Predicting User Choice in Video Games

A. Khandelwal[1] W. Ma[2] D. Simchi-Levi[2]

[1]Department of Mathematics
MIT

[2]Operations Research Center
MIT

6/30/17

Introduction

- Given a collection of comparable items $\mathcal{I}$, when a customer c is presented with a subset $S \subset \mathcal{I}$, predict the item they choose
- Many existing methods:
  - Multinomial Logit (MNL)
  - Nested MNL

Question

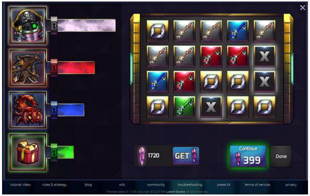
Given a customer c, how can we incorporate information from their past transactions to improve prediction?

Prismata

- Real-time strategy game from Lunarch Studios
- Two minigames in which users must choose skins/emotes of different rarities
- Customers compare items within a given rarity

Prismata

- Choice data from both Black Lab and Armory minigames – 5,000-30,000 transactions for any rarity, 50-300 items per rarity
- Some notable features:
    - No prices
    - No no-purchase option
    - Once an item is obtained, it won't appear again
    - Considerable heterogeneity in transactions per user



Multinomial Logit (MNL)

- Model consists of weights w_j for each item j
- For an item $i \in S$, the probability of choosing i is

$$\frac{e^{w_i}}{\sum_{j \in S} e^{w_j}}$$

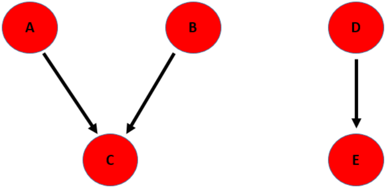
- We solve the convex maximum-likelihood estimation problem (with a regularization term) to find w

Generative Model (Jagabathula and Vulcano 16)

- Set of items $\mathcal{I}$, K customers, $M = |\mathcal{I}|$ items, $n_c \in \{1, \dots, N\}$ transactions per customer.
- We are given fixed underlying MNL weights w, which consists of $w_i \in \mathbb{R}$ for each item $i \in \mathcal{I}$
- For each customer c, draw a latent *preference ordering* $\pi^{(c)}$ from the distribution over permutations of items implied by weights
- For each transaction, choose a subset of the items S, choose $j \in S$ that appears first in $\pi^{(c)}$

Observed Customer Preferences

- We can represent the (incomplete) preference information of a customer as a directed acyclic graph (DAG).
- As we observe transactions, we update the graph accordingly



Prior over Preferences

Lemma

$$P(\pi|W) = \prod_{i=1}^{M} \frac{\exp(w_{\pi_i})}{\sum_{k=i}^{M} \exp(w_{\pi_k})}$$

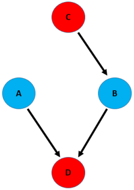
- Select most preferred item first, then next most preferred, etc.
- Conditioning on item π_1 being better than rest of items *does not* affect remaining weights
  - Characteristic of Gumbel distribution
- Does not work in the reverse direction – symmetry doesn't hold!

Computing the Posterior

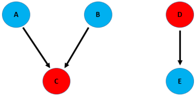
- Given the prior decribed above and the digraph G_c for customer c, the posterior is $P(\pi_c|G_c) \propto P(G_c|\pi_c)P(\pi_c)$.
- $P(G_c|\pi_c) = \mathbb{1}[G_c \text{ consistent with } \pi_c]$
- No nice closed form!
    - Closed form approximation?
    - Approximate sampling?

Closed-form Tree Approximation (Jagabathula and Vulcano 16)

- For an item i, let $\Phi(i) = e^{w_i} + \sum_{k \text{ desc. of } i} e^{w_k}$
- For an assortment S, the probability of choosing $i \in S$ is proportional to $\Phi(i)$
- Exact when graph G_c is a directed forest and S is a subset of the roots



Connected Component Sampling

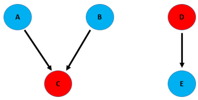


$(a = e^{w_A})$

What is the probability of selecting A out of {A, B, E}?

