

ユーザの嗜好を考慮した非負値行列因子分解の提案

1X13C048-8 齊藤 央樹
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が増加し、ユーザの嗜好も多様化している。そのため、対象ユーザの嗜好に適したアイテムを自動的に推薦するシステムの重要性が高まっている。そのような推薦システムの中に、ユーザとアイテムの評価履歴データに基づいて未評価アイテムに対する評価値を予測し、推薦を行う協調フィルタリングがある。モデルベースでの協調フィルタリングには様々な手法があるが、本研究では、行列分解に基づく手法である EM Non-negative Matrix Factorization[1](以下、EM-NMF)に着目する。EM-NMF とは、アイテムとユーザがそれぞれ共通のクラスタに所属すると考え、それらへの所属度合いを要素に持つ 2 つの低次元非負値行列に評価値行列を分解し、分解後の行列の積により評価値を近似し予測を行う手法である。しかし一般に評価値行列は、被評価数の多い少数の人気アイテムと、大多数の被評価数が少ないアイテムから構成されているという特徴を持つ。そのため、EM-NMF では評価データ数の多い人気アイテムの情報に適応するように行列分解が行われ、人気アイテムを嗜好するユーザに対しては適切な推薦を行うことができる。一方で、被評価数の少ないアイテムを嗜好するユーザに対する予測評価値は、相対的に予測精度が低くなってしまいう問題がある。

そこで本研究では、評価値行列全体から推定した予測値とユーザの嗜好を基に分割した部分評価値行列から推定した予測値をアンサンブルする手法を提案し、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り本研究の提案手法の有効性を示す。

2 推薦システム

推薦システムとは、ユーザの評価履歴や閲覧履歴などの情報からユーザが興味を持つと思われる情報を予測し、推薦対象ユーザに情報の推薦を行う手法である。いま、 N 個からなるアイテム集合を $\mathcal{I} = \{I_i : 1 \leq i \leq N\}$ 、 M 人からなるユーザ集合を $\mathcal{J} = \{J_j : 1 \leq j \leq M\}$ 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 $a_{ij} = y$ を要素、未評価の場合の a_{ij} は欠損とする評価値行列を $A = [a_{ij}] \in \mathcal{R}^{N \times M}$ と定義する。評価値行列 A に対し、適当なアルゴリズムを適用することで未評価であるアイテムの評価値を予測し、その値が高いアイテムを対象ユーザに推薦する。

3 従来手法

3.1 EM-NMF

EM-NMF とは評価値行列 A を 2 つの低次元かつ非負の特徴行列の積で近似し、未評価要素の予測を行う手法である。いま、この 2 つの行列をそれぞれアイテム I_i がクラスタ $k(1 \leq k \leq K)$ に所属する度合いを要素に持つ行列を $U = [u_{ik}] \in \mathcal{R}^{N \times K}$ 、ユーザ J_j がクラスタ k に所属する度合いを要素に持つ行列を $V = [v_{kj}] \in \mathcal{R}^{K \times M}$ とし、 U, V の積によって算出される評価値行列の近似行列を $X = [X_{ij}] \in \mathcal{R}^{N \times M}$ と定義する。EM-NMF の学習では、 A' が更新される都度、式 (1) を最小化するように行列 U, V を更新する。なお、 $\|\cdot\|_F$ はフロベニウスノルムを表す。

$$\min_{U, V} \|A' - UV\|_F^2 \quad (1)$$

$$s.t. \forall u_{ik} \geq 0, \forall v_{kj} \geq 0$$

いま、評価値行列 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。式 (1) の解を得るために、 $a_{ij} \in \mathcal{A}^o$ に該当する要素は評価値で固定し、 $a_{ij} \in \mathcal{A}^u$ に該当する要素を繰り返し演算により更新して求める。具体的には、 t 回目の更新により算出された行列 X の値を $x_{ij}^{(t)}$ 、行列 A の未評

価要素に $x_{ij}^{(t-1)}$ を代入した行列を $A'^{(t)} = [a'_{ij}{}^{(t)}] \in \mathcal{R}^{N \times M}$ とすると、次の式 (2)~(4) を用いて行列 A' と U, V を繰り返し更新することで式 (1) の問題を解くことができる。行列 A', U, V, X について、 t 回目の更新後の各行列を $A'^{(t)}, U^{(t)}, V^{(t)}, X^{(t)}$ とし、および各行列の (i, j) 要素を $a'_{ij}{}^{(t)}, u_{ik}^{(t)}, v_{kj}^{(t)}, x_{ij}^{(t)}$ とする。また $t = 0$ は各行列の初期値を表す。

$$u_{ik}^{(t)} \leftarrow u_{ik}^{(t-1)} \frac{(A'^{(t-1)} V^{(t-1)T})_{ik}}{(U^{(t-1)} V^{(t-1)} V^{(t-1)T})_{ik}} \quad (2)$$

$$v_{kj}^{(t)} \leftarrow v_{kj}^{(t-1)} \frac{(U^{(t-1)T} X^{(t-1)})_{kj}}{(U^{(t-1)T} U^{(t-1)} V^{(t-1)})_{kj}} \quad (3)$$

$$a'_{ij}{}^{(t)} = \begin{cases} a_{ij} & (a_{ij} \in \mathcal{A}^o) \\ x_{ij}^{(t-1)} & (a_{ij} \in \mathcal{A}^u) \end{cases} \quad (4)$$

