

なんとなく協調フィルタリング — 順序応答に基づく推薦

Nantonac Collaborative Filtering — Recommendation Based on Order Responses

神島 敏弘* (Toshihiro KAMISHIMA)

産業技術総合研究所 (National Institute of Advanced Industrial Science and Technology (AIST))

Abstract: A recommender system suggests the items expected to be preferred by the users. Recommender systems use collaborative filtering to recommend items by summarizing the preferences of people who have tendencies similar to the user preference. Traditionally, the degree of preference is represented by a scale, for example, one that ranges from one to five. This type of measuring technique is called the semantic differential (SD) method. We adopted the ranking method, however, rather than the SD method, since the SD method is intrinsically not suited for representing individual preferences. In the ranking method, the preferences are represented by orders, which are sorted item sequences according to the users' preferences. We here propose some methods to recommend items based on these order responses, and carry out the comparison experiments of these methods.

1 はじめに

推薦システム (Recommender System) は、利用者が好むであろうアイテムを見つけ出すもので [23], 電子商取引の拡大にともなって、その重要性が増している [26]. 次に述べる協調フィルタリング (Collaborative Filtering; CF) は、「口コミ」の原理を用いて、推薦システムを実装したものである: 最初に、システムを利用するユーザは自身の嗜好を提示する. システムは、提示された嗜好と類似した嗜好をもつ他のユーザをデータベースから見つけ出す. これらの見つけ出したユーザが好むアイテムを利用ユーザに推薦する.

CFの手法は数多く研究されているが、ほとんどのものは Semantic Differential 法 (SD 法) [21] を用いて利用者の嗜好を計測する. この方法では、利用者の嗜好を、両端を対義語で表した次のような物差しから、自身の嗜好の度合いを選択させることで計測する.

好き 5 4 3 2 1 嫌い

だが、この方法は個人の嗜好を計測するには、1.1 節で述べるような問題点がある.

そこで、利用者の嗜好を順序によって提示する順位法を用いた CF [10] の枠組みを提案する. ここでいう順序とは、利用者の嗜好の強さの順にアイテムを整理し

たものである. このような順序応答を扱う手法をいくつか提案し、これらの手法が SD 法を用いるものより予測性能が良いことを示す. なお、絶対的な尺度ではなく、相対的な嗜好の強弱を利用者は示すだけで、より直観的に利用できるという意味で、この手法を「なんとなく CF」と呼ぶことにする.

以下、2 節では、従来のもとの、提案する CF の枠組みについて、3 節では、なんとなく CF の手法について、4 節では実験結果を示し、最後に 5 節にまとめを述べる.

1.1 順序応答の利点

ここでは、SD 法が CF に適さない理由をいくつか示すが、その前に、Stevens の尺度水準 (scale level) [27] について述べる. Stevens は尺度を四つの水準に分類した: 名義 (nominal), 順序 (ordinal), 間隔 (interval), 比率 (ratio). これらの中で順序尺度と間隔尺度を取り上げる. 順序尺度の値には、大小の判断だけが可能である. 例えば、順序尺度の二つの値 2 と 4 があったとき、「4 は 2 より大きい」という解釈は適切だが、「4 は 2 の 2 倍である」は不適切である. 順序尺度で示される意味は、任意の単調変換について不変である. 段位、階級、モース硬度などが順序尺度である. 一方、間隔尺度の値には、間隔の等しさの判断が可能である. 例え

*<http://www.kamishima.net/>

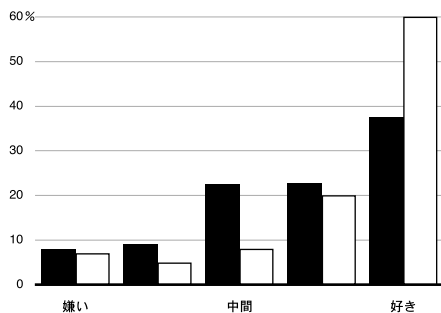


Fig. 1: SD 法による調査で、利用者が選んだスコアの分布

ば、間隔尺度の三つの値 2, 3, 4 があつたとき、「2 と 4 の差は、3 と 4 の差の 2 倍」という解釈は適切だが、「4 は 2 の 2 倍」は不適切である。間隔尺度で示される意味は、任意の線形変換に対して不変で、エネルギーや時刻などが間隔尺度である。

