Supplementary Material of "Doubly Robust Joint Learning for Recommendation on Data Missing Not at Random"

Xiaojie Wang [1] **Rui Zhang** [1] **Yu Sun** [2] **Jianzhong Qi** [1]

1 Proofs of Lemmas and Theorems

Lemma 1.1 (Zero-Expectation Term). *Given any imputed errors $\hat{\mathbf{E}}$, suppose learned propensities $\hat{\mathbf{P}}$ are accurate, the expectation of the term $\mathcal{E}_{\text{DR}} - \mathcal{E}_{\text{IPS}}$ over all the possible instances of the indicator matrix $\mathbf{O}$ is equal to zero.*

Proof. When the learned propensities are accurate, we have $\Delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. Therefore, we can compute the expectation of the term $\mathcal{E}_{\text{DR}} - \mathcal{E}_{\text{IPS}}$ over all the possible instances of the indicator matrix $\mathbf{O}$ by

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}} - \mathcal{E}_{\text{IPS}}] \\
&\quad = \mathbb{E}_{\mathbf{O}}\left[\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&\quad = \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{\mathbf{O}}\left[\frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&\quad = \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{o_{u,i}}\left[\frac{(\hat{p}_{u,i} - o_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}\right], \\
&\quad = \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \frac{(\hat{p}_{u,i} - p_{u,i})\hat{e}_{u,i}}{\hat{p}_{u,i}}, \\
&\quad = \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\hat{e}_{u,i}, \\
&\quad = 0.
\end{aligned}
$$

This completes the proof. □

Lemma 1.2 (Bias of EIB Estimator). *Given imputed errors $\hat{\mathbf{E}}$, the bias of the EIB estimator is given by*

$$
\text{Bias}\left(\mathcal{E}_{\text{EIB}}\right) = \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} (1 - p_{u,i})\delta_{u,i}\right|.
$$

Proof. According to the definition of the bias, we can derive the bias of the EIB estimator as follows

$$
\begin{aligned}
&\text{Bias}\left(\mathcal{E}_{\text{EIB}}\right) \\
&\quad = |\mathcal{P} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{EIB}}]|, \\
&\quad = \left|\mathcal{P} - \mathbb{E}_{\mathbf{O}}\left[\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} o_{u,i}e_{u,i} + (1 - o_{u,i})\hat{e}_{u,i}\right]\right|, \\
&\quad = \left|\mathcal{P} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} \mathbb{E}_{o_{u,i}}[o_{u,i}e_{u,i} + (1 - o_{u,i})\hat{e}_{u,i}]\right|, \\
&\quad = \left|\mathcal{P} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} (p_{u,i}e_{u,i} + (1 - p_{u,i})\hat{e}_{u,i})\right|, \\
&\quad = \left|\frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} e_{u,i} - \frac{1}{|\mathcal{D}|}\sum_{u,i\in\mathcal{D}} (p_{u,i}e_{u,i} + (1 - p_{u,i})\hat{e}_{u,i})\right|, \\
&\quad = \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} (1 - p_{u,i})\delta_{u,i}\right|,
\end{aligned}
$$

which completes the proof. □

Corollary 3.1 (Double Robustness). *The DR estimator is unbiased when either imputed errors $\hat{\mathbf{E}}$ or learned propensities $\hat{\mathbf{P}}$ are accurate for all user-item pairs.*

Proof. On one hand, when the imputed errors are accurate, we have $\delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. In such case, we can compute the bias of the DR estimator by

$$
\begin{aligned}
&\text{Bias}(\mathcal{E}_{\text{DR}}) \\
&\quad = \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\delta_{u,i}\right|, \\
&\quad = \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\cdot 0\right|, \\
&\quad = 0\,.
\end{aligned}
$$

On the other hand, when the learned propensities are accurate, we have $\Delta_{u,i} = 0$ for $u, i \in \mathcal{D}$. In this case, we can

[1]University of Melbourne, Parkville, Victoria, Australia. [2]Twitter, San Francisco, California, United States of America. Correspondence to: Rui Zhang <rui.zhang@unimelb.edu.au>.

