

# 動的環境下における誘因両立性を考慮した推薦方式の提案

二ノ宮仁 松原繁夫

京都大学大学院情報学研究科

## 1. はじめに

コミュニティ志向の情報推薦システムを考える。例えば、経路推薦サービス Waze などが挙げられる。そこでは、利用者はシステムから経路推薦を受けるという点で情報の利用者でもあり、また、道路の混雑状況をシステムに伝えるという点で情報の生産者でもある。推薦者は利用者に多くの道路の混雑状況を調べてもらって情報収集したいという要求を持つ。一方、利用者は早く目的地に到着したいという要求を持つ。ここに目的の不一致が存在する。推薦者は、利用者が自ら進んで推薦された経路を選ぶという制約のもとで、推薦を行う必要がある。

上記の状況は誘因両立性を満たす多腕バンディット問題として定式化できる。多腕バンディット問題はいかに探索と活用を適切に釣り合わせるかという問題であり、様々なアルゴリズムが提案されている。ただし、多くの多腕バンディット問題に関する研究は、推薦者が任意の選択肢を選択できると仮定している。ここで、上記のように推薦者と利用者が分かれている場合、利用者は価値の期待値が低い選択肢を進んで探索することはしない。利用者が自ら進んで探索するという性質を誘因両立性制約と呼び、誘因両立な多腕バンディット問題に関する研究も近年拡大している[1]。しかし、従来研究では選択肢の価値は一定であることが仮定され、動的環境という点が十分に考慮されていない。

そこで本稿では、動的環境下における誘因両立性を考慮した推薦方針の分析を行う。選択肢の価値の変化に関して、変化確率と変化後の価値の範囲が共通知識であるとの仮定のもと、推薦内容にどのような違いが生じるか数学的に分析し、その結果について議論する。

## 2. 関連研究

誘因両立制約を満たす多腕バンディット[1]で前提とされる推薦者、利用者の持つ情報と行動、それに対する推薦方針は以下である。

### ➤ 推薦者の持つ情報と行動：

情報：

- ・ 選択肢の評価予測区間
- ・ 各利用者から得た実際の評価値

行動：

- ・ 推薦方針によって利用者に推薦を行う。

### ➤ 利用者の持つ情報と行動：

情報：

- ・ 選択肢の評価予測区間
- ・ 自分が何番目の利用者なのか

行動：

- ・ 自己が持つ情報から各選択肢の期待値を算出し、最も良い期待値を持つ選択肢が推薦されている場合のみ推薦を受け入れる。
- ・ 推薦を受け入れた場合、推薦者に選択肢の実際の評価値を伝える。

つまり、利用者は過去の利用者からの探索で得た選択肢の実際の評価値を知らないという点で推薦者と情報の非対称性がある。この非対称性を用いることで推薦者は以下の推薦方針で誘因両立制約を満たしながら多くの選択肢の探索を可能とする。

### ➤ 推薦方針：

最初の推薦では、選択肢の予想評価区間から期待値を計算し、最も高い期待値のものから推薦を行い、その後はユーザから報告される探索済みの選択肢の実際の評価値と、未探索の選択肢の予想評価区間を元に推薦を行う。

## 3. 動的環境下での誘因両立を満たす推薦方式

手法[1]に次のような動的環境を組み込み、推薦を行う。推薦システムは利用者に選択肢 A, B のいずれかを推薦する。選択肢のそれぞれの価値を  $R_A, R_B$  で表す。 $R_A \in [-1, 5], R_B \in [-5, 5]$  である。1人目が推薦を受けたあと、選択肢 A の価値が確率  $1 - x_1$  で  $R_A$  から  $R'_A$  に変化する ( $0 \leq x_1 \leq 1$ )。  $R'_A \in [-1, 5]$  である。2人目以降の選択肢の価値は変化しない。これらは、利用者と推薦者どちらに

Proposal of recommendation considering incentive compatibility in dynamic environments

† Jin Ninomiyam, Kyoto University

‡ Shigeo Matsubara, Kyoto University

対しても共通知識であるとする。最初の利用者に対しては、 $E[R_A] > E[R_B]$ であるため、利用者には A が推薦される。ここで、推薦システムは利用者から A の値について報告を受け、その値を知る。

➤ 2番目の利用者:

2番目の利用者に推薦する際、推薦システムにとって以下の場合が考えられる。

・Case 1: 1人目に A が推薦され、かつ、A に変化が生じなかった場合

・Case 2: 1人目に A が推薦され、かつ、A に変化が生じた場合

Case 1:

