

ECサイトにおける閲覧端末を考慮したクーポン効果最大化モデルの構築

Construction of a coupon effect maximization model considering viewing terminals of the EC site

遠藤 海太郎[†] 山下 遥[‡] 後藤正幸[‡]
Kaitaro Endo Haruka Yamashita Masayuki Goto

1. 研究目的・背景

近年のECサイトの市場規模拡大に伴い、ユーザのアクセスログデータが大量に蓄積されるようになった。蓄積された膨大な量のデータを活用するWebマーケティング技術の重要性が高まっており、その施策の一つとして、サイトにアクセスしているユーザに対する割引クーポンの発行が行われている。しかし、全てのユーザに対して無条件にクーポンを発行した場合、クーポンがなくても購入する顧客、すなわち定価でも購入する予定だったユーザに対しても発行をしてしまうことになり、機会損失が発生する可能性がある。以上から、「購入を迷っていたが、クーポンが表示されることで購入を決心するようなユーザ」を対象としたクーポンの発行が、売上増加の観点から重要であると考えられる。

他方、タブレットやスマートフォンの利用者拡大により、ユーザのWeb閲覧行動も大きく変化しつつある。しかし、閲覧端末の特性に差異があるにも関わらず、多くの研究ではそれらの違いを考慮せずに購買行動の分析が行われている。売上向上のためのWebマーケティングという観点からすれば、ユーザが閲覧する端末の差異を考慮したクーポン発行モデルの確立が必要不可欠であろう。

本研究では上記の問題点の解決のため、ECサイトに蓄積されている膨大な量のアクセスログデータを用いて、ユーザの閲覧端末を考慮したクーポンの発行のためのユーザ分析モデルの構築を行う。具体的には、ユーザの異質性を考慮するために潜在クラスモデル、クーポンの発行による購買を分析するためにロジットモデルを導入するとともに、2つのモデルを別々に考えるのではなく、それらを統合した分析モデルの提案を行う。また、構築された分析モデルを用いて、クーポンの発行が効果的なユーザの購買行動を明らかにする。

2. 準備

2.1 分析対象データ

本研究では、ECサイトの購買行動解析および販促を行う企業であるEmotion Intelligence株式会社の保有しているデータ(大手総合通販カタログサイトのアクセスログ)を用いることでユーザの購買分析モデルの構築を行う。アクセスログデータはユーザID、ページ遷移開始および終了時間、アクセスページに関する情報のほかに、購買の有無、購買金額やクーポン発行の有無、ECサイトにアクセスしてから、離脱するまでの一連の行動に対する一連番号であるセッションIDなどの情報が付与されている。例えば、1ユーザが時間を空けてECサイトに2アクセスした場合、セッション数は2であると考えられる。

[†] 早稲田大学

2.2 潜在クラスモデル

ユーザの購買行動はアイテムへの潜在的嗜好に従うとしてユーザとアイテムの間に潜在クラスを仮定し、購買傾向をモデル化する様々な潜在クラスモデルがマーケティングモデルとして提案されている。その一つであるAspect Model[3]は大量の購買データに対しても、容易に適用可能であり、各ユーザのアイテムへの嗜好を適切に表現できるとして、多くの事例に適用されている。このモデルは似た購買傾向を持つユーザとこれらのユーザに購買されやすいアイテムは同じ潜在クラスに属することを仮定している。また、このモデルではユーザとアイテムが複数の異なる潜在クラスに所属することを許容するという特徴を有する。いま、 I 人からなるユーザ集合 $U = \{u_i; 1 \leq i \leq I\}$ 、 N 個のアイテム集合を $\mathcal{A} = \{a_n; 1 \leq n \leq N\}$ 、 L 個の潜在クラス集合を $\mathcal{Z} = \{z_l; 1 \leq l \leq L\}$ とする。このとき、ユーザ u_i がアイテム a_n を購買する事象 (u_i, a_n) の確率は式(1)で表される。

$$P(u_i, a_n) = \sum_{l=1}^L P(z_l) P(u_i|z_l) P(a_n|z_l) \quad (1)$$

式(1)のパラメータ $P(z_l)$ 、 $P(u_i|z_l)$ 、 $P(a_n|z_l)$ はEMアルゴリズム[4]により推定できる。Aspect Modelでは1購買毎のユーザとアイテムの共起関係をモデル化としている。