Proceedings of the 36th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.

compute the bias of the DR estimator by

$$\begin{aligned}\text{Bias}(\mathcal{E}_{\text{DR}}) &= \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\delta_{u,i}\right|, \\ &= \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} 0\cdot\delta_{u,i}\right|, \\ &= 0\,.\end{aligned}$$

In both cases, the bias of the DR estimator is zero, which means that the expectation of the DR estimator over all the possible instances of $\mathbf{O}$ is exactly the same as the prediction inaccuracy. This completes the proof. □

Lemma 3.2 (Tail Bound of DR Estimator)**.** *Given imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$, for any prediction matrix $\hat{\mathbf{R}}$, with probability $1-\eta$, the deviation of the DR estimator from its expectation has the following tail bound*

$$\left|\mathcal{E}_{\text{DR}} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}]\right| \le \sqrt{\frac{\log\left(\frac{2}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}}{\hat{p}_{u,i}}\right)^2}.$$

Proof. To keep the notation uncluttered, we define random variable $\ell_{u,i}$ that equals

$$\ell_{u,i} = \hat{e}_{u,i} + \frac{o_{u,i}(e_{u,i}-\hat{e}_{u,i})}{\hat{p}_{u,i}}\,.$$

Since we assume that each observation indicator $o_{u,i}$ follows a Bernoulli distribution with probability $p_{u,i}$, we can rewrite the random variable $\ell_{u,i}$ as follows

$$\begin{cases} P(\ell_{u,i} = \hat{e}_{u,i} + \rho_{u,i}) = p_{u,i}\,, \\ P(\ell_{u,i} = \hat{e}_{u,i}) = 1 - p_{u,i}\,, \end{cases}$$

where $\rho_{u,i}$ is given by

$$\rho_{u,i} = \frac{e_{u,i}-\hat{e}_{u,i}}{\hat{p}_{u,i}} = \frac{\delta_{u,i}}{\hat{p}_{u,i}}\,.$$

We can see that the random variable $\ell_{u,i}$ takes value in the interval $[\hat{e}_{u,i}, \hat{e}_{u,i}+\rho_{u,i}]$ of size $\rho_{u,i}$ with probability 1. Recall that we assume that the observation indicators $\{o_{u,i}|u,i\in\mathcal{D}\}$ are independent random variables, so the random variables $\{\ell_{u,i}|u,i\in\mathcal{D}\}$ are also independent. Therefore, according to the Hoeffding's inequality, for any $\tilde{\epsilon}>0$, we have the following inequality

$$P\left(\left|\sum_{u,i}\ell_{u,i} - \mathbb{E}_{\mathbf{O}}\left[\sum_{u,i}\ell_{u,i}\right]\right| \ge \tilde{\epsilon}\right) \le 2\exp\left(\frac{-2\tilde{\epsilon}^2}{\sum_{u,i}\rho_{u,i}^2}\right).$$

Here, the three summations are summed over all user-item pairs $u,i\in\mathcal{D}$. Setting $\tilde{\epsilon} = \epsilon|\mathcal{D}|$ ($\tilde{\epsilon}>0 \Leftrightarrow \epsilon>0$) in the above inequality and simplifying the inequality with the definition of the DR estimator yield

$$P\left(\left|\mathcal{E}_{\text{DR}} - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\text{DR}}]\right| \ge \epsilon\right) \le 2\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i}\rho_{u,i}^2}\right).$$

Setting the right hand side of the inequality to η and solving for ϵ complete the proof. □

Corollary 3.2 (Tail Bound Comparison)**.** *Suppose imputed errors $\hat{\mathbf{E}}$ are such that $0\le\hat{e}_{u,i}\le 2e_{u,i}$ for $u,i\in\mathcal{D}$, then for any learned propensities $\hat{\mathbf{P}}$, the tail bound of the DR estimator will be lower than that of the IPS estimator.*

Proof. To simplify the notation, we define a constant C as

$$C = \frac{\log\left(\frac{2}{\eta}\right)}{2|\mathcal{D}|^2}\,.$$

Then, we can derive the following inequalities

$$\begin{aligned} & 0\le\hat{e}_{u,i}\le 2e_{u,i} \text{ for } u,i\in\mathcal{D}\,, \\ \Rightarrow\; & \delta_{u,i}^2 \le e_{u,i}^2 \text{ for } u,i\in\mathcal{D}\,, \\ \Rightarrow\; & \sqrt{C\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}}{\hat{p}_{u,i}}\right)^2} \le \sqrt{C\sum_{u,i\in\mathcal{D}}\left(\frac{e_{u,i}}{\hat{p}_{u,i}}\right)^2}\,. \end{aligned}$$

