

電子商店街における店舗情報を考慮した商品推薦アルゴリズムに関する研究

経営情報学研究

5218F008-1 大堀祐一
指導教員 後藤正幸

Study on Item Recommendation Considering Store Information in Cyber Mall

OHORI Yuichi

1. はじめに

近年、EC サイト上では大量の商品が扱われるようになり、消費者の嗜好も多様化している。これに伴い、自動で各顧客の嗜好に合致した商品を提示する推薦システムの重要性が高まっている。推薦するアイテムを決定する手法の枠組みとして協調フィルタリングがある。これに属する手法は、基本的にユーザの購買履歴を用いて未購買アイテムの購買確率を推定し、購買の可能性が高いアイテムを推薦する。しかし現実において、例えば直近でキッチン用品を多く購買しているユーザは、その後もキッチン用品を購買する可能性が高いといったように、各アイテムの購買には時間的関係性があると考えられる。そのため、推薦アイテムを決定する際には、ユーザの購買順序も考慮する必要がある。以上を踏まえ、本研究では購買順序も考慮したモデルである Translation-based Recommendation[1] (以下、Transrec) に着目する。TransRec はユーザがあるアイテムを買った後に、次に買うアイテムを選択する際の嗜好をモデル化している。そして、ユーザが最後に買ったアイテムから、次に買う可能性の高いアイテムを推薦する。

ここで、EC サイトの一形態として、電子商店街(ショッピングモール型 EC サイト)がある。これは、一つのサイト上に複数の異なる出店者が商店ページをオープンし、それぞれの商店が商品販売を行うというものである。従来の EC サイトとは異なり、ユーザは商店ページごとに閲覧し購買を行うことが可能となっている。この電子商店街の購買履歴データに対しても TransRec を適用することは可能である。しかし、購買履歴データのみでは特定の商店ページのみで購買する傾向のあるユーザに対して、他の商店のアイテムを推薦してしてしまうなど、電子商店街におけるユーザの嗜好を十分に学習できない可能性がある。

そこで、本研究では商店情報を考慮した TransRec モデルを提案する。そのために商店と商店が与えられたもとのアイテムに対する嗜好をモデルに加えることで、アイテムベクトルの学習に商店のもとのユーザの嗜好も反映させる。この方法で従来手法のユーザの嗜好を電子商店街におけるユーザの嗜好に近づけ、予測精度の向上を目指す。さらに実データを用いた実験により、モデルの有用性を示す。

2. 関連研究

2.1. 嗜好を考慮した推薦システム

各ユーザの嗜好を考慮して、アイテムを推薦する手法は多数ある。これらの手法は、好みのジャンルなどからアイテ

ムを決定する内容ベースフィルタリングと類似した他のユーザを参考にアイテムを決定する協調フィルタリングに分類される。

協調フィルタリングにおいて、以前は行列因子分解に関する手法が低ランク近似を用いることで効率良くデータを処理することが可能となり、高い予測精度を出せるため注目を集めていた。近年では、大量のデータを用いてデータ同士の関係から特徴空間を構成し、データをその空間に埋め込むことでデータを低ランク近似同様に低次元ベクトルで表現する埋め込み空間モデルが注目を集めている。埋め込み空間モデルには Item2vec[2] などの手法があるが、これらの手法は時系列を考慮しておらず、ユーザにとって既に興味のないアイテムを推薦してしまう可能性がある。そこで、本研究は埋め込み表現モデルであり、かつ時系列を考慮している TransRec に着目する。

2.2. Translation-based Recommendation

TransRec とは、アイテムとユーザの嗜好を同一空間(埋め込み空間)上のベクトルとした時に、最後にユーザが買ったアイテムとユーザの嗜好ベクトルの和と距離が近く、かつ人気度が高いアイテムを推薦する手法である。いま、次元数を K 、ユーザを $u \in \mathcal{U}$ 、アイテムを $i, j \in \mathcal{I}$ 、ユーザ u が購買したアイテムリストを $S^u = \{S_1^u, S_2^u, \dots, S_{|S^u|}^u\}$ と定義する。