2.3 二項ロジットモデル

ECサイトにアクセスしたユーザの各セッションにおける購買確率を求めるために、確率的選択モデルの一つであるロジットモデルを用いる。セッション s ($s = 1, 2, \dots, S$)における購買確率 P_s は式(2)で表される。

$$P_s = \frac{\exp\{V_s\}}{\exp\{V_s\} + 1} \quad (2)$$

ただし、 V_s はセッション s の購買に関する確定的効用で、式(3)で表すような共変量の一次関数とする。

$$V_s = \sum_{l=1}^L \alpha_{ls} \beta_l^T \mathbf{x}_s + \varepsilon \quad (3)$$

ここで、 $\mathbf{x}_s = (x_1^s, x_2^s, \dots, x_K^s)^T$ はモデル作成に用いる K 個の効用変数からなる K 次元ベクトル、 α_{ls} はAspect Modelより求められたクラス l ($1 \leq l \leq L$)におけるセッション s の所属確率、 β_l はクラス l ごとのパラメータベクトル、 ε はランダム効用である。また、 T は転置を表す。上記のモデルについて、最尤推定法によりそれぞれのクラスの変数のパラメータベクトル β_l を求める。

3. 提案モデル

3.1 提案モデル概要

本研究は上記の二手法をベースとすることで、クーポン発行が効果的なユーザの分析モデルを構築する。具体的に

は、ユーザのセッションにおける、ロジットモデルを適用する購買行動とアイテム閲覧履歴との間に潜在クラスを仮定し、潜在クラスごとにクーポンの効果を表現するモデルを提案する。効用 \mathbf{x}_s と購買・非購買の関係とアイテム閲覧履歴 \mathbf{y}_s の共起を事象とし、所属するアイテム閲覧履歴の生起確率に対して潜在クラスごとにロジットモデルと多項分布を仮定した潜在モデルを定式化する。図 1 に本研究における提案モデルのグラフィカルモデルを示す。

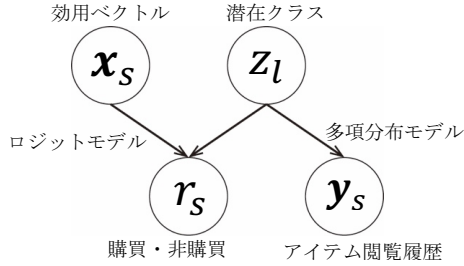


図 1: 本研究における提案モデルのグラフィカルモデル

3.2 提案モデルの定式化

ユーザが EC サイトを閲覧する 1 セッションを 1 閲覧行動データとし、全解析対象データ数を S とする。また、 S 個の閲覧行動データで出現する二項ロジットモデルにおける効用変数ベクトル集合を $\mathcal{X} = \{\mathbf{x}_s; 1 \leq s \leq S\}$ とする。 $r_s (1 \leq s \leq S)$ は s 番目のセッションが購買するとき 1、それ以外で 0 をとる。ここで、閲覧行動 \mathbf{x}_s の特徴量を、全アイテム N 個のうちの任意の n 個目 ($1 \leq n \leq N$) のアイテムの閲覧回数 y_n^s を要素とした N 次元ベクトル $\mathbf{y}_s = (y_1^s, \dots, y_n^s, \dots, y_N^s)$ で表すことにする。さらに、 L 個の潜在クラス集合を $\mathcal{Z} = \{z_l; 1 \leq l \leq L\}$ とし、各潜在クラス z_l ごとに、 n 個目のアイテム閲覧回数の割合 θ_{ln} を要素とした多項分布のパラメータを $\boldsymbol{\theta}_l = (\theta_{l1}, \dots, \theta_{ln}, \dots, \theta_{lN})$ と定義する ($\sum_{n=1}^N \theta_{ln} = 1$)。このとき、効用ベクトル \mathbf{x}_s とアイテム閲覧履歴 \mathbf{y}_s で表される s 番目の閲覧行動データの共起確率 $P(r_s, \mathbf{y}_s)$ は式(4)で表現される。

$$P(r_s, \mathbf{y}_s) = \sum_{l=1}^L P(z_l) P(r_s | z_l) P(\mathbf{y}_s | z_l) \quad (4)$$

$$= \sum_{l=1}^L P(z_l) P(r_s | z_l) \frac{(\sum_{n=1}^N y_n^s)!}{\prod_{n=1}^N y_n^s!} \prod_{n=1}^N \theta_{ln}^{y_n^s}$$