Semantic Differential (SD) 法 [21] に話を戻す。SD 法では、順序尺度で表された観測値が得られる。例えば、利用者がアイテム A と B をそれぞれ 5 段階中 4 と 2 と評価したとき、適切な解釈は A より B が好きということだけである。だが、解析手法の不足から、これらのスコアはしかたなく間隔尺度としてみなされることが多い。そのため、「各被験者ごとに目盛の間隔は等しい」と「全被験者の間で尺度の全長は一致する」という二つの非現実的な仮定が導入される [20]。SD 法は、本来、被験者の集団の概念を定量化するためのものである。尺度の間隔や極値の変動は、集団中で平均化されることで弱められるので、この目的では効果的な手法である。しかし、CF では個人の嗜好を扱うので、この変動はアイテムの推薦の品質に大きく影響する。なぜなら、利用者の応答が標準的な尺度からずれていれば、データベース中の利用者との嗜好の類似性を正しく評価できなくなるからである。それに対し、順序を用いれば、尺度の間隔は考慮する必要はなく、相対的な大小関係だけを扱うため極値の概念を共有する必要もない。

この問題点をより明らかにするために、SD 法で質問した場合に、利用者が選んだスコアの分布を図 1 に示す。黒は 4 節の実験で用いた寿司の嗜好調査データ、白は Amazon.com でユーザがつけた商品に対する評価のデータ¹である。このスコアを間隔尺度として扱う仮定によれば「中間」を中心とする単峰で対称な分布が期待されるが、実際の分布は図のように非常に片寄って

いる。また、尺度中の非常に狭い範囲のスコアが使われていて、獲得できる情報量も少ない。

他にも、質問によっては、評価尺度の中立的な部分を被験者が利用しがちになってしまう Central Tendency 効果などの他、多くの rator 効果が知られており、これらによっても利用者の嗜好の正確な計測は阻害される。

1.2 関連研究

順序を扱う研究には以下のようなものがある。18 世紀には、投票によって候補の順序を決定する社会選択 [30] の研究が、J.-C. de Borda, M. de Condorcet, C. L. Dodgson²らによって始められた。20 世紀には K. Arrow が、民主的な決定に必要な 4 条件を示し、それを満たして決定を行う関数が数学的に存在しないという著名な Arrow's impossibility 定理を示した。Thurstone [28] は順序が与えられた対象対の集合から、実数値の尺度を構成する方法を提案した。順序を対象とするモデルについてまとめた文献には [19] が、順序付きカテゴリ変数を目的変数とする回帰分析である ordinal regression についてまとめた文献には [1] がある。

近年の研究には以下のようなものがある。Cohen ら [3], Herbrich ら [7], Joachims [8] は対ごとの順序関係から、属性ベクトルで記述された対象を順序付けする手法を提案した。神島と赤穂 [11] や賀沢ら [14] は順序付けされた対象集合からの学習問題について研究した。Freund は順序付けを目的とする boosting である RankBoost を考案した [5]。彼らは、この手法を協調フィルタリングに適用している。目標が順序である点は本研究と同じだが、嗜好の計測には従来どおり SD 法を用いている。Mannila と Meek [18] は順序の集合から特徴的な順序構造を見つける手法を提案した。Sai ら [24] は順序変数に対する相関ルールを提案した。神島ら [12, 13] は順序をクラスタリングした。佐藤ら [25] は、与えられた類似性の順序関係を保持する類似度関数を構成した。Lebanon と Lafferty [16, 17] は順序変数を条件とする順序変数の条件付き確率分布を推定した。メタ検索エンジンを順序の統合問題として扱った Dwork ら [4] の研究もある。鷲尾ら [29] は、尺度水準を制約として用いて知識発見を行った。Gionis ら [6] は、0/1 属性の共起性から属性間の順序関係を導く手法を提案した。