TransRec の学習は \mathcal{R}^K で表される K 次元ユークリッド空間上で行われる。この時、アイテム i のアイテムベクトルを $\vec{r}_i \in \mathcal{R}^K$ とする。また、あるアイテムから次に買うアイテムを選択するユーザ u の嗜好ベクトルを $\vec{T}_u = \vec{t} + \vec{t}_u \in \mathcal{R}^K$ とする。 \vec{t} は全ユーザに共通する嗜好であり、 \vec{t}_u は各々のユーザ固有の嗜好である。ユーザ u における、あるアイテム i からその次に買うアイテム j への遷移関係を以下のように定義する。

$$\vec{r}_i + \vec{T}_u \approx \vec{r}_j \quad (1)$$

このように定式化することで、ユーザ固有であり、かつユーザが購入したすべてのアイテム S^u を考慮した嗜好のモデル化が可能となる。そして、購買確率はアイテムの人気と(1)式の遷移関係によって決まると考え、ユーザ u がアイテム i を買った後にアイテム j を購買する確率を以下のように定義する。

$$p_{u,i,j} = \beta_j - d(\vec{r}_i + \vec{T}_u, \vec{r}_j) \quad (2)$$
$$s.t. \vec{r}_i \in \Psi \subseteq \mathcal{R}^K, \forall i \in \mathcal{I}$$

$d(x, y)$ は x と y の距離を表し、 Ψ は中心点からの距離が 1 以下の K 次元空間 (単位球) を表している。これは次元の呪いの影響を減衰させる効果がある。 β_j はアイテム j の人気度を表しており、 $p_{u,i,j}$ が最も大きくなるアイテムを推薦する。

(2) 式のパラメータは Sequential Bayesian Personalized Ranking (S-BPR) [3] を用いたランキング最適化によって推定する。ユーザが今までに購入したことのあるアイテム j が購入したことのないアイテム $j' \in \mathcal{I} \setminus j$ よりも (2) 式の購買確率が高くなるようにパラメータを学習する。定式化すると以下ようになる。

$$\begin{aligned} \hat{\Theta} &= \arg \max_{\Theta} \ln \prod_{u \in \mathcal{U}} \prod_{j \in S^u} \prod_{j' \notin S^u} p(j_{u,i}, j' | \Theta) p(\Theta) \\ &= \arg \max_{\Theta} \sum_{u \in \mathcal{U}} \sum_{j \in S^u} \sum_{j' \notin S^u} \ln \sigma(p_{u,i,j} - p_{u,i,j'}) - \Omega(\Theta) \end{aligned} \quad (3)$$

Θ は推定すべきパラメータであり、 $\Theta = (\beta_i, \vec{r}_i, \vec{t}_u, \vec{t} | i \in \mathcal{I}, u \in \mathcal{U})$ で与えられる。 $j_{u,i}$ はユーザ u がアイテム i を買った後に購入したアイテムを表す。 $\sigma(\cdot)$ はシグモイド関数であることを表す。 $\Omega(\Theta)$ は L_2 正則化項を表す。

式 (3) を最大化するようにパラメータを更新する。パラメータは確率的勾配上昇法によって更新される。

$$\Theta \leftarrow \Theta + \epsilon \left(\sigma(p_{u,i,j} - p_{u,i,j'}) \frac{\partial(p_{u,i,j} - p_{u,i,j'})}{\partial \Theta} - \lambda_{\Theta} \Omega(\Theta) \right) \quad (4)$$