$R_A$ の大小により A が推薦されるか B が推薦されるか決まり、その閾値を $X_2$ とする。すなわち、 $R_A \in (-1, X_2]$ である時に B が推薦されると考える。この条件が成立する確率は以下で表される。

$$p_1 = \frac{X_2 - (-1)}{5 - (-1)} x_1 = \frac{x_1(X_2 + 1)}{6}$$

また、 $R_A$ と $R_B$ の差の期待値は以下で計算される。

$$\Delta u_1 = E[R_B - R_A] = -\frac{-1 + X_2}{2}$$

Case 2:

$R'_A$ の値は未知であるため、 $R'_A$ を判断基準に用いることはできない。推薦システムは Case1 と Case2 のどちらににいるか区別できないため、Case1 で用いたの同様の基準により、A を推薦するためか B を推薦するか決定する。Case1 と同様にして、

$$p_2 = \frac{(1 - x_1)(X_2 + 1)}{6}, \Delta u_2 = -2$$

以上により、B が推薦された場合の $R_B$ と $R'_A$ の差の期待値は以下で表される。

$$\Delta u = \frac{-(X_2 + 1)(x_1 X_2 - 5x_1 + 4)}{12(p_1 + p_2)}$$

これにより、 $\Delta u \geq 0$ となる。つまり、B が推薦される条件は以下で表される。

$$X_2 \leq \frac{5x_1 - 4}{x_1}$$

$X_2 > -1$ であるので、 $x_1 > \frac{2}{3}$ である場合に、B が推薦される可能性が生じる。

➤ 3番目の利用者:

3番目の利用者に推薦する際、推薦システムにとって以下の場合が考えられる。

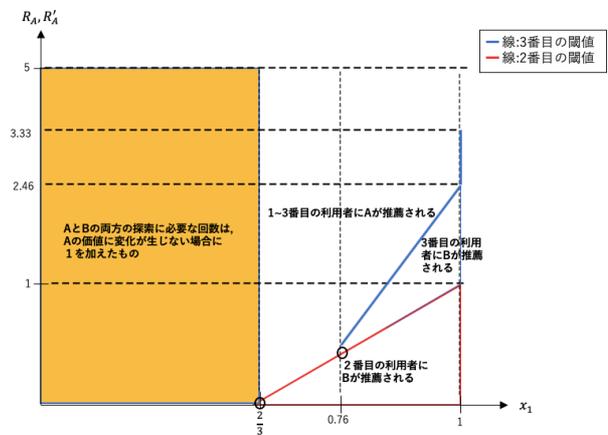
・ $x_1 \leq \frac{2}{3}$ の場合:

この場合、2番目の利用者に B が推薦される可能性はない。推薦システムは2番目の利用者の報告を聞くことで、 $R_A$ に変化があったかどうか知る。これは、利用者の報告を受けた状況と同一である。よって、 $R_A$ に変化がなかった場合、 $R_A \leq$

1であれば B が推薦され、それ以外の場合は A が推薦される。 $R_A$ に変化があった場合も、 $R'_A \leq 1$ であれば B が推薦され、それ以外の場合では A が推薦される。よって、この場合、A と B の両方の探索に必要な回数は、A の価値に変化が生じない場合に 1 を加えたものとなる。

・ $x_1 > \frac{2}{3}$ の場合:

この場合、2番目の利用者に B が推薦される可能性がある。その点を考慮して3番目の利用者に B が推薦される可能性がある場合を求めると、 $x_1 > 0.76$ となる。よってまとめると、1番目から3番目までの利用者に B を推薦するための A の閾値と $x_1$ の関係は以下のような図にまとめられる。



図：1 から 3 番目の利用者に B を推薦するための  $R_A, R'_A$  の閾値と  $x_1$  の関係

#### 4. むすび

本稿では、真の選択肢の価値が非定常であった場合に、動的環境下での誘因両立制約を満たす推薦についての分析を行なった。それにより、変化がある場合とない場合では、変化する確率に依存して探索される範囲が増減することを具体的な数値で分析した。

謝辞：本研究は JSPS 科研費 JP17H00759, JP19H04170 の助成を受けた。

#### 参考文献

[1] Ilan Kremer, Yishay Mansour, and Motty Perry. Implementing the “wisdom of the crowd”. Journal of Political Economy, Vol. 122, No. 5, pp. 988–1012, 2014.