

予測評価値の信頼性を考慮した Hybrid-NMF の提案

1X12C016-0 井原龍太郎
指導教員 後藤正幸

1 研究背景・目的

近年、EC サイトで扱われるアイテム数の増加に伴い、ユーザの嗜好も多様化する一方である。そのため、有効な Web マーケティングツールとして、各ユーザの嗜好に合った商品を自動で推薦するシステムの重要性が高まっている。代表的な推薦システムには、類似ユーザの評価履歴からユーザの未評価アイテムに対する予測評価値を算出し、推薦を行う協調フィルタリング (以下、CF) があるが、評価値の予測モデルの精度をいかに高められるかが肝要である。そのため、本研究では行列分解に基づく手法である Non-negative Matrix Factorization (以下、NMF) に着目する。NMF は評価値行列を二つの行列に分解し、その積により未評価要素の予測を行う手法であり、様々な適用例において、その有効性が示されている。

NMF に関しては、評価要素のみに対して重み付けを行うことによって、精度向上を図る Weighted-NMF (以下、WNMF) という改良手法があるが、精度が初期値に大きく依存することが知られている。他方、未評価要素に予測評価値を補完しながら学習を行う EM-NMF という手法も提案されている。これらの手法は、WNMF の初期値依存の問題を、EM-NMF で求めた行列を初期値として用いることで解決する Hybrid-NMF において融合され、その有効性が示されている。

しかし、Hybrid-NMF ではユーザによって評価傾向の推測のし易さが異なり、期待される EM-NMF の予測精度は異なるにも関わらず、WNMF フェーズで全てのユーザの評価項目に対して同様の重み付けを行っている。そのため、ユーザ毎に適切な重み付けを行うことで、さらなる精度向上の可能性がある。そこで本研究では従来の Hybrid-NMF とは異なり、EM-NMF で算出したユーザ毎の期待される予測精度を考慮した、重み付けを行う手法を提案し予測精度の向上を目指す。また、ベンチマークデータを用いた評価実験によってその有効性を示す。

2 推薦システム

推薦システムとはユーザの評価履歴からユーザの嗜好を特定し、アイテムの推薦を行う手法である。いま、アイテム集合を $\mathcal{I} = \{I_i : 1 \leq i \leq N\}$ 、ユーザ集合を $\mathcal{J} = \{J_j : 1 \leq j \leq M\}$ と定義する。アイテム I_i に対し、ユーザ J_j が Y 段階評価で y 点の評価をした場合は y を要素とし、未評価の場合は欠損とするアイテム・ユーザ行列を $A = [a_{ij}] \in \mathcal{R}^{N \times M}$ と定義する。何らかのアルゴリズムを適用することで欠損である未評価アイテムの評価値を補完し、その値が高いアイテムを各ユーザに推薦する。

3 従来手法

3.1 NMF

NMF は評価値行列 A を二つの行列 U, V に分解し、その積により未評価要素の予測を行う手法である。いま、アイテムのジャンルなどを示す潜在クラスタの集合を $\mathcal{K} = \{K_k : 1 \leq k \leq L\}$ とし、アイテム I_i が潜在クラスタ K_k に所属する度合いを要素に持つ行列 $U = [u_{ik}] \in \mathcal{R}^{N \times L}$ とし、ユーザ J_j のクラスタ k に所属する度合いを要素に持つ行列 $V = [v_{kj}] \in \mathcal{R}^{L \times M}$ と定義する。また、 U と V の積で算出される行列を $X = [x_{ij}] \in \mathcal{R}^{N \times M}$ で表す。NMF は X と A のフロベニウスノルムの 2 乗の差を最小化するように U, V の更新を行い、最終的に求めた X を予測評価値行列とする。

3.2 EM-NMF

EM-NMF は、更新により得られた予測評価値を未評価要素に補完する手法である。この手法は尤度関数が収束するまでパラメータを繰り返し更新するため、計算時間が多くかかるが、予測精度が高いことが知られている。いま、行列 A の評価要素の集合を \mathcal{A}^o 、未評価要素の集合を \mathcal{A}^u と定義する。また、 t 回目の更新ごとに得られた予測評価値を行列 A の未評価要素に補完した行列を $A^{(t)}$ と定義する。 $A^{(t)}$ の要素は $a_{ij} \in \mathcal{A}^o$ に該当する要素は定数とし、 $a_{ij} \in \mathcal{A}^u$ に該当する要素は EM アルゴリズムの $t-1$ 回目の更新により算出された行列 $X^{(t-1)}$ の要素 $x_{ij}^{(t-1)}$ とする。すなわち、 $A^{(t)}$ の要素は以下の式で表される。

$$a^{(t)}_{ij} = \begin{cases} a_{ij} & (a_{ij} \in \mathcal{A}^o) \\ x_{ij}^{(t-1)} & (a_{ij} \in \mathcal{A}^u) \end{cases} \quad (1)$$