ϵ は学習率であり、 λ_{Θ} は正則化のハイパーパラメータである。

3. 提案手法

3.1. 着想

電子商店街から得られる購買履歴データに対して、従来手法である TransRec はそのまま適用可能である。しかし、ユーザによっては商店ページを閲覧してからアイテムを比較検討し購買に至る。このようなユーザの購買履歴データに対してそのまま従来手法を適用した場合、商店が与えられたもとのアイテムに対する嗜好を学習しておらず、商店に関する嗜好が推薦に反映されない。そこで、商店ベクトルを埋め込み空間に布置すると同時に商店が与えられたもとのアイテムに対する嗜好ベクトルも埋め込み空間に追加することを考える。商店を $l \in \mathcal{L}$ 、商店を表すベクトルを $\vec{s}_l \in \mathcal{R}^K$ 、商店が与えられたもとのアイテムに対する嗜好ベクトルを $\vec{T}'_u = \vec{v} + \vec{t}'_u \in \mathcal{R}^K$ とする。 \vec{v} は全ユーザに共通する商店が与えられたもとのアイテムに対する嗜好ベクトルを表しており、 \vec{t}'_u は各ユーザの商店が与えられたもとのアイテムに対する嗜好ベクトルを表現している。この時、ユーザ u が商店 l においてアイテム i を購入した後に j を購入する確率を以下のように定義する。

$$\begin{aligned} p_{u,i,j,l} &= \beta_j - d(\vec{r}_i + \vec{T}'_u, \vec{r}_j) - d(\vec{s}_l + \vec{T}'_u, \vec{r}_j) \\ \text{s.t. } r_i &\in \Psi \subseteq \mathcal{R}^K, \forall i \in \mathcal{I} \\ s_l &\in \Psi \subseteq \mathcal{R}^K, \forall l \in \mathcal{L} \end{aligned} \quad (5)$$

このように定式化することで、従来手法に商店ベクトル \vec{s}_l と商店が与えられたもとのアイテムに対する嗜好 \vec{T}'_u を加えることが可能となる。これにより、ある商品が人気であり、かつ購買順序の観点からはユーザの嗜好に合致した商品であったとしても、商店が与えられたもとのアイテムに対する嗜好に合致していない場合は予測購買確率が低くなるなど、電子商店街におけるユーザの嗜好を考慮した推薦が可能になると考えられる。

ここで、図 1 に (5) 式のイメージ図を示す。図 1 は一つの商店のもので連続で三つのアイテムを購入したユーザを学習した際の (5) 式の全てのベクトルを示している。図 1 のように同じ商店における購買が多いユーザを考えた時、同じ商店で購入された全てのアイテムに商店ベクトルと商店が与えられたもとのアイテムに対する嗜好ベクトルの和を十分に近づけることは不可能である。そのため、同じ商店で購買を続けるユーザの嗜好を埋め込み空間上で表現することができない。

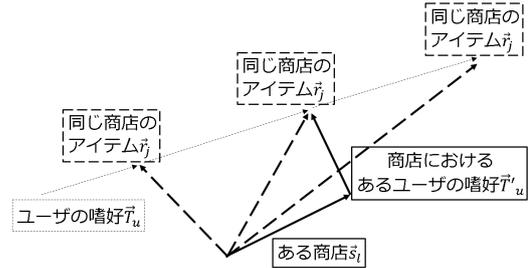


図 1: (5) 式のイメージ

そこで、本研究では TransRec をベースとして、直線で表現される商店、および商店が与えられたもとのアイテムに対する嗜好ベクトルを追加したモデルを提案する。このモデルは、アイテムに対する嗜好と商店が与えられたもとのアイテムに対する嗜好を同時に表現している。さらに、商店をベクトルによる直線で表現することにより、同じ商店で購入されたすべてのアイテムに商店ベクトルと商店が与えられたもとのアイテムに対する嗜好ベクトルの和を十分に近づけることが可能となる。これにより、商店が与えられたもとのアイテムに対する嗜好も十分に学習され、その嗜好を反映した推薦が可能となる。

3.2. 提案手法

商店 l を表す直線を $\vec{s}_l = \alpha \vec{m}_l + \vec{n}_l$ とする。 α は媒介変数である。ユーザ u が購入した商店リストを $S'^u = \{S'_1{}^u, S'_2{}^u, \dots, S'_{|S'^u|}{}^u\}$ とする。図 2 に提案手法のイメージを示す。図 2 は図 1 同様に一つの商店のもので連続で三つのアイテムを購入したユーザを対象にしており、 α が変化した際の商店を表す \vec{s}_l を示している。図 2 のように α が変化する