以上の式より求めた取束後の行列 $A'^{(T)}$ の値を用いて、対象ユーザに推薦を行う。

3.2 一般的な推薦システムにおける問題点

以下に公開データセット MovieLens[2] の映画評価データ 100 万件から抽出した評価実験において用いる学習データ 80 万件に対して被評価数のヒストグラムを図 1 に示した。

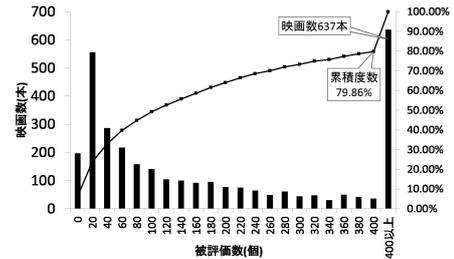


図 1: 映画被評価数のヒストグラム

この図 1 に示されるとおり、一般的な評価履歴データでは、全アイテムのうち、そのほとんどが被評価数の少ないアイテムによって占められる分布になりやすい。

一般的に、被評価数の多いアイテムと被評価数の少ないアイテムでは評価傾向が異なると考えられる。前者では、そのアイテムに精通していない多様なユーザからの評価を受けやすく、ユーザごとの嗜好が大きく異なると考えられる。一方後者では、そのアイテムに精通した一部の熱狂的なユーザからの評価を受けていることも考えられる。

従って、一般的な推薦システムにおいては、特性の異なるアイテム同士をまとめてモデル化してしまい、両者の特性を反映していないモデルが構築されている可能性がある。

3.3 EM-NMF における問題点

図 1 のような被評価数に大きな偏りのあるデータに対する EM-NMF による学習は、被評価数の多いデータが重視されて行列分解の大半が決定されてしまうため、被評価数の少ないアイテムに対しては相対的に予測精度が悪くなるという問題がある。いま、先に示した図 1 から、被評価数上位 2 割 (被評価数 400 件以上) の 637 アイテムを MAJOR 映画、それ以外の 2,525 アイテムを MINOR 映画と定義する。以下の表 1 に MAJOR 映画と MINOR 映画の予測評価値の平均絶対誤差 (以下、MAE) の計算結果を示す。なお、MAE は式 (5) で定義される。

$$MAE = \frac{1}{D} \sum_{j=1}^M \sum_{i=1}^N |T_{ij} - \hat{x}_{ij}| \delta_{ij} \quad (5)$$

ただし、 T_{ij} はテストデータ中のユーザ J_j がアイテム I_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、 δ_{ij} は T_{ij} がテストデータに存在する場合は 1, それ以外は 0 を示すインジケータ関数とする。

表 1: アイテム区分における MAE

	MAJOR 映画	MINOR 映画
MAE	0.758	0.823

表 1 より, 被評価数の少ない MINOR 映画の MAE は, 被評価数の多い MAJOR 映画と比べ劣っていることが分かる。すなわち, MAJOR 映画を好むユーザと比較して, MINOR 映画を好むユーザへの推定精度も悪いといえる。

4 提案手法

4.1 着眼点

上記の問題を解消するため, MAJOR 映画と MINOR 映画といったアイテムの被評価数の観点から学習データを分割してモデル化することで, それぞれのアイテムの特性を考慮することを考える。一方で, MAJOR 映画と MINOR 映画の両方を好むユーザが存在することも考えられるが, 単純にアイテムの人気度の観点からデータを分割すると, そのようなユーザの嗜好を考慮することができない。さらに, 個々のモデルの学習データが乏しくなってしまう, 推定精度の低下も懸念される。そこで本研究では, 全学習データで行列分解を行うモデルをベースとし, これにユーザの嗜好ごとに別々に推定したモデルを適切にアンサンブルする手法を提案する。これにより, 十分なデータ数を確保しながら, 被評価数の少ないアイテムからもモデルを構築することができ, 予測精度が向上すると考えられる。

4.2 実行手順

ユーザ数を M , アイテム数を N とする。ユーザの分割を, 被評価の多いアイテムを多く評価するユーザを MAJOR ユーザ, 被評価数の少ないアイテムを多く評価するユーザを MINOR ユーザとし, ユーザを分割した行列をそれぞれ $\bar{A} (\in \mathcal{R}^{N \times \bar{M}})$, $\underline{A} (\in \mathcal{R}^{N \times \underline{M}})$ とする。ただし, \bar{M} は MAJOR 映画を嗜好するユーザ数, \underline{M} は MINOR 映画を嗜好するユーザ数とし, すなわち, $\bar{M} = M - \underline{M}$ である。また, 分割した行列の (i, j) 要素を, それぞれ \bar{a}_{ij} , \underline{a}_{ij} とする。そして, \bar{A} , \underline{A} それぞれに対し EM-NMF を行い, 予測値を混合することによって, 評価値を予測し, アイテムの推薦を行う。分割した各予測評価値行列を \bar{A}' , \underline{A}' とし, それらの行列の (i, j) 要素をそれぞれ \bar{a}'_{ij} , \underline{a}'_{ij} とする。その際のパラメータを MAJOR ユーザの混合比率 $\alpha (0 \leq \alpha \leq 1)$, MINOR ユーザの混合比率 $\beta (0 \leq \beta \leq 1)$ とする。以下の式 (6) より, 予測評価値行列 $\hat{A} = [\hat{a}_{ij}] \in \mathcal{R}^{N \times M}$ を算出する。