3.3 EM アルゴリズムによるパラメータ推定

式(4)に対する対数尤度関数 LL は次の式(5)で与えられる。

$$LL = \sum_{l=1}^L \log P(r_s, \mathbf{y}_s) \quad (5)$$

式(5)を最大化するパラメータを $P(z_l)$, θ_{ln} , $P(r_s | z_l)$ は以下の EM アルゴリズムを用いて推定することができる。

[E-step]

$$P(z_l | r_s, \mathbf{y}_s) = \frac{P(z_l) P(r_s | \mathbf{x}_s, z_l) P(\mathbf{y}_s | z_l)}{\sum_{l=1}^L P(z_l) P(r_s | \mathbf{x}_s, z_l) P(\mathbf{y}_s | z_l)} \quad (6)$$

[M-step]

$$P(z_l) = \frac{\sum_{s=1}^S P(z_l | \mathbf{x}_s, r_s, \mathbf{y}_s)}{\sum_{l=1}^L \sum_{s=1}^S P(z_l | \mathbf{x}_s, r_s, \mathbf{y}_s)} \quad (7)$$

$$\theta_{ln} = \frac{\sum_{s=1}^S P(z_l | r_s, \mathbf{y}_s) y_n^s}{\sum_{l=1}^L \sum_{s=1}^S P(z_l | r_s, \mathbf{y}_s) y_n^s} \quad (8)$$

ここで、 $P(r_s | z_l)$ は EM アルゴリズムの M-step において、二項ロジットモデルのパラメータ最適化の過程で推定を行う。以下にそのアルゴリズムを示す。

[step 1]

d_s , α_{ls} , \mathbf{x}_s が与えられた下で二項ロジットモデルにおける対数尤度関数 LL_{logit} を最大にする $\boldsymbol{\beta}_l$ を求める。

$$LL_{logit} = \log \prod_{s=1}^S p_s^{d_s} (1 - p_s)^{1-d_s}$$

$$= \sum_{s=1}^S \left[d_s \left(\sum_{l=1}^L \alpha_{ls} \boldsymbol{\beta}_l^T \mathbf{x}_s \right) - \log \exp \left\{ \left(\sum_{l=1}^L \alpha_{ls} \boldsymbol{\beta}_l^T \mathbf{x}_s \right) + 1 \right\} \right] \quad (9)$$

ただし、 d_s はセッション s において購買したとき 1、それ以外 0 をとる。

[step 2]

\hat{p}_s , $\boldsymbol{\beta}_l$ を固定した下で、式(10)において α_{ls} の非線形最適化を行う。

$$\hat{p}_s = \frac{\exp \left\{ \sum_{l=1}^L \alpha_{ls} \boldsymbol{\beta}_l^T \mathbf{x}_s \right\}}{\exp \left\{ \sum_{l=1}^L \alpha_{ls} \boldsymbol{\beta}_l^T \mathbf{x}_s \right\} + 1} \quad (10)$$

ただし、 \hat{p}_s は目標購買確率を表し、セッション s において購買したとき 1、それ以外 0 をとる。

[step 3]

$$\alpha_{ls, new} = P(z_l | r_s) = \hat{p}_s$$

として、ベイズの定理より式(11)を用いて $P(r_s | z_l)$ を得る。

$$P(r_s | z_l) = \frac{P(z_l | r_s) P(r_s)}{P(z_l)} \quad (11)$$

[step 4]

M-step により得られた式(7)(8)(11)を、E-step へ代入する。E-step, M-step を式(5)が収束するまで繰り返し、パラメータ推定を行う。

4. 提案モデルにおけるクーポンの発行

ユーザのセッションごとにクーポンの効果的な発行を行うために、提案モデルにより得られた各潜在クラスの効用ベクトルパラメータ $\boldsymbol{\beta}_l$ を参照する。その際に $\boldsymbol{\beta}_l$ においてクーポンによる効用が高いクラスに対して発行を行う。また、閲覧端末の効用も同様に参照することで、閲覧端末による閲覧行動の差異を考慮したクーポンの発行が可能になる。実際の EC サイトから得られた閲覧行動データを用いた実験結果においては、当日に発表する。

参考文献

- [1] 久松俊道, 外川隆司, 朝日弓末 "EC サイトにおける購買予兆発見モデルの提案", オペレーションズ・リサーチ: 経営の科学, vol.52, pp. 93-100, 2013.
- [2] D. Van den Poel and W. Buckinx, "Predicting Online-purchasing Behaviour," European Journal of Operational Research, 166(2), pp. 557-575, 2005.
- [3] T. Hofmann, "Probabilistic Latent Semantic Analysis," UAI, 289-296, 1999.
- [4] 宮川雅巳., "EM アルゴリズムとその周辺," 応用統計学, Vol.16, No.1, pp.1-21, 1987.