

1 はじめに

本資料では、ランキング問題におけるオフ方策評価を扱う。

2 推薦枠内におけるランキングのオフ方策評価

2.1 KPIを設定する

KPI は、推薦枠において発生する総 CV 数とする。

2.2 データの観測構造をモデル化する

対象となる推薦枠においてコンバージョンが発生するまでの一連の流れを以下のように整理する。

1. ユーザーがサービスに訪問する。
2. 訪問ユーザーに対して、推薦枠内に複数のアイテムがランキング形式で表示される。
3. ユーザーの興味において、推薦枠内のいずれかのアイテムがクリックされる。クリックが発生しないケースもあり得る。
4. クリックされたアイテムについて、アイテム詳細ページに遷移した後、ユーザーの意思決定に基づきコンバージョンが発生するか否かが決定される。

ログデータ \mathcal{D} は、以下のように定義する。

$$\mathcal{D} = \{(u_i, S_i, C_i, C_i R_i)\}_{i=1}^n \quad (2.1)$$

ログデータ \mathcal{D} の各要素は、独立同分布に従うとする。ログデータ \mathcal{D} の従う確率分布は、

$$p(\mathcal{D}) = \prod_{i=1}^n p(u_i, S_i, C_i, C_i R_i) = \prod_{i=1}^n p(u_i) \pi_0(S_i | x_{u_i}) p(C_i, R_i | x_{u_i}, S_i) \quad (2.2)$$

とする。

2.3 解くべき問題を特定する

KPI は、推薦枠において発生する総 CV 数とした。これより、方策の価値は推薦枠における期待 CV 数を表す

$$V(\pi) = \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)} \left[\sum_{k=1}^K C_k R_k \right] \quad (2.3)$$

とする。

2.4 観測データのみを用いて問題を解く方法を考える

既存の推薦モデル π_0 の下で収集されたログデータ D_0 のみを用いて、新たな推薦モデル π を導入した場合の方策の価値 $V(\pi)$ を推定するための推定量を考える。また、各推定量について、方策の価値 $V(\pi)$ に対するバイアスとバリエーションを導出し、推定量の性質を考察する。

2.4.1 IPS 推定量

IPS 推定量は、

$$\hat{V}_{\text{IPS}}(\pi; \mathcal{D}) = \frac{1}{n} \sum_{i=1}^n \frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \quad (2.4)$$

で定義される推定量である。

まず、IPS 推定量のバイアスを求めるため、 $p(\mathcal{D})$ の下での期待値を求める。

$$\mathbb{E}_{p(\mathcal{D})} [\hat{V}_{\text{IPS}}(\pi; \mathcal{D})] = \mathbb{E}_{p(\mathcal{D})} \left[\frac{1}{n} \sum_{i=1}^n \frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \right] \quad (2.5)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{p(\mathcal{D})} \left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \right] \quad (2.6)$$

$$= \frac{1}{n} \sum_{i=1}^n \mathbb{E}_{p(u_i)\pi_0(S_i|x_{u_i})p(C_i,R_i|x_{u_i},S_i)} \left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \right] \quad (2.7)$$

$$= \mathbb{E}_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \sum_{k=1}^K C(k) R(k) \right] \quad (2.8)$$

$$= \mathbb{E}_{p(u)\pi_0(S|x_u)} \left[\mathbb{E}_{p(C,R|x_u,S)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \sum_{k=1}^K C(k) R(k) \right] \right] \quad (2.9)$$

$$= \mathbb{E}_{p(u)\pi_0(S|x_u)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \quad (2.10)$$

$$= \mathbb{E}_{p(u)} \left[\mathbb{E}_{\pi_0(S|x_u)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \right] \quad (2.11)$$

$$= \mathbb{E}_{p(u)} \left[\sum_{S \in \mathcal{S}} \pi_0(S|x_u) \frac{\pi(S|x_u)}{\pi_0(S|x_u)} \mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \quad (2.12)$$

$$= \mathbb{E}_{p(u)} \left[\sum_{S \in \mathcal{S}} \pi(S|x_u) \mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \quad (2.13)$$

$$= \mathbb{E}_{p(u)} \left[\mathbb{E}_{\pi(S|x_u)} \left[\mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \right] \quad (2.14)$$

$$= \mathbb{E}_{p(u)\pi(S|x_u)p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \quad (2.15)$$

$$= V(\pi) \quad (2.16)$$

これより、IPS 推定量のバイアスは0である。

次に、IPS 推定量のバリエーションを求める。

$$\mathbb{V}_{p(\mathcal{D})} [\hat{V}_{\text{IPS}}(\pi; \mathcal{D})] = \mathbb{V}_{p(\mathcal{D})} \left[\frac{1}{n} \sum_{i=1}^n \frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \right] \quad (2.17)$$

$$= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}_{p(\mathcal{D})} \left[\frac{\pi(S_i|x_{u_i})}{\pi_0(S_i|x_{u_i})} \sum_{k=1}^K C_i(k) R_i(k) \right] \quad (2.18)$$

$$= \frac{1}{n} \mathbb{V}_{p(u)\pi_0(S|x_u)p(C,R|x_u,S)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \sum_{k=1}^K C(k) R(k) \right] \quad (2.19)$$

$$= \frac{1}{n} \left(\mathbb{E}_{p(u)\pi_0(S|x_u)} \left[\mathbb{V}_{p(C,R|x_u,S)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \sum_{k=1}^K C(k) R(k) \right] \right] \right. \quad (2.20)$$

$$\left. + \mathbb{V}_{p(u)\pi_0(S|x_u)} \left[\mathbb{E}_{p(C,R|x_u,S)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \sum_{k=1}^K C(k) R(k) \right] \right] \right) \\ = \frac{1}{n} \left(\mathbb{E}_{p(u)\pi_0(S|x_u)} \left[\left(\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \right)^2 \mathbb{V}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \right. \quad (2.21)$$

$$\left. + \mathbb{V}_{p(u)\pi_0(S|x_u)} \left[\frac{\pi(S|x_u)}{\pi_0(S|x_u)} \mathbb{E}_{p(C,R|x_u,S)} \left[\sum_{k=1}^K C(k) R(k) \right] \right] \right)$$

アイテム数が多い場合にはランキングの組み合わせも多くなるため、 $\pi_0(S|x_u)$ が小さくなることでバリエーションが大きくなることが予想される。

2.4.2 IIPS 推定量

2.4.3 SNIIPS 推定量

2.4.4 Naive 推定量

2.4.5 DM 推定量

2.4.6 DR 推定量

3 プラットフォーム全体におけるランキングのオフ方策評価