Unbiased Pairwise Learning from Implicit Feedback

Yuta Saito

Department of Industrial Engineering and Economics, Tokyo Institute of Technology

Tokyo Tech

Overview

- **Implicit Feedback** is prevalent in the real-world and is often used to construct recommender systems.
- However, this type of feedback is generally **biased**, and de-biasing is necessary to obtain a well-performing recommender.
- The previous solutions have been proved to be biased toward the ideal losses or are based on the simple pointwise approach.
- We developed **a first unbiased estimator for the ideal pairwise loss function** that can be estimated from only observable implicit feedback.

Problem Setting and Challenges

$Y_{u,i}$ is an implicit feedback (e.g., click, view, or log-in) of a pair of user u and item i. Following the exposure model, we assume the following feedback generation model:

$$Y_{u,i} = O_{u,i} \cdot R_{u,i} \tag{1}$$

$$P(Y_{u,i}=1) = P(O_{u,i}=1) \cdot P(R_{u,i}=1) = \theta_{u,i} \cdot \gamma_{u,i} \tag{2}$$

where $O_{u,i}$ is an **exposure random variable** and $R_{u,i}$ is a **relevance random variable**. $\theta_{u,i} = P(O_{u,i}=1)$ and $\gamma_{u,i} = P(R_{u,i}=1)$ are defined as exposure and relevance parameters, respectively.

Under the formulation, **the ideal pointwise and pairwise losses** are defined as

Definition 1. *The ideal* **pointwise** *loss function is defined as*

$$\mathcal{L}^{point}_{ideal}\left(\widehat{\boldsymbol{R}}\right) = \frac{1}{|\mathcal{D}_{point}|} \sum_{(u,i)\in\mathcal{D}_{point}} \gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\,\delta^{(0)}_{u,i} \tag{3}$$

By optimizing this ideal **pointwise** loss function, we can obtain a model that directly predicts the relevance level of user–item pairs.

Definition 2. *The ideal* **pairwise** *loss function is defined as*

$$\mathcal{L}^{pair}_{ideal}\left(\widehat{\boldsymbol{X}}\right) = \frac{1}{|\mathcal{D}_{pair}|} \sum_{(u,i,j)\in\mathcal{D}_{pair}} \gamma_{u,i}\,(1-\gamma_{u,j})\,\ell_{uij} \tag{4}$$

By optimizing this ideal **pairwise** loss function, we can obtain a model that accurately ranks items for each user by their relevance levels.

These losses are defined using the ground truth relevance information. However, it cannot be calculated from implicit feedback. Thus, **estimating the ideal losses from implicit feedback is critical to constructing a de-biased recommender offline**.

To achieve the goal, we have to address the following two difficulties.

- **Positive-Unlabeled Problem**: In the implicit feedback setting, one can only observe $Y_{u,i}$ and both $O_{u,i}$ and $R_{u,i}$ are unobserved. Thus, negative feedback is always unobserved because $Y_{u,i}=0 \Rightarrow O_{u,i}=0$ or $R_{u,i}=0$.
- **Missing-Not-At-Random Problem**: Exposure parameter $\theta_{u,i}$ is not uniform among instances. This introduces troublesome biases such as item popularity bias.

Existing Baseline Estimators

Exposure Matrix Factorization (ExpoMF): ExpoMF adopts a loss function different from the naive MF model to address the positive-unlabeled problem of implicit feedback.

$$\widehat{\mathcal{L}}_{\mathrm{ExpoMF}}\left(\widehat{\boldsymbol{R}}\right) = \frac{1}{|\mathcal{D}_{\mathrm{point}}|} \sum_{(u,i)\in\mathcal{D}_{\mathrm{point}}} \theta'_{u,i}\left(Y_{u,i}\delta^{(1)}_{u,i} + (1-Y_{u,i})\,\delta^{(0)}_{u,i}\right) \tag{5}$$

where $\theta'_{u,i} = \mathbb{E}[O_{u,i}\,|\,Y_{u,i}]$ is the posterior probability represents the **confidence of how much relevance information an interaction indicator** $Y_{u,i}$ **includes**. This estimator is **biased** toward the ideal pointwise loss (i.e., $\mathbb{E}[\widehat{\mathcal{L}}_{\mathrm{ExpoMF}}(\widehat{\boldsymbol{R}})] \neq \mathcal{L}^{\mathrm{point}}_{\mathrm{ideal}}(\widehat{\boldsymbol{R}})$).

