

## 投票者モデルに基づくレビュー回数ユーザ重み影響度分析 Influence Analysis of users' review frequency weight based on Voter Model

鈴木 優人<sup>†</sup>  
Yuto Suzuki

斉藤 和巳<sup>†</sup>  
Kazumi Saito

### 1. はじめに

近年、商品やサービスについてのレビューを投稿できるレビューサイトを利用するユーザが増えており盛んに活用され、レビューサイト分析や研究が多く展開されている [Salganik 06]. これらレビューサイトにおけるユーザのレビューは、他のユーザの購買活動に強く影響を与えるものであり、レビューサイトは重要なメディアといえる。その中でも、他のユーザに影響を与えるユーザ抽出技術は重要な研究課題である。これまでの我々の研究では、6つのレビューサイトを用いてユーザのスコアを計算し、影響力のあるユーザを分析した。分析結果より、ユーザごとに影響度が異なるとしたユーザモデルと、レビュー順序ごとに影響度が異なる傾向を示すことを確認した [Suzuki 16].

そこで本研究では、違う角度からのユーザ重みの重要性を検証する。そこでレビュー回数が同程度のユーザは同程度のユーザ重みと仮定し、新ユーザモデルを構築する。構築モデルの有効性は、投票者モデル [Even-Dar 07] に基づき、6つのレビューサイトを用いて検証する。

### 2. 提案指標

ユーザ集合を  $U = \{u, v, \dots\}$ 、アイテム集合を  $I = \{i, \dots\}$  とし、ユーザ  $u$  のレビューしたアイテム集合を  $I(u) \subset I$ 、アイテム  $i$  にレビューしたユーザ集合を  $U(i) \subset U$  とする。また、ユーザ集合  $U$  において、レビュー回数  $k \in \{1, 2, \dots\}$  で同じユーザを同一視した仮想ユーザ集合  $V_k = \{u \in U : |I(u)| = k\}$  とする。

レビュー影響の算出には、以下の2つ方法を用いる。第一の方法では、各ユーザの重みをパラメータ  $X = \{x_u, \dots\}$  として設定し、アイテム  $i$  にユーザ  $u$  が評点  $s$  を与える確率をモデル化し、パラメータ集合  $X$  を最尤推定で求める。すなわち、アイテム  $i$  にユーザ  $u$  より前にレビューしたユーザ集合を  $U(i, u)$  とし、 $U(i, u)$  の中で評点  $s$  を与えたユーザ集合を  $U(i, u, s) \subset U(i, u)$  とすると、アイテム  $i$  にユーザ  $u$  が評点  $s$  を与える確率は次のようにモデル化できる。

$$P(i, u, s; X) = \frac{\sum_{v \in U(i, u, s)} x_v}{\sum_{v \in U(i, u)} x_v} \quad (1)$$

そして、収集したレビューデータ  $D$  で、実際にアイテム  $i$  に対しユーザ  $u$  が評点  $s$  を付けたことより、次式の尤度  $L(X)$  を最大化するパラメータ  $X$  を求める。

$$L(X) = \sum_{(i, u, s(u, i)) \in D} \log P(i, u, s(i, u); X) - \frac{r}{2} \sum_{u \in U} \log(x_u)^2 \quad (2)$$

<sup>†</sup>静岡県立大学 経営情報学部

ここで  $r$  は正則化係数を表し、特に過学習を防ぐために用いる。  $L(X)$  の最大化については、任意のユーザ  $u$  のパラメータを  $x_u = 1$  として初期値  $X^{(0)}$  を設定し、EM アルゴリズムの第  $n$  反復目での尤度の改善が  $L(X^{(n)}) < (1 + \epsilon)L(X^{(n-1)})$  となるまで学習させた。なお実験では、 $\epsilon = 0.0001$  に設定した。以下、この方法はユーザ重み付け法 (user 法) とする。