ることで、直線上の点であればどの点においても同一の商店ベクトルとして扱うことができる。

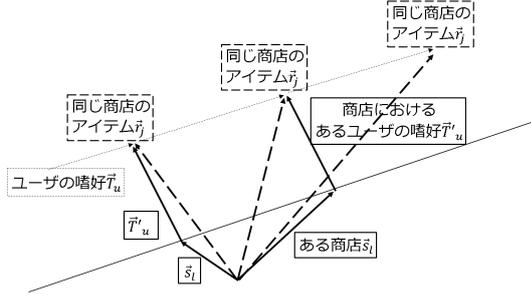


図 2: 提案手法のイメージ

このように、同一の商店で買われたアイテムすべてに対して十分な学習が可能となる。

ユーザ u と商店 l において購入するアイテム j との関係を (6) 式のように定義する。

$$\alpha \vec{m}_l + \vec{n}_l + \vec{T}'_u \approx \vec{r}_j \quad (6)$$

このときの α は以下のように定義する。

$$\alpha = \frac{-\sum_k \vec{m}_l \cdot (\vec{n}_l - (\vec{r}_i - \vec{T}'_u))}{\sum_k \vec{m}_l^2} \quad (7)$$

最終的なユーザ u が商店 l においてアイテム i の後に j を購入する確率を以下のように定義する。

$$\begin{aligned} p_{u,i,j,l} &= \beta_j - d(\vec{r}_i + \vec{T}'_u, \vec{r}_j) - d(\alpha \vec{m}_l + \vec{n}_l + \vec{T}'_u, \vec{r}_j) \\ &\text{s.t. } r_i \in \Psi \subseteq \mathcal{R}^K, \forall i \in \mathcal{I} \\ &\quad m_l \in \Psi \subseteq \mathcal{R}^K, \forall l \in \mathcal{L} \\ &\quad n_l \in \Psi \subseteq \mathcal{R}^K, \forall l \in \mathcal{L} \\ \alpha &= \frac{-\sum_k \vec{m}_l \cdot (\vec{n}_l - (\vec{r}_i - \vec{T}'_u))}{\sum_k \vec{m}_l^2} \end{aligned} \quad (8)$$

(8) 式のパラメータを推定することを考える。従来手法ではユーザが買っていないアイテムよりも買ったアイテムの方が購買確率が高くなるように学習を行った。しかし、提案手法においては、①ユーザが利用する商店において買っていないアイテム（ユーザの購買履歴に存在しないアイテム）と②利用していない商店において買っているアイテム（ユーザの購買履歴に存在するアイテム）の両方よりも、利用する商店において実際に購入したアイテムの方が購買確率が高くなるように学習を行う必要がある。

そこで、以下のように①と②に対応する二つの目的関数を設定し、交互に二つの目的関数を最大化するようにパラメータを更新する。

$$\begin{aligned} \Theta_1 &= \\ \arg \max_{\Theta_1} &\sum_{u \in \mathcal{U}} \sum_{j \in S^u} \sum_{j' \notin S^u} \sum_{l \in \mathcal{L}} \ln \sigma(p_{u,i,j,l} - p_{u,i,j',l}) \\ &\quad - \Omega(\Theta_1) \end{aligned} \quad (9)$$

$$\begin{aligned} \Theta_2 &= \\ \arg \max_{\Theta_2} &\sum_{u \in \mathcal{U}} \sum_{j \in S^u} \sum_{l \in \mathcal{L}} \sum_{l' \notin S^u} \ln \sigma(p_{u,j,l} - p_{u,j,l'}) - \Omega(\Theta_2) \end{aligned} \quad (10)$$

Θ_1 はパラメータ $(\beta_i, \vec{r}_i, \vec{t}_u, \vec{t}, \vec{m}_l, \vec{n}_l, \vec{t}'_u, \vec{t}' | i \in \mathcal{I}, u \in \mathcal{U}, l \in \mathcal{L})$ を表し、 Θ_2 はパラメータ $(\vec{r}_i, \vec{m}_l, \vec{n}_l, \vec{t}'_u, \vec{t}' | i \in \mathcal{I}, u \in \mathcal{U}, l \in \mathcal{L})$ を表す。