In the last inequality, the left hand side is the tail bound of the DR estimator and the right hand side is the tail bound of the IPS estimator (Schnabel et al., 2016). This completes the proof. □

Theorem 4.1 (Generalization Bound)**.** *For any finite hypothesis space $\mathcal{H}$ [1] of prediction matrices, with probability $1-\eta$, the prediction inaccuracy $\mathcal{P}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^{f})$ of the optimal prediction matrix using the DR estimator with imputed errors $\hat{\mathbf{E}}$ and learned propensities $\hat{\mathbf{P}}$ has the upper bound*

$$\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}^{\ddagger},\mathbf{R}^{o}) + \underbrace{\sum_{u,i\in\mathcal{D}}\frac{|\Delta_{u,i}\delta_{u,i}^{\ddagger}|}{|\mathcal{D}|}}_{\textit{Bias Term}} + \underbrace{\sqrt{\frac{\log\left(\frac{2|\mathcal{H}|}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}^{\S}}{\hat{p}_{u,i}}\right)^2}}_{\textit{Variance Term}},$$

where $\delta_{u,i}^{\S}$ is the error deviation corresponding to the prediction matrix $\hat{\mathbf{R}}^{\S} = \operatorname{argmax}_{\hat{\mathbf{R}}^h\in\mathcal{H}}\left\{\sum_{u,i\in\mathcal{D}}\left(\frac{\delta_{u,i}^h}{\hat{p}_{u,i}}\right)^2\right\}$.

Proof. To simplify notation, let $\mathcal{E}_{\text{DR}}(\hat{\mathbf{R}}) = \mathcal{E}_{\text{DR}}(\hat{\mathbf{R}},\mathbf{R}^{o})$. Note that we can rewrite the finite hypothesis space with $\mathcal{H} = \{\hat{\mathbf{R}}^1,\ldots,\hat{\mathbf{R}}^{|\mathcal{H}|}\}$. We define $\rho_{u,i}^h$ as follows

$$\rho_{u,i}^h = \frac{e_{u,i}^h - \hat{e}_{u,i}^h}{\hat{p}_{u,i}} = \frac{\delta_{u,i}^h}{\hat{p}_{u,i}}\,,$$

[1] For infinite hypothesis spaces, a similar generalization bound can be derived using covering numbers (Anthony & Bartlett, 2009) or other related measures (Maurer & Pontil, 2009).

where $e^h_{u,i}$ and $\hat{e}^h_{u,i}$ are the prediction and imputed errors corresponding to a prediction matrix $\hat{\mathbf{R}}^h \in \mathcal{H}$, respectively. Similarly, we define $\rho^{\S}_{u,i}$ as follows

$$\rho^{\S}_{u,i} = \frac{e^{\S}_{u,i} - \hat{e}^{\S}_{u,i}}{\hat{p}_{u,i}} = \frac{\delta^{\S}_{u,i}}{\hat{p}_{u,i}}\,,$$

where $e^{\S}_{u,i}$ and $\hat{e}^{\S}_{u,i}$ are the prediction and imputed errors corresponding to the prediction matrix $\hat{\mathbf{R}}^{\S} \in \mathcal{H}$, respectively. By making the arguments of uniform convergence and union bound, for any $\epsilon > 0$, we have

$$\begin{aligned}
&P\Big(\Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^{\ddagger}) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^{\ddagger})]\Big| \le \epsilon\Big) \ge 1 - \eta\,,\\
&\Leftarrow P\left(\max_{\hat{\mathbf{R}}^h \in \mathcal{H}} \Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \le \epsilon\right) \ge 1 - \eta\\
&\quad \text{(by the argument of uniform convergence)}\,,\\
&\Leftrightarrow P\left(\bigcup_{\hat{\mathbf{R}}^h \in \mathcal{H}} \Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \ge \epsilon\right) < \eta\,,\\
&\Leftarrow \sum_{h=1}^{|\mathcal{H}|} P\Big(\Big|\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^h)]\Big| \ge \epsilon\Big) < \eta\\
&\quad \text{(by the argument of union bound)}\,,\\
&\Leftarrow \sum_{h=1}^{|\mathcal{H}|} 2\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i\in\mathcal{D}}\{\rho^h_{u,i}\}^2}\right) < \eta\\
&\quad \text{(by Hoeffding's inequality in Lemma 3.2)}\,,\\
&\Leftarrow 2|\mathcal{H}|\exp\left(\frac{-2\epsilon^2|\mathcal{D}|^2}{\sum_{u,i\in\mathcal{D}}\{\rho^{\S}_{u,i}\}^2}\right) < \eta\\
&\quad \text{(by the definition of the prediction matrix } \hat{\mathbf{R}}^{\S})\,.
\end{aligned}$$