¹A.S.Weigend の KDD2003 招待講演より

²「不思議の国のアリス」を著した Lewis Carroll のペンネームでも知られる

2 協調フィルタリング

SD法を用いた従来の協調フィルタリングと順位法を用いたなんとなく協調フィルタリングについて述べる。

2.1 従来の協調フィルタリング

RiedlらのGroupLensシステム[22]が採用しているCFの枠組みについて述べる。協調フィルタリングは、ある特定の利用者(活動利用者)の嗜好を、他の利用者の集積した嗜好データ(利用者DB)から予測する問題で、形式的には次のとおり。\$s_{ij}\$は、利用者\$i\$がアイテム\$j\$を評価したスコア。このスコアは利用者の嗜好を表し、\$\{1, 2, 3, 4, 5\}\$のようないくつかの候補から一つを選択する。\$X_i = \{x_1, \dots, x_{|X_i|}\}\$は、利用者\$i\$が評価をしたアイテム集合で、そのスコアの集合が\$S_i = \{s_{ij} | x_j \in X_i\}\$。\$X^* = \{x_1, \dots, x_{|X^*|}\}\$はアイテムの全集合。ただし、\$|X|\$は集合\$X\$の大きさ。利用者DB\$D_S = \{S_1, \dots, S_{|D_S|}\}\$は、全ての\$S_i\$の集合。ここでは、DB中の利用者を標本利用者と呼ぶ。\$S_a\$は活動利用者が評価したスコアの集合で、\$X_a\$はすでに活動利用者が評価したアイテムの集合。\$S_a\$と\$D_S\$から、活動利用者が高く評価するであろうアイテムを予測し、それらを活動利用者に推薦するのがCFの目標である。

GroupLensの推定手法を以下に示す：まず、活動利用者と\$D_S\$中の標本利用者\$i\$の類似度を相関係数で測る：

$$r_{ai} = \frac{\sum_{j \in I(X_{ai})} (s_{aj} - \bar{s}_a)(s_{ij} - \bar{s}_i)}{\sqrt{\sum_{j \in I(X_{ai})} (s_{aj} - \bar{s}_a)^2} \sqrt{\sum_{j \in I(X_{ai})} (s_{ij} - \bar{s}_i)^2}}$$

ただし、\$X_{ai} = X_a \cap X_i\$、\$I(X) = \{j | x_j \in X\}\$で、\$\bar{s}_i\$は、\$X_{ai}\$中のアイテムについての、利用者\$i\$の平均スコア。活動利用者のアイテム\$j\$の予測スコアは

$$\hat{s}_{aj} = \bar{s}_a + \frac{\sum_{i \in I_j} r_{ai} (s_{ij} - \bar{s}_i)}{\sum_{i \in I_j} |r_{ai}|} \quad (1)$$

ただし、\$\bar{s}_a = |S_a|^{-1} \sum_{s_{aj} \in S_a} s_{aj}\$、\$I_j = \{i | S_i \in D_S \text{ s.t. } s_{ij} \in S_i\}\$。システムは、予測スコア\$\hat{s}_{aj}\$が高いアイテムを推薦する。

このアルゴリズムは非常に単純だが、効果的である。文献[2]の比較研究では、このような明示的な評価を用いる場合では最も優れていた。暗黙的で2値の評価スコアを利用する場合にはベイジアンネットを用いる手法が優れているが、その差はわずかである。ここでは明示的な多段階の評価を利用するので、このGroupLensの手法を評価対象とした。

2.2 なんとなく協調フィルタリング

次に我々の提案する枠組みである「なんとなく協調フィルタリング」について述べる。従来の枠組みとは、利用者の嗜好の表現方法だけが異なり、SD法で得られたスコアの集合を用いる代わりに、順位法により得られた順序を用いる。システムがアイテム集合\$X_i\$を利用者\$i\$に提示すると、利用者はこれらのアイテムを嗜好の度合いによって整列した順序\$O_i = x_1 \succ x_2 \succ \dots \succ x_{|X_i|}\$を返す。この順序の意味は、利用者\$i\$は\$x_1\$より\$x_2\$を好むといったことを表す。順位\$r(O_i, x_j)\$は、アイテム\$x_j\$の順序\$O_i\$中の位置を示す基数である。例えば、順序\$O_i = x_1 \succ x_3 \succ x_2\$について\$r(O_i, x_2)\$は3である。この枠組みでは、利用者DBは、ユーザが整列した順序の全集合\$D_S = \{O_1, \dots, O_{|D_S|}\}\$になる。\$X_a\$を活動利用者が整列したアイテムの集合、\$O_a\$をその順序とする。与えられた\$O_a\$と\$D_S\$から、活動利用者が好むアイテムを予測する問題がなんとなくCFである。