第二の方法では、ユーザ  $k \in U$  のレビュー回数を  $k(u) = k$  で表し、各  $k$  ごとにパラメータ  $y_k$  を設定すると重みパラメータ  $Y = \{y_1, \dots, y_k, \dots\}$  とすると、アイテム  $i$  にユーザ  $u$  が評点  $s$  を与える確率は次のようにモデル化できる。

$$P(i, u, s; Y) = \frac{\sum_{v \in U(i, u, s)} y_{k(v)}}{\sum_{v \in U(i, u)} y_{k(v)}} \quad (3)$$

そして、収集したレビューデータ  $D$  で、実際にアイテム  $i$  に対しユーザ  $u$  が評点  $s$  を付けたことより、次式の尤度  $L(Y)$  を最大化するパラメータ  $Y$  を求める。

$$L(Y) = \sum_{(i, u, s(u, i)) \in D} \log P(i, u, s(i, u); Y) - \frac{r}{2} \sum_{u \in U} \log(y_{k(u)})^2 \quad (4)$$

以下この手法は、レビュー回数ごとのユーザ重み付け法 (新 user 法) とする。

### 3. 実験設定

実験には6つのレビューデータを用いる。アニコレは、アニメを題材としたランキング、口コミサイトであり、1,790 アイテム、13,111 ユーザー、300,327 レビューを有する。@cosme は、化粧品をアイテムとしたレビューサイトであり、46,398 アイテム、10,403 ユーザ、297,453 レビューを有する。ブックログは、ブックレビューサイトであり、178,432 アイテム、157,325 ユーザー、2,752,817 レビューを有する。食べログは、全国飲食店の口コミ、ランキングサイトであり、443,573 アイテム、291,052 ユーザー、3,114,507 レビューを有する。価格.com は、あらゆる製品を比較検討できる買い物支援サイトであり、75,519 アイテム、170,645 ユーザー、475,327 レビューを有する。MovieLens は、映画の評価がまとめられており、10,677 アイテム、69,878 ユーザー、10,000,054 レビューを有する。表1にこれらレビューサイトのデータ概要のまとめを示す。

### 4. 実験結果

図1のグラフは、横軸に反復回数、縦軸に尤度をプロットした実験結果を示す。(a) アニコレ (b) @cosme (c) ブックログ (d) 食べログ (e) 価格.com (f) MovieLens の結果を示す。u1 はユーザー一人ひとりに重み付けをしたユーザ法、u2 がレビュー回数ごとにユーザを振り分けし重み付けをした新ユーザ法である。(a) アニコレ、(e)

表 1: レビューサイト詳細表

	アニコレ	@cosme	ブックログ	食べログ	価格.com	MovieLens
アイテム数	1,790	46,398	178,432	443,573	75,519	10,677
ユーザ数	13,111	10,403	157,325	291,052	170,645	69,878
レビュー数	300,327	297,453	2,752,817	3,114,507	475,327	10,000,054