We solve the inequality in the last line for ϵ and obtain, with probability $1 - \eta$, the following inequality

$$\begin{aligned}
&\mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^{\ddagger})] - \mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^{\ddagger})\\
&\le \sqrt{\frac{\log\left(\frac{2|\mathcal{H}|}{\eta}\right)}{2|\mathcal{D}|^2}\sum_{u,i\in\mathcal{D}}\left(\frac{\delta^{\S}_{u,i}}{\hat{p}_{u,i}}\right)^2}\,.
\end{aligned} \tag{1}$$

Given the optimal prediction matrix $\hat{\mathbf{R}}^{\ddagger}$, imputed errors $\hat{\mathbf{E}}^{\ddagger} = \{\hat{e}^{\ddagger}_{u,i} | u, i \in \mathcal{D}\}$, and learned propensities $\hat{\mathbf{P}}$, the bias of the DR estimator can be upper bounded as follows

$$\begin{aligned}
&\mathcal{P}(\hat{\mathbf{R}}^{\ddagger}, \mathbf{R}^f) - \mathbb{E}_{\mathbf{O}}[\mathcal{E}_{\mathrm{DR}}(\hat{\mathbf{R}}^{\ddagger})]\\
&\le \frac{1}{|\mathcal{D}|}\left|\sum_{u,i\in\mathcal{D}} \Delta_{u,i}\delta^{\ddagger}_{u,i}\right|,\\
&\le \sum_{u,i\in\mathcal{D}} \frac{|\Delta_{u,i}\delta^{\ddagger}_{u,i}|}{|\mathcal{D}|}\,.
\end{aligned} \tag{2}$$

After adding the two inequalities in Eq. 1 and Eq. 2, we can rearrange the terms to obtain the stated results. □

2 Additional Experiment Results

Table 1 reports the results of rating prediction measured by MAE and MAE-SNIPS on AMAZON and MOVIE.

Table 1: Inaccuracy of rating prediction on MNAR test ratings.

	AMAZON		MOVIE	
	MAE	MAE-SNIPS	MAE	MAE-SNIPS
MF	0.764	0.761	0.745	0.743
PMF	0.767	0.764	0.754	0.748
AutoRec	0.759	0.755	0.743	0.737
Gaussian-VAE	0.730	0.727	0.733	0.728
CPT-v	1.001	0.991	0.956	0.939
MF-HI	0.773	0.770	0.749	0.742
MF-MNAR	0.747	0.739	0.741	0.727
MF-IPS	0.768	0.759	0.752	0.738
MF-JL	**0.721**	0.720	**0.694**	0.693
MF-DR-JL	0.725	**0.717**	0.703	**0.689**

* MF-JL and MF-DR-JL are the proposed approaches.

Table 2 shows the bias and standard deviation of the EIB, IPS, SNIPS, NCIS, and DR estimators in terms of the percentage over the prediction inaccuracy under MAE on the synthetic dataset (see Section 5.2 for details).

Table 2: Bias and standard deviation in terms of percentage over the prediction inaccuracy under MAE. DR is the proposed estimator.

	EIB	IPS	SNIPS	NCIS	DR
ONE	21.7±1.7	20.5±1.8	20.5±1.8	25.8±1.6	**9.8±0.8**
FOUR	66.9±1.7	66.8±1.8	66.8±1.8	84.0±1.8	**24.1±0.6**
ROT	12.7±0.2	12.7±0.4	12.8±0.2	16.0±0.2	**7.1±0.1**
SKEW	10.9±0.3	10.3±0.7	10.5±0.3	12.2±0.2	**7.1**±0.2
CRS	13.5±0.2	12.4±0.6	12.5±0.4	16.5±0.2	**6.9±0.1**

References

Anthony, M. and Bartlett, P. L. *Neural network learning: Theoretical foundations*. cambridge university press, 2009.

Maurer, A. and Pontil, M. Empirical bernstein bounds and sample variance penalization. In *COLT*, 2009.

Schnabel, T., Swaminathan, A., Singh, A., Chandak, N., and Joachims, T. Recommendations as treatments: Debiasing learning and evaluation. In *ICML*, 2016.