1. Select weakly connected component with probability proportional to the sum of the weights in the component
2. Compute the set K of items with no predecessors
3. Choose item i from K with probability proportional to $\Phi(i)$
4. Insert i into the permutation. Delete i from the graph, then repeat

Connected Component Sampling (example)



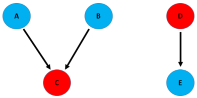
$(a = e^{w_A})$

What is the probability of selecting A out of {A, B, E}?

- Probability of inserting A first: $\frac{a+b+c}{a+b+c+d+e} \cdot \frac{a+c}{a+b+2c}$
- Probability of inserting A after D: $\frac{d+e}{a+b+c+d+e} \cdot \frac{a+b+c}{a+b+c+e} \cdot \frac{a+c}{a+b+2c}$
- Probability of choosing A:

$$\frac{a(a+b+c)(a+b+c+d+2e)}{(a+b+2c)(a+b+c+e)(a+b+c+d+e)}$$

Trickle-up Sampling

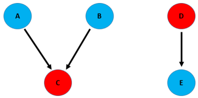


$(a = e^{w_A})$

What is the probability of selecting A out of {A, B, E}?

1. Choose item i with probability proportional to e^{w_i}
2. If i has no predecessors, insert i; otherwise, compute the set L of parent nodes of i.
3. For each item $l \in L$, pick l with probability proportional to
$$\sum_{j \in L, j \neq l} \Phi(j)$$
.
4. Repeat "trickle-up" procedure until termination
5. Delete inserted item from graph; repeat from beginning

Trickle-Up Sampling (example)



$(a = e^{w_A})$

What is the probability of selecting A out of $\{$A, B, E$\}$?

- Probability of choosing A first: $\frac{a}{a+b+c+d+e}$
- Probability of choosing C, then A: $\frac{c}{a+b+c+d+e} \cdot \frac{b+c}{a+b+2c}$
- Probability of choosing E, then inserting A: $\frac{e}{a+b+c+d+e} \cdot \left(\frac{a}{a+b+c+e} + \frac{c}{a+b+c+e} \cdot \frac{b+c}{a+b+2c}\right)$
- Probability of choosing A:

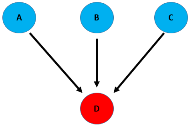
$$\frac{a(a+b+c)(a+b+c+d+2e)}{(a+b+2c)(a+b+c+e)(a+b+c+d+e)}$$

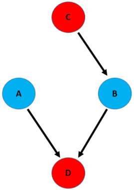
Performance on Prismata Dataset

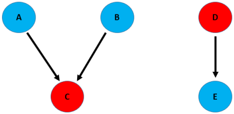
	MNL	ConnComp	Trickle-up	Tree	Naive
Accuracy	0.509	0.525	**0.532**	0.509	0.250
Brier Score	0.619	0.597	**0.595**	0.619	0.750
Log-likelihood	-1.119	-1.090	**-1.085**	-1.119	-1.386

- Average out-of-sample scores for users which have ≥ 2 transactions
- Separate training and test chronologically by user (80:20 split)

Evaluation on Toy Dataset







	Item		
Method	**A**	**B**	**C**
Exact	**0.261**	**0.388**	**0.351**
ConnComp	0.258	0.384	0.359
Trickle-up	0.254	0.4	0.347
Tree	0.266	0.384	0.35

	Item	
Method	**A**	**B**
Exact	**0.225**	**0.775**
ConnComp	0.221	0.779
Trickle-up	0.228	0.772
Tree	0.329	0.671

	Item		
Method	**A**	**B**	**E**
Exact	**0.316**	**0.545**	**0.139**
ConnComp	0.319	0.542	0.139
Trickle-up	0.321	0.545	0.134
Tree	0.287	0.496	0.217

Randomized Experiment

- We can construct random incomplete preference graphs by simulating the generative process
- We fix an item universe, assortment size, number of users, and number of "transactions" (edges in observed preference graph)
- For small item universes (< 10), we can compute the exact distribution
- We compute the distribution when the assortment is the entire universe using the methods described above

	ConnComp	Trickle-up	Tree
Average TV Distance	0.0245	**0.0222**	0.0795

Conclusion

- Introduced two novel sampling algorithms which allow efficient approximation of the posterior over preference orderings given an incomplete preference graph
- Improved prediction performance on Prismata dataset
- Randomized simulated experiments indicate approximation is "close"
- Questions?