3 手法

2.2節のなんとなくCFの解法をいくつか示す。比較研究論文[2]では、CF手法を、利用者DB全体を用いて利用者の嗜好を予測するメモリベースと、事前に利用者DBから学習したモデルを用いて予測するモデルベースとに分類している。これらの手法に加えて、ここではこれらを合成した手法も提案する。

3.1 メモリベース法

GroupLensの方法に代表される方法で、これらの手法を順序に適用できるように修正したものを示す。

CFを実行するためには、活動利用者と標本利用者の嗜好の類似性を測る必要がある。この手法では、アドホックに、順位をそのまま、GroupLensの手法におけるスコアと置き換える。すなわち、類似度は次式

$$R_{ai} = \frac{\sum_{x_j \in X_{ai}} (r(O_a, x_j) - \bar{r}_a)(r(O_i, x_j) - \bar{r}_i)}{\sqrt{\sum_{x_j \in X_{ai}} (r(O_a, x_j) - \bar{r}_a)^2} \sqrt{\sum_{x_j \in X_{ai}} (r(O_i, x_j) - \bar{r}_i)^2}} \quad (2)$$

ただし、\$X_{ai} = X_a \cap X_i\$、\$\bar{r}_i = |X_{ai}|^{-1} \sum_{x_j \in X_{ai}} r(O_i, x_j)\$である。他方の順序に含まれないアイテムは除外するが、順位の付け直しはしない。例えば、\$O_a = x_1 \succ x_2 \succ x_3\$において、\$x_2\$が\$O_i\$に含まれないので計算から除外した

アルゴリズム k -o'means($S, k, maxIter$)

$D_S = \{O_1, \dots, O_{|D_S|}\}$: サンプル集合

k : クラスタの数

$maxIter$: 反復回数の上限

- 1) 初期分割: D_S をランダムに分割 $\pi = \{C_1, \dots, C_k\}$,
 $\pi' := \pi, t := 0$
- 2) $t := t + 1$, もし $t > maxIter$ ならステップ 6 へ
- 3) 各クラスタ $C_j \in \pi$ について
 中心順序 \bar{O}_j を Thustone の方法で求める
- 4) S 中の各順序 O_i を次のクラスタに割り当て:
 $\arg \min_{C_j} d_{Spear}(\bar{O}_j, O_i)$
- 5) もし $\pi = \pi'$ ならばステップ 6 へ
 でなければ $\pi' := \pi$, ステップ 2 へ
- 6) π を出力

Fig. 2: k -o'means 法

場合に、活動利用者の順序を $x_1 \succ x_3$ として x_3 の順位を 2 とするのではなく、もとの順序の順位である 3 をそのまま用いる。活動利用者のアイテム j についての嗜好は次式で予測する。

$$\tilde{r}_a + \frac{\sum_{i \in I_j} R_{ai}(r(O_i, x_j) - \tilde{r}_i)}{\sum_{i \in I_j} |R_{ai}|} \quad (3)$$

ただし、 $\tilde{r}_a = |X_a|^{-1} \sum_{x_j \in X_a} r(O_a, x_j)$ であり、 $I_j = \{i | O_i \in D_S \text{ s.t. } x_j \in X_i\}$ 。アイテムをこの式の値で整列し、上位に順位付けされたものを推薦する。

3.2 モデルベース法

文献 [2] では、クラスタモデルとベイジアンネットの二つのモデルベース手法が提案されているが、ベイジアンネットは順序に適用できないので、クラスタモデルのみを実装した。

ここでは順序を k -o'means 法 [12, 13] でクラスタリングする。 k -o'means 法は、クラスタの中心として Thustone の一対比較法で求めた中心順序を、距離として Spearman の距離 [19] を用いること以外は k -means 法 [9] と同じである。アルゴリズムを図 2 に示す。最初に、 D_S をランダムに分割して初期分割を得る。クラスタの中心順序の再計算 (ステップ 3) と、順序のクラスタへの再割り当て (ステップ 4) の二つのステップを反復することで、クラスタは改良される。反復回数が上限 $maxIter$ をこえるか、分割に変化が無かった場合に停止して、現在の分割を出力する。 k -means と同様