Relevance Matrix Factorization (Rel-MF): Rel-MF is currently **the only method that utilizes an unbiased estimator for the ideal pointwise loss as its loss function**. The loss function of Rel-MF is defined as follows.

$$\widehat{\mathcal{L}}_{\mathrm{Rel\text{-}MF}}\left(\widehat{\boldsymbol{R}}\right) = \frac{1}{|\mathcal{D}_{\mathrm{point}}|} \sum_{(u,i)\in\mathcal{D}_{\mathrm{point}}} \frac{Y_{u,i}}{\theta_{u,i}}\delta^{(1)}_{u,i} + \left(1-\frac{Y_{u,i}}{\theta_{u,i}}\right)\delta^{(0)}_{u,i} \tag{6}$$

This estimator is **unbiased** toward the ideal pointwise loss (i.e., $\mathbb{E}[\widehat{\mathcal{L}}_{\mathrm{Rel\text{-}MF}}(\widehat{\boldsymbol{R}})] = \mathcal{L}^{\mathrm{point}}_{\mathrm{ideal}}(\widehat{\boldsymbol{R}})$). However, the pointwise approach is not suitable for the ranking task. Also, **the unbiased pairwise loss function has not yet been proposed**.

Bayesian Personalized Ranking (BPR): BPR is a well-established pairwise algorithm for the top-N recommendations based on implicit feedback [1]. BPR assumes that **interacted items should be ranked higher than all the other non-interacted items** and optimizes the following loss function to obtain latent factors.

$$\widehat{\mathcal{L}}_{\mathrm{BPR}}\left(\widehat{\boldsymbol{X}}\right) = \frac{1}{|\mathcal{D}_{\mathrm{pair}}|} \sum_{(u,i,j)\in\mathcal{D}_{\mathrm{pair}}} Y_{u,i}(1-Y_{u,j})\,\ell\left(\widehat{X}_{uij}\right) \tag{7}$$

where $\ell(\cdot) = -\ln(\sigma(\cdot))$ is generally used[1], and $\widehat{X}_{uij} = U_u^T V_i - U_u^T V_j$ is the difference of predicted scores. This estimator is **biased** toward the ideal pairwise loss (i.e., $\mathbb{E}[\widehat{\mathcal{L}}_{\mathrm{BPR}}(\widehat{\boldsymbol{X}})] = \mathcal{L}^{\mathrm{point}}_{\mathrm{ideal}}(\widehat{\boldsymbol{X}})$).

Proposed Estimator

We propose the following **unbiased estimator for the ideal pairwise loss function**.

$$\widehat{\mathcal{L}}_{\mathrm{unbiased}}\left(\widehat{\boldsymbol{X}}\right) = \frac{1}{|\mathcal{D}_{\mathrm{pair}}|} \sum_{(u,i,j)\in\mathcal{D}_{\mathrm{pair}}} \frac{Y_{u,i}}{\theta_{u,i}}\left(1-\frac{Y_{u,j}}{\theta_{u,j}}\right)\ell_{uij}$$

The proposed estimator weights each data by the inverse of the propensity score.

Proposition 1. *The unbiased pairwise loss is unbiased against the ideal pairwise loss.*

$$\mathbb{E}\left[\widehat{\mathcal{L}}_{unbiased}\left(\widehat{\boldsymbol{X}}\right)\right] = \mathcal{L}^{pair}_{ideal}\left(\widehat{\boldsymbol{X}}\right)$$

We also present the variance of the proposed unbiased estimator.

Theorem 1. *(Variance of the unbiased pairwise loss) Given sets of independent random variables* $\{(Y_{u,i}, O_{u,i}, R_{u,i})\}$, *propensity scores* $\{\theta_{u,i}\}$, *and predicted scoring set* $\widehat{\boldsymbol{X}}$, *the variance of the unbiased pairwise loss is*

$$\mathbb{V}\left(\widehat{\mathcal{L}}_{unbiased}\left(\widehat{\boldsymbol{X}}\right)\right) = \frac{1}{|\mathcal{D}_{pair}|^2}\left(\sum_{(u,i,j)\in\mathcal{D}_{pair}} v_{uij}\ell_{uij} + \sum_{u\in\mathcal{U}}\sum_{i\in\mathcal{I}}\sum_{\substack{(j,k)\in\mathcal{I}\times\mathcal{I}\\ j\neq k}} w_{uijk}\ell_{uij}\ell_{uik}\right)$$

where

$$v_{uij} = \gamma_{u,i}(1-2\gamma_{u,j})\left(\frac{1}{\theta_{u,i}} - \gamma_{u,i}\right) + \gamma_{u,i}\gamma_{u,j}\left(\frac{1}{\theta_{u,i}\theta_{u,j}} - \gamma_{u,i}\gamma_{u,j}\right)$$

$$w_{uijk} = \left(\frac{1}{\theta_{u,i}} - \gamma_{u,i}\right)\gamma_{u,i}(1-\gamma_{u,j})(1-\gamma_{u,k})\ell_{uik}\ell_{uij}$$

The variance depends on the inverse of the exposure parameter that can be small. Thus, we finally introduce the following **non-negative estimator** as a variance reduction technique.