3.3. 購買確率

現実では、電子商店街においても通常の EC サイトと同じようにユーザが好むであろうアイテムが推薦され、商店ごとに推薦アイテムは決定されない。そのため、提案手法は商店 l においてユーザ u がアイテム i を買った後にアイテム j の予測購買確率を推定しているが、実応用を考えた時、商店関係なく予測購買確率を算出する必要がある。そこで、本研究では提案手法の予測購買確率を以下のように定義する。

$$\hat{p}_{u,i,j} = \beta_j - d(r_i + T_u, r_j) \quad (11)$$

ここでは商店に関する距離の項を考慮せず、従来の TransRec と同様のパラメータを用いて予測購買確率が計算される。この予測購買確率を用いて、実験では従来手法である TransRec との比較を行う。

4. 予測精度の評価実験

提案手法の有効性を示すため、電子商店街の実データを用いた評価実験を行い、従来手法である TransRec との比較を行った。

4.1. 実験条件

実験では、2010 年度楽天市場データセット [4] の購買履歴データ 3,666,073 件を用いた。ユーザ数は 333,543 (人)、アイテム数は 257,082 (件)、商店数は 17,715 (件) である。このうち、評価数が 5 件以下のユーザ、もしくは被評価数が 5 件以下のアイテムのデータは取り除いた。評価件数が 3 件以上のユーザが最後に購入したアイテム $S_{|S^u|}^u$ とその前に購入したアイテム $S_{|S^u|-1}^u$ をテストデータ、 $S_1^u, S_2^u, \dots, S_{|S^u|-2}^u$ を学習データとした。パラメータである次元数は全ての手法において $K = 50$ とした。実験はユーザをランダムに 5 つのデータセットに振り分け、それぞれのデータセットにおいて評価指標を算出し、その平均値で評価を行う 5 分割交差検定を用いる。また、評価指標には以下の Area Under the ROC Curve (AUC) と Hit@50 を用いる。

AUC :

$$AUC = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \frac{1}{|\mathcal{I} \setminus S^u|} \sum_{j' \in \mathcal{I} \setminus S^u} 1(R_{u,g_u} < R_{u,f'}) \quad (12)$$

Hit@50 :

$$Hit@50 = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} 1(R_{u,g_u} \leq 50) \quad (13)$$

g_u は実際にユーザが最後に購入したアイテムであり、 $R_{u,i}$ はユーザ u のアイテム i に対する予測購買確率のランキング（値が小さいほど良い）、 $1(b)$ はインジケータ関数であり、 b が真であれば 1、そうでなければ 0 を返す。

4.2. 実験結果

実験結果を以下に示す。

表 1. テストデータ全体に対する結果

| | TransRec | 提案手法 |
|--------|----------|--------|
| AUC | 0.6784 | 0.7290 |
| Hit@50 | 0.0389 | 0.0812 |

表 1 より、提案手法は、AUC と Hit@50 の両方で最も良い結果を示していることが分かる。これより、電子商店街における購買履歴データに TransRec を適用する場合、商店情報を考慮する方が良いことがわかる。また、商店を直線で表現できていることがわかる。

5. 実データ分析

本研究では、同一の商店において購買を続けるユーザを考慮するために商店を直線で表現した。これに対して、ユーザ個人単位で見のではなく空間全体で見た時、商店がベクトルによる直線で表現されているため、一部のアイテムだけでも商店間で関係があれば、その商店の直線同士の最短距離は近くなると考えられる。例えば、寝具とアクセサリーを売っている商店があった場合、その商店に対して寝具専門店やアクセサリー専門店が近づくといったように、異なるカテゴリのアイテムを販売している商店同士でも、一部で似たようなアイテムを販売しているならば、その距離は近くなるという関係性があると考えられる。そこで商店を表現する直線を分析することで、その関係性が現れているか確認する。

具体的には、ある商店に着目し、その商店が実際に販売しているアイテムのベクトル \vec{r}_j と全ユーザに共通する商店が与えられたもとのアイテムに対する嗜好ベクトル \vec{t}^i の逆ベクトルの和の距離 $d_{j,l}$ が短い商店 l を列挙する。 $d_{j,l}$ は以下のように定義される。

$$d_{j,l} = d(\vec{r}_j - \vec{t}^i, \alpha \vec{m}_l + \vec{n}_l) \quad (14)$$

このときの α は以下のように定義される。

$$\alpha = \frac{-\sum_k \vec{m}_k \cdot (\vec{n}_l - (\vec{r}_j - \vec{t}^i))}{\sum_k \vec{m}_k^2} \quad (15)$$