3.3 WNMF

WNMF は評価要素のみに着目して学習を行う手法である。この手法は計算に要する時間は短い³、予測精度が初期値に大きく依存することが知られている。WNMF は、以下の式 (2) で定義された最適化問題を解くことで行列 U, V の学習を行う。行列 $W = [w_{ij}] \in \mathcal{R}^{N \times M}$ はアイテム i のユーザ j による評価要素に対する重み行列であり、式 (3) により定義される。 w_{ij} が 1 の場合は学習に用いることを表し、0 の場合は学習に用いないことを表す。

$$\min_{U, V} \|W * (A - UV)\|_F^2 \quad (2)$$

$$w_{ij} = \begin{cases} 1 & (a_{ij} \in \mathcal{A}^o) \\ 0 & (a_{ij} \in \mathcal{A}^u) \end{cases} \quad (3)$$

ただし、式 (2) における $*$ は行列の要素ごとの積、 $\|\cdot\|_F$ は行列のフロベニウスノルムを表す。式 (2) の最適化問題は、以下の式 (4)、(5) を用いて U, V を収束するまで繰り返し更新することで解くことができる。

$$u_{ik}^{(t)} = u_{ik}^{(t-1)} \frac{\left((W * A^{(t-1)}) V^{(t-1)T} \right)_{ik}}{\left((W * U^{(t-1)} V^{(t-1)T}) V^{(t-1)T} \right)_{ik}} \quad (4)$$

$$v_{kj}^{(t)} = v_{kj}^{(t-1)} \frac{\left(U^{(t-1)T} (W * A^{(t-1)T}) \right)_{kj}}{\left(U^{(t-1)T} (W * U^{(t-1)T} V^{(t-1)}) \right)_{kj}} \quad (5)$$

3.4 Hybrid-NMF

Hybrid-NMF は、計算量を要する EM-NMF の更新を最初の数回のみ行い、得られた予測評価値を WNMF の初期値として用いることで、2 つの手法と比較して計算時間が短く、かつ精度が初期値に依存しづらい評価予測を可能にする手法である。すなわち、EM-NMF の更新を t_{EM} 回繰り返して得られた U, V を初期値として、WNMF を実行するという 2 つのフェーズにより評価予測を行う。

4 提案手法

4.1 着眼点

Hybrid-NMF では EM-NMF フェーズで得られた予測評価値においてユーザによって評価傾向の推測のし易さが異なり、期待される予測精度が異なるにも関わらず、WNMF フェーズで全ユーザの評価要素に対して同様の重みを与えている。すなわち、期待される予測精度が高いユーザを重点的に学習に用いた方が予測精度は向上すると考えられるにも関わらず、予測精度が低いユーザも等しく学習に用いている。

そのため、EM-NMF フェーズでユーザ毎の予測評価値の信頼性を数値化し、その値が高いユーザを重点的に学習することができれば、より予測精度を向上できる可能性がある。

信頼性の基準を考える際、EM-NMF 終了後に算出した X は学習データ A を近似した行列であることに着目する。すなわち、 X がどの程度 A に近似できたかを測定することによって、近似精度が高いユーザほど信頼性が高いと仮定し、そのユーザを重点的に学習するために、式 (3) の行列 W の要素を $\{0, 1\}$ ではなく重要度に合わせた連続値にする方法を提案する。

4.2 重み付けの基準・方法

ユーザ毎の信頼性を算出する基準として、EM-NMF フェーズ終了後に算出した X と行列 A の評価要素の平均 2 乗誤差の値を使用する。この値は、以下の式 (6) で示されるユーザ毎の評価要素の 2 乗誤差の和をユーザ毎の評価要素数で割った値であり、小さいほど、そのユーザの信頼性は高いとする。

$$\alpha_j = \frac{\sum_i \rho_{ij} (a_{ij} - x_{ij})^2}{\sum_i \rho_{ij}} \quad (6)$$

$$\rho_{ij} = \begin{cases} 1 & (a_{ij} \in \mathcal{A}^o) \\ 0 & (a_{ij} \in \mathcal{A}^u) \end{cases} \quad (7)$$

信頼性の高いユーザを重視して WNMF で学習するため、最適化問題でも用いた平均 2 乗誤差の値が全体で小さい方から $\beta (0 \leq \beta \leq 1)$ の割合のユーザの評価値の要素集合を \mathcal{A}^{o1} 、 $(1 - \beta)$ の割合のユーザの評価値の要素集合を \mathcal{A}^{o2} と定義し、異なる重みを与えることを考える。すなわち、信頼性を考慮した重み行列 $W^p = [w_{ij}^p] \in R^{N \times M}$ を式 (8) により定義する。

$$w_{ij}^p = \begin{cases} \gamma & (a_{ij} \in \mathcal{A}^{o1}) \\ 1 & (a_{ij} \in \mathcal{A}^{o2}) \\ 0 & (a_{ij} \in \mathcal{A}^u) \end{cases} \quad (8)$$