Definition 3. *(Non-negative estimator) When propensity scores and a constant* $\beta \geq 0$ *are given, then the non-negative estimator is defined as*

$$\widehat{\mathcal{L}}_{non\text{-}neg}\left(\widehat{\boldsymbol{X}}\right) = \frac{1}{|\mathcal{D}_{pair}|} \sum_{(u,i,j)\in\mathcal{D}_{pair}} \max\left\{\frac{Y_{u,i}}{\theta_{u,i}}\left(1-\frac{Y_{u,j}}{\theta_{u,j}}\right)\ell_{uij}\left(\widehat{X}_{uij}\right), 0\right\} \tag{8}$$

This technique reduces the variance of the estimator at the cost of introducing some bias.

Experimental Results

We used the **Yahoo! R3 dataset**. This is **an explicit feedback dataset** collected from a song recommendation service. Besides, **it contains users' ratings towards randomly selected sets of music as a test set**. Therefore, **this dataset is suitable for simulating the biased implicit feedback setting**. In the experiment, we treated items rated greater than or equal to 4 as relevant, and the other observed feedback was considered irrelevant.

We compared **WMF, ExpoMF, Rel-MF (with propensity clipping), BPR, and the proposed UBPR (with non-negative)**. The performances of these methods were measured by MAP@K and DCG@K on the test set where $K \in \{1,3,5\}$.

The results are reported in the following figures. The key observations are as follows.

- The previously proposed **unbiased pointwise approach reveals the slightly worse performance than the naive pairwise algorithm (BPR)**, which validates the empirical strength of the pairwise algorithm.
- **The proposed UBPR algorithm significantly outperforms the other methods in all the settings**. The results validate the effectiveness of the proposed de-biasing approach on biased implicit feedback.

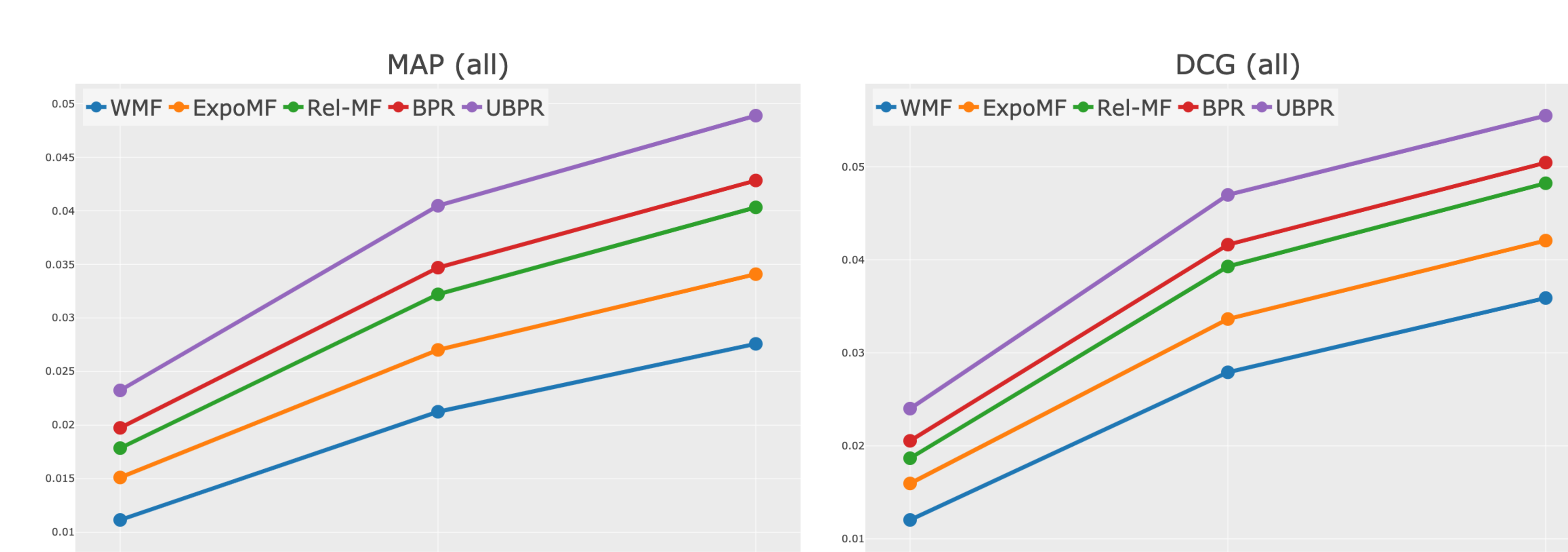

Fig. 1: Results on the Yahoo! R3 dataset with respect to the ranking metrics.

Future Work

The future research directions can be

- Developing an accurate estimation method for the exposure parameter.
- Seeking possible connection with other counterfactual estimators.
- Theoretical Analysis of the variance reduction technique.