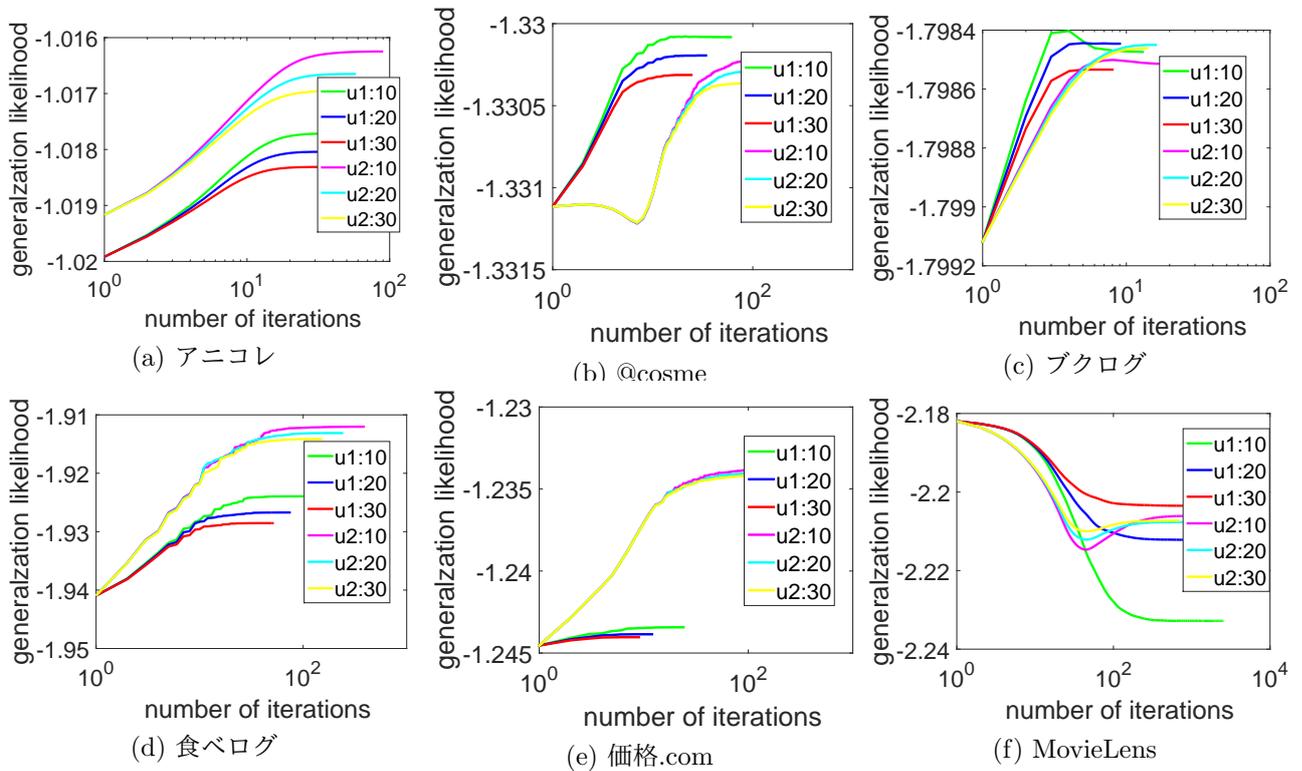


図 1: 反復回数と尤度による各レビューサイトの評価結果

価格.com, (d) 食べログでは新ユーザ法が有意な結果を示した。(b)@cosme, (c) ブックログ, (f)Movie Lens10kでは同程度の結果を示した。特に (e) 価格.com では、レビュー回数と同じユーザは同じくらいの影響度を持つ、と仮定する我々のモデルと最も親和性が高かったと考えられる。(f)MovieLens10kでは既存のユーザ法、新ユーザ法ともにレビュー評点の予測性能が低かった。これは (f) は完全なアンケートでのデータで、他のユーザーの評価は見えない。そのためこの結果は不自然ではないと考えられる。

### 5. おわりに

本研究では、投票者モデルに基づき、6つのレビューサイトを用いてユーザのスコアを計算し、レビューサイトにおいて影響力のあるユーザを分析した。分析結果より、今回調査した6つのレビューサイト全てにおいて新ユーザ法が従来のユーザ法より精度が悪くならなかった。これにより新ユーザ法は有効性を示した。今後は多様なレビューサイトを用いて、活動レベルなどユーザー属性と影響度との関係分析なども進めたい。

謝辞 本研究は、科学研究費補助金基盤研究(C)(No.15K00429)の助成を受けた。

### 参考文献

[Salganik 06] M.J.Salganik, P.S.Dodds, and D.J.Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market", Science 10, pp.854-856, February 2006, (2006).

[Even-Dar 07] Even-Dar.E. and Shapira.A, "A note on maximizing the spread of influence in social networks", In Proc. of WINE 2007, pp.281-286, 2007, (2007) .

[Suzuki 16] 鈴木 優人, 齊藤 和巳, "投票者モデルに基づく複数レビューサイトでの影響度分析", 第15回情報科学技術フォーラム (FIT2016)