ただし t_{EM} を用いて、以下の式のように γ を定義する。

$$\gamma = 0.2 \times e^{-\frac{1}{4} \log t_{EM}} + 1 \quad (9)$$

4.3 アルゴリズム

提案手法のアルゴリズムを以下に示す。

Step1) EM-NMF の更新を t_{EM} 回行い、行列 $U^{(t_{EM})}, V^{(t_{EM})}$ を得る。

Step2) 式 (6) より行列 $X^{(t_{EM})}$ と行列 A の評価要素のユーザ毎の平均 2 乗誤差を算出する。

Step3) 式 (9) を用いて重み γ の計算を行い、 W^p を式 (8) により導出する。

Step4) Step3 で得られた W^p を用いて WNMF を行い、予測評価値の高いアイテムをユーザに推薦する。 □

5 実験

提案手法の有効性を示すため、推薦システムのベンチマークデータを用いた評価実験を行った。

5.1 実験条件

実験では、公開データセット MovieLens の映画評価データ 10 万件を用いた。ユーザ数 $M = 943$ 、アイテム数 $N = 1682$ 、学習データを 8 万件、テストデータを 2 万件とした。まず、従来研究で最も予測精度が高いと示されている $t_{EM} = 5$ で β を 0.1 から 1 まで変化させて、最適な β を算出した。また、更新回数による提案手法と従来手法の精度の差の変化について考察を行うために t_{EM} の値を変化させ、予測精度の比較を行う。評価指標としては式 (10) で定義されるテストデータと予測評価値の MAE (平均絶対誤差) を用いた。

$$MAE = \frac{1}{D} \sum_{j=1}^M \sum_{i=1}^N |T_{ij} - \hat{x}_{ij}| \delta_{T_{ij}} \quad (10)$$

ただし、式 (10) における T_{ij} はユーザ j がアイテム i に対してつけた評価値、 \hat{x}_{ij} は得られた予測評価値、 D はテストデータ数である。また、 $\delta_{T_{ij}}$ は T_{ij} がテストデータの場合は 1、それ以外は 0 を示すインジケータ関数とする。上記条件で実験を 3 回行い、その平均値で予測精度の比較を行う。

5.2 実験結果と考察

$t_{EM} = 5$ において、 β の値を 0.1 から 1.0 まで変化させて提案手法を用いた結果を図 1 に示す。図 1 より $\beta = 0.9$ の時に最も高い予測精度が得られ、 β の値が小さくなるにつれて予測精度は低下した。また $\beta = 0.4 \sim 1.0$ の時に提案手法が従来手法を上回った。この結果より、従来の Hybrid-NMF の予測精度を低下させるのは信頼性が低いごく一部のユーザが原因であることが考えられる。また、 β の値が 1 の時にも予測精度が低下したことから、EM-NMF において信頼性が高いとされたユーザのみに 1 以上の重み付けを行い、重点的に学習に用いることで精度向上が達成されると考えられる。

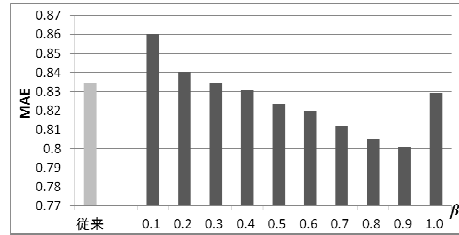


図 1. β と MAE の関係

次に上記の実験で得られた結果から β を 0.9 とし、従来手法と提案手法における t_{EM} を 1, 3, 5, 7 と変化させたときの結果を図 2 に示す。図 2 から、すべての t_{EM} において提案手法が従来手法を上回っていることがわかる。また、 t_{EM} の値が大きくなるにつれて従来手法と提案手法共に予測精度が高くなっている。これは EM-NMF で算出された初期値の精度が向上するためであると考えられる。

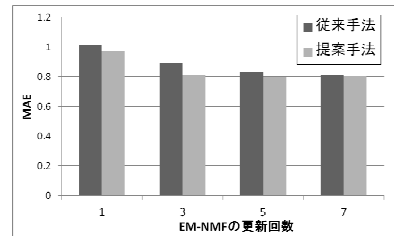


図 2. $t_{EM} = 1, 3, 5, 7$ の時の従来手法と提案手法の比較

上記の図 1, 図 2 の結果より、EM-NMF において信頼性が高いとされたユーザに対して 1 以上の重み付けを行い、重点的に学習することの有効性が示された。

6 まとめと今後の課題

本研究では Hybrid-NMF において EM-NMF フェーズで得られた信頼性が高いユーザの初期値に対して、WNMF フェーズで 1 以上の重み付けを行うことにより、重点的に学習する手法を提案した。また、ベンチマークデータを用いた実験により提案手法の有効性が示された。

今後の課題として、信頼性が高いユーザに対して一定の 1 以上の重み付けをするのではなく、各ユーザの信頼性の度合いを考慮した重み付けを行う手法の検討が挙げられる。

参考文献

- [1] Zhang S., Wang W., Ford J., Makedon F., "Learning from Incomplete Ratings Using Non-negative Matrix Factorization," *6th SIAM Conference on Data Mining(SDM)*, pp.549-553, 2006.