tic
- Can *extremely* efficiently approximate to within 1.5 [A, Ch ikar, N an 2005] Algo : C t!

- Can a ximate to wit (1+ε) time [Kenyon, Schudy 2006]

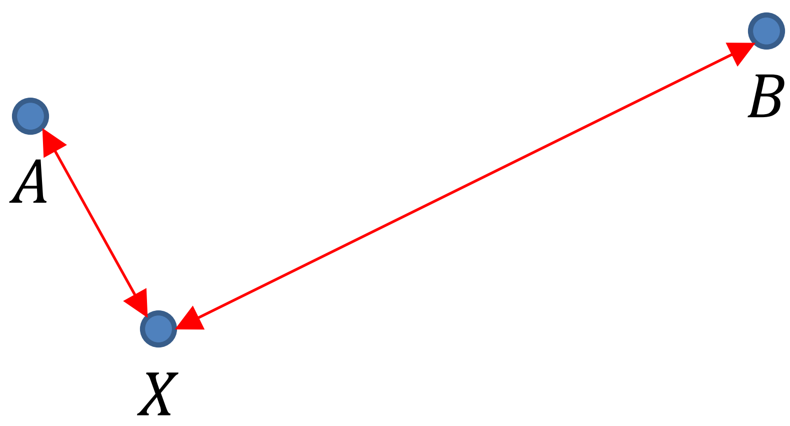
1st 2nd 3rd 4th 5th

Problem II
Supervised Learning From Preferences

- Training data:
  - For each $i = 1..n$
    - associated ground set U_i
    - set of features for each element in U_i
    - list of information bits of the type "$A \succ B$", $A, B \in U_i$
      - Each bit may come from different source (person)
- Learn how to order U_i so as to minimize inconsistencies
- Generalize to unseen sets $U_{i+1} \dots U_{n+m}$
- This is almost standard binary classification, with hypothesis space defined by transitivity constraint
- [Ailon Mohri 2009], [Ailon 2009]
  - By "fixing" inconsistencies, lose accuracy by at most 2

Problem II
Supervised Learning From Preferences

- A particular case of interest: Metric learning
- Given points A, B in the context of point X
  - Is A closer to X than B is?
- [Come to Elad Hoffer's talk later]



Problem III
Active Learning from Preferences

- Same as before, but…
- We are free to choose $A's$ and $B's$
- Each preference bit comes at a price
- Which preferences to choose?
- Standard active learning methods fail

[Ailon Begleiter Ezra 2013]

- At stage i
  - Obtain ground set U_i
  - Start with some guess permutation π_0
  - For $j = 1 \ldots$ [stopping condition]
    - Choose k pairs from U_i, assigning higher weight to pairs that are close w.r.t. π_{j-1}
      - Weight roughly proportional to $1/|rank\_difference|$
    - Fit best permutation to data, obtain π_j
- Do not need more than $O(|U_i| \log^3 |U_i|)$ pairs for each set U_i

Problem IV
Dueling Bandits

- Same as before, but…
- Some pairs $A, B \in U_i$ are more expensive than others
- Think "Recommendations".
  - We want to learn how to recommend items
  - A click on A (given choice of A, B) is form of preference
  - To learn, we must occasionally present pairs A, B one of which is possibly ranked low, but we are not quite sure
  - When we present items that are ranked highly, we *gain* (e.g. by high clickthrough)

exploration

exploitation

Dueling Bandits

- Iterate:
  - User comes at time t
  - We present A_t, B_t
  - Each of A_t, B_t has **hidden** utility $u(A_t), u(B_t)$
  - We observe the outcome of a *duel* between A_t, B_t
    - $Prob[A_t\ wins] = \Phi(u(A_t)\ - u(B_t))$
  - We gain (unobserved) utility $u(A_t) + u(B_t)$
- Goal: Reduce regret
- How to model? What can be proven? [Ailon, Joachims, Karnin 2013], [Ailon, Wolfenfeld 2015]

Summary and Future

- Analysis of preference judgments is fundamental to social choice, economic theory, combinatorics, optimization and machine learning, with applications in ad, search and recommendation optimization.
- Billions of traces of preferences are available through cyberspace every day. It’s not just about what you selected, but also about what you saw but didn’t select.
- Choosing which alternative set to present to a user is an interesting combinatorial and information theoretical problem, with a natural underlying exploration-exploitation tradeoff.
- Lots of theoretical open questions
- Implications to real world data (?)