

Clicking Data

1 General Notations

- c_j : number of items clicked by user j
- \mathcal{A}_j : set of items clicked by user j , $|\mathcal{A}_j| = c_j$
- \mathcal{A}_j^c : set of items NOT clicked by user j , $|\mathcal{A}_j^c| = n - c_j$
- $\mathbf{B}^j = \{b_1^j, \dots, b_n^j\}$: a binary vector of length n , indicating user j 's clicks. $b_i^j = 1$ if $i \in \mathcal{A}_j$, and 0 otherwise

2 Likelihood function for clicking data

3 Simulation results

3.1 Toy example set up

In this section, we first simulate a dataset of $N = 20$ users, $n = 20$ items, $\boldsymbol{\rho}^0 = 1, \dots, n$, and $\alpha^0 = 5$, by sampling from the Mallows distribution. The number of clicks for each user c_j is drawn from a truncated poisson distribution, with a mean of 5 items, minimum of 1 item and maximum of 17 items. We recommend $k = 2$ items for each user.

3.2 Comparison of $\boldsymbol{\rho}$

A comparison of the distribution of $\boldsymbol{\rho}$ using the Mallows posterior and the pseudo-likelihood is shown on **Figure 1**.

3.3 Heat plots of selected individual users

Heat plots of 2 selected users are shown in **Figure 2**. The distributions appear quite similar, except for a few items, in which using the pseudo-likelihood function produces a somewhat flatter distribution compared to using Mallows directly.

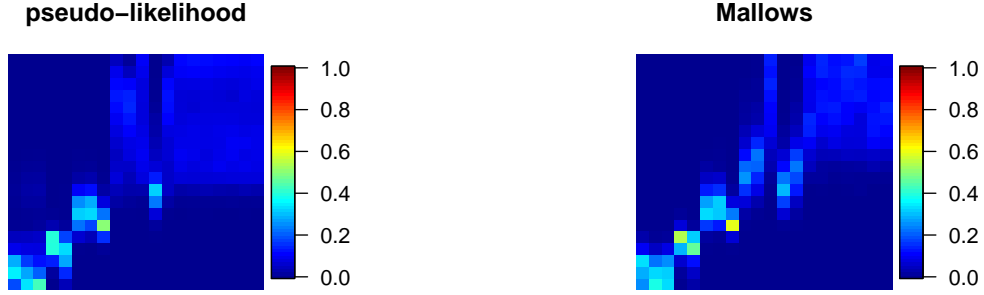


Figure 1: left: pseudo-likelihood, right: Mallows

3.4 Recommendation accuracy

In this particular toy example, the Pseudolikelihood 45% of items are correctly recommended, while Mallows performs slightly better at 50%.

3.5 Special notes

No Gaussian variation is introduced at this stage for the pseudo-likelihood; i.e., $\sigma = 0$. For each user j , the sequence for which items are to be sampled are based on a uniform distribution. (The V-ordering is tried out but don't seem to improve recommendation accuracy here). The α^0 used for sampling is set at 10, instead of its real value 5. But the different choices of α^0 don't seem to affect the resulted heat plots much.

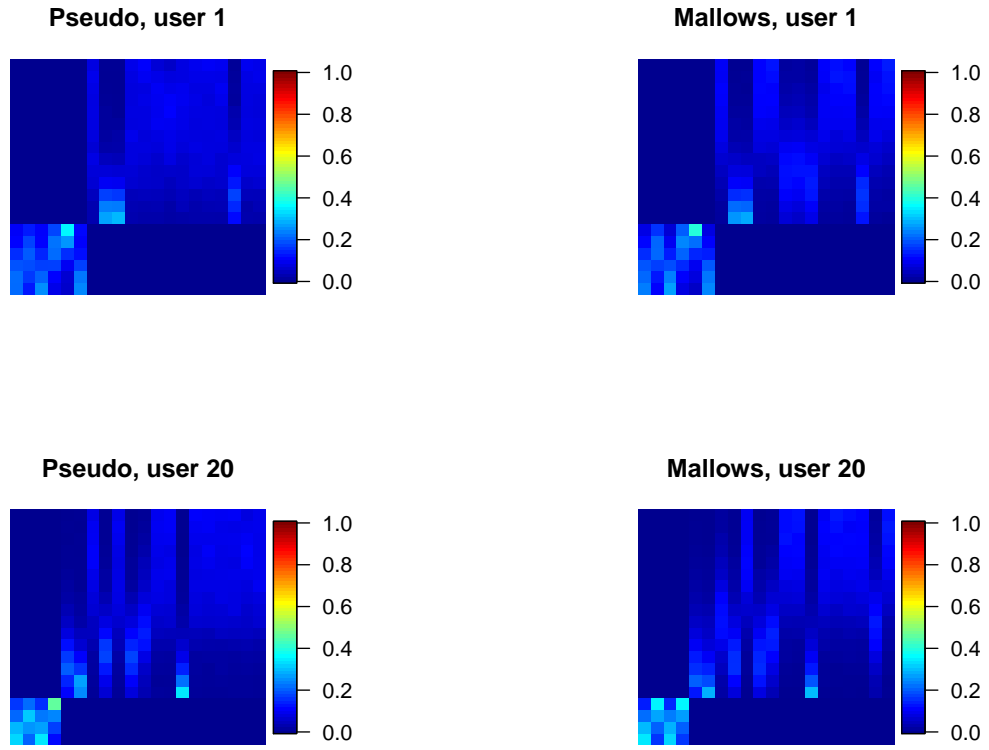


Figure 2: left: pseudo-likelihood, right: Mallows. The items are sorted according to the truth of the individual