**Problem**

- Collaborative Ranking
  - Recommender system problem
  - Focus on ranking of items rather than ratings in the model
  - Performance measured by ranking order of top k items for each user
- State-of-arts are using pairwise loss (such as BPR, and Primal-CR++).
- **But pairwise loss is not the only ranking loss.**
- We will show a new listwise loss works better than pairwise loss in collaborative ranking with implicit feedback.

**STOCHASTIC QUEUING (SQ)**

We denote the set of valid permutations as $\Pi \subseteq S(R, \Omega)$, where $\Omega$ is the set of all pairs $(i, j)$ such that $R_{i,j}$ is observed.

**Model**

- Permutation probability for a single user’s top $k$ ranked items:
  \[
  P_s(k, \bar{m}) (\pi) = \prod_{j=1}^{m} \phi(s_{\pi_j}),
  \]
  where $\pi$ is a particular permutation (or ordering) of the $m$ items, $s$ are underlying true scores for all items, and $\phi$ is some increasing function.
- Can easily be extended to multiple users even with a lot of ties (e.g. 0/1 implicit feedback data).
- Minimize the negative log-likelihood:
  \[
  \min_{\bar{X} \in \mathcal{X}} - \log \sum_{\Pi \in S(R, \Omega)} P_s(k, \bar{m}) (\Pi).
  \]
- The non-convex version can easily be optimized using SGD:
  \[
  \sum_{\Pi \in S(R, \Omega)} - \frac{m}{\sum_{\Pi \in S(R, \Omega)}} \sum_{i=1}^{m} \sum_{j=1}^{m} \phi(u_i^T v_{\pi_j}) + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2).
  \]
  $g = \log \phi$ is the sigmoid function.
- For implicit feedback data, we sample $\bar{m}$ unobserved entries uniformly and append to the back of the list $\rightarrow \bar{m} = (1 + \rho)\tilde{m}$ (For each user (row of $R$), assume there are $\tilde{m}$ 1’s).

**Theory**

The problem of the constrained form

\[
\hat{X} := \arg \min - \log P_X (\Pi) \text{ such that } X \in \mathcal{X}.
\] (1)

The personalized setting:

\[
X_{ij} = u_i^T v_j, u_i, v_j \in \mathbb{R}^r, \|U\|_F \leq c_u, \|V\|_F \leq c_v.
\] (2)

**Corollary 1.** Consider the minimizer, $\hat{X}$, to the constrained optimization, (1). Suppose that there exists a $X^* \in \mathcal{X}$ such that $\Pi_i \sim P_{X_i}$ independently for all $i = 1, \ldots, n$. If $\log \phi$ is $1$-Lipschitz, then in the personalized ranking setting, (2), the KL-divergence between the estimate and the truth is bounded:

\[
D(X^*, \hat{X}) = O_p \left( \frac{\sqrt{m \log m}}{n} \right).
\]

**Conclusion**

- We propose a new collaborative filtering algorithm using listwise loss.
- Our algorithm is faster and more accurate than the state-of-the-art methods on implicit feedback data.
- We provide a theoretical framework for analyzing listwise methods.

**Source Code**

- Julia codes: [https://github.com/wuliwei9278/SQL-Rank](https://github.com/wuliwei9278/SQL-Rank)