$$\hat{a}_{ij} = \begin{cases} \alpha \bar{a}'_{ij} + (1 - \alpha) \underline{a}'_{ij}, & \text{if } a'_{ij} \in \bar{A} \\ \beta \underline{a}'_{ij} + (1 - \beta) \bar{a}'_{ij}, & \text{if } a'_{ij} \in \underline{A} \end{cases} \quad (6)$$

また, 以下に提案手法の手順を以下に示す。

- Step1: 評価値行列 A を \bar{A} , \underline{A} に分割する。
- Step2: 全体の評価値行列 A と, \bar{A} , \underline{A} に対して EM-NMF の更新が収束するまで行う。
- Step3: 各 \bar{A}' , \underline{A}' を式 (6) を用いアンサンブルする。
- Step4: 予測評価値行列 \hat{A} を用い, 対象ユーザに推薦を行う。

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い, 従来手法との比較を行った。

5.1 実験条件

実験では, 公開データセット MovieLens[2] の映画評価データ 100 万件を用いた。ユーザ数 $M = 6,040$ (人), アイテム数 $N = 3,952$ (本), 評価段階数は (1~5) であり, 学習データを 80 万件, テストデータを 20 万件とした。また, 与えられた評価値行列において各映画の被評価数の合計を算出し, 被評価数上位 2 割を MAJOR 映画, それ以外を MINOR 映画と定義した。ユーザに関しては, 評価している映画の 8 割が MAJOR 映画であれば MAJOR ユーザ, それ以外の

ユーザを MINOR ユーザと定義した。それぞれのユーザ数は, $\bar{M} = 2,018$, $\underline{M} = 4,022$ である。また, K は予備実験の結果から $K = 10$ とした。

上記条件で実験を 10 回行い, その平均値により予測精度の比較を行う。本実験では評価指標として式 (5) で与えられる MAE を用いる。

5.2 実験結果と考察

$K = 10$ において, α , β 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 2, 3 に示す。

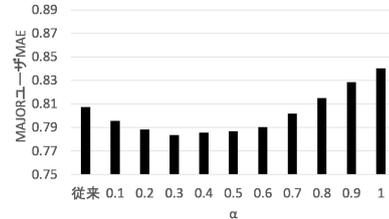


図 2: パラメータ α と MAJOR ユーザの MAE の関係

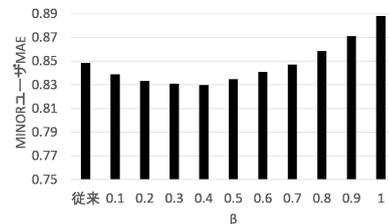


図 3: パラメータ β と MINOR ユーザの MAE の関係

図 2, 3 が示すように, 提案手法の MAE は α , β ともに 0.1~0.7 のときに従来手法を上回っている。MAJOR ユーザの MAE は $\alpha = 0.3$ で, MINOR ユーザの MAE は $\beta = 0.4$ で最小となった。この結果より, $\alpha = 0.3$, $\beta = 0.4$ とし, 従来手法と提案手法の実験結果を表 2 に示す。

表 2: 実験結果

	MAE		
	全体	MAJOR ユーザ	MINOR ユーザ
従来手法	0.832	0.807	0.845
提案手法	0.813	0.785	0.827

表 2 から, MAJOR ユーザと MINOR ユーザ両方で MAE が改善されていることがわかる。従って, 提案手法では従来手法で捉えきれなかった各ユーザの嗜好を, 評価値行列を分割することで捉えることができたと考えられる。

実験から得られた適切なパラメータの値は, MINOR ユーザの混合比率が $\beta = 0.4$ と MAJOR ユーザの混合比率 $\alpha = 0.3$ よりも高い。従って, MINOR ユーザに対して細かな特徴を捉えるためには, ユーザ分割が必要であったと考えられる。

5.3 まとめと今後の課題

本研究では, MAJOR ユーザと MINOR ユーザに分割した評価値行列に EM-NMF を適用して算出した予測値と評価値行列全体に EM-NMF を用いて算出した予測値をアンサンブル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を示すことができた。今後の課題として, 本実験ではアンサンブルに用いるパラメータを経験的に求めたが, そのパラメータの適切な選択方法の検討, また, 行列の適切な分割方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.
- [2] MovieLens. “<http://www.movielens.org/>”, 2017/1/12 アクセス。