着目する商店はネックレスや靴が主に売られている商店とする。着目するアイテムはネックレス A, 靴 A とする。

表 2. ネckレス A に近い TOP10 の商店

| | |
|----|--------------------------|
| 1 | 対象の商店 |
| 2 | スポーツ用品店 A |
| 3 | アクセサリーや靴、服を販売しているアパレル店 A |
| 4 | 雑貨店 A |
| 5 | アクセサリーや靴、服を販売しているアパレル店 B |
| 6 | 靴や靴、アクセサリーを販売している店 A |
| 7 | アクセサリーや雑貨を販売している店 A |
| 8 | アクセサリーや靴などを販売している雑貨店 A |
| 9 | アクセサリーや服を販売しているアパレル店 C |
| 10 | ペット用品店 A |

表 2 より、対象の商店と同様にネックレスが売られている商店が多く出てきたことがわかる。これより、同様の商店が多く出現したことから、商店の学習はある程度うまくできていることが分かる。

また、ネックレスや靴と距離の近い商店 TOP20 の中で、ネックレスか靴のどちらかのみを扱っている商店を列挙した。

表 3. ネckレスか靴のどちらかのみを扱っている商店

| | |
|---|------------------------|
| 1 | アクセサリーや雑貨を販売している店 A |
| 2 | アクセサリーや服を販売しているアパレル店 C |
| 3 | 服や靴を販売しているアパレル店 E |
| 4 | アクセサリーも販売している天然石販売店 A |
| 5 | 服や靴を販売しているアパレル店 G |
| 6 | 服や靴を販売しているアパレル店 E |
| 7 | 靴も売っているスポーツ用品店 B |

表 3 より、アクセサリーも販売している天然石販売店 A や靴も売っているスポーツ用品店 B などが抽出されていることが分かる。このように、異なるカテゴリのアイテムを販売している商店同士でも、一部で似たようなアイテムを販売しているならば、その距離が近づくように学習できていることが分かる。

6. 考察

実データを用いた実験により、電子商店街の購買履歴データに対して提案手法が有効であることを示した。一方で、提案手法は従来手法と比較してパラメータ数が多いため、データ数が少ないユーザに対しては有効でない可能性がある。そのため、データ数が少ないユーザも考慮したモデルに改良することでさらなる精度の向上が可能であると考えられる。

7. まとめ

本研究では、従来手法である TransRec に直線で表現される商店とある商店でアイテムを選択するユーザの嗜好をモデルに加えた。そして、商店ベクトルを点ではなく線で表現することにより、電子商店街におけるユーザの嗜好を表現したモデルを提案した。さらに、電子商店街の実データを用いた実験を通じ、提案手法の有効性を示すことができた。

本研究ではある商店においてアイテムを選択するユーザの嗜好を追加したが、さらにあるアイテムを購入した後に商店を選択するユーザの嗜好を加えることで、より電子商店街におけるユーザの嗜好が十分に表現され予測精度が向上すると考えられる。

謝辞

本研究で用いた楽天市場データセットは国立情報学研究所の IDR データセット提供サービスにより楽天株式会社様からご提供いただきました。ご協力に心より感謝申し上げます。参考文献

- [1] Ruining He, Wang-Cheng Kang, and Julian McAuley, "Translation-based Recommendation," *Eleventh ACM Conference on Recommender Systems (RecSys '17)*. ACM, New York, NY, USA, pp.161-169, 2017.
- [2] Vitali Kuzmin, "Item2Vec-based Approach to a Recommender System," *University of Tarutu*, (2017).
- [3] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme, "Factorizing personalized markov chains for next-basket recommendation," *International Conference on World Wide Web (WWW)*, pp.811-820, 2010.
- [4] 楽天株式会社 (2014): 楽天市場データ. 国立情報学研究所情報学研究データリポジトリ. (データセット). <https://doi.org/10.32130/idr.2.1>