に、このアルゴリズムは局所最適解にしか収束しないため、初期分割を変えて数回繰り返し、各サンプル順序から中心順序までの距離の総和が最小になるものを選択する。

ステップ 3 の中心順序は次の Thustone の方法により求める。アイテム $x_l \in \bar{X}_j$ (ただし、 $\bar{X}_j = \cup_{O_i \in C_j} X_i$) について次式の値を求める。

$$\bar{\mu}(x_l) = \frac{1}{|\bar{X}_j|} \sum_{x' \in \bar{X}_j} \Phi^{-1}(\Pr[x_l > x']) \quad (4)$$

ただし、 $\Phi(\cdot)$ は正規分布の分布関数。また、 C_j 中の順序を順序対に分解し、これらの順序対の頻度を求めることで、 $\Pr[x_l > x']$ は計算できる。ステップ 4 の中心順序と O_i の距離 d_{Spear} は、順序中のアイテムの順位の差の 2 乗和である Spearman の距離を用いる。

推薦の方法は次のように行う。推薦の前に、 k -o'means 法によって利用者 DB D_S を k 個のクラスタ $\{C_1, \dots, C_k\}$ に分割する。活動利用者の順序 O_a が与えられると、各クラスタの中心順序と O_a の距離を求めて、活動利用者と最も類似した嗜好をもつ標本利用者のクラスタを見つける。そのクラスタの中心順序で上位にあるアイテムを推薦する。

3.3 合成法

クラスタモデルと単純相関法を結合させたものが合成法である。メモリベース法では、活動利用者と、利用者 DB 全体中の各利用者との類似度を計算した。合成法では、まず嗜好が最も類似したクラスタ C^* を 3.2 節の方法で見つけ、活動利用者の嗜好を C^* 中の利用者の嗜好のみを用いて 3.1 節の方法で予測し、他のデータは無視する。

4 実験

上述の方法を寿司の嗜好データを用いて比較する。

4.1 データの収集方法

収集方法について述べる前に、寿司の嗜好調査を実験領域として選んだ理由を述べておく。まず、寿司の種類は多数で、嗜好は個人ごとに異なっており、嗜好の予測実験に適している。さらに、プライバシー上の問題は寿司の嗜好についてはあまりないため、データの収集が容易なためでもある。

データの収集方法は以下に述べる．WWW から 25 軒の寿司店のメニューを抽出し，これから各寿司が提供される頻度を求めた．頻度の低いものや，あまり知られていない寿司を取り除き 100 種の寿司を選びだし，これを X^* とした．

各利用者に対し 2 種類のアイテム集合 X_i を生成した．集合 A (X^A) は全ての利用者に共通で，次の 10 種の寿司で構成される：

海老	穴子	鮪	いか	うに
いくら	玉子	とろ	鉄火巻	かつば巻

この集合に対する応答はテストに利用した．もう一つの集合 B (X_i^B) は，寿司店での提供頻度に比例する確率で 10 種のアイテムを，各被験者ごとにランダムに選んだ．この集合は活動利用者と標本利用者の類似性を測るために用いた． X^A と X_i^B は平均 2.58 個のアイテムが重複している．

WWW アンケート調査会社を利用して以下の質問項目のデータを収集した．

-
- 1: 順位法 1) 利用者 i に X^A 中の寿司を嗜好順に整列させ，その応答順序を O_i^A とした．
 - 2: SD 法) 利用者 i に， X_i^B 中の寿司について 5 段階の SD 法で嗜好を回答させ，その応答スコアを S_i^B とした．
 - 3, 4) 続いて，嗜好とは無関係な質問を二つ行い，質問 2 の質問 5 への影響を弱めた．
 - 5: 順位法 2) 利用者 i に X_i^B 中の寿司を嗜好順に整列させ，その応答順序を O_i^B とした．
-

回答時間が短すぎたり長すぎるデータは信頼性が低いと考えて取り除き，5000 個の三つ組データ (O_i^A, O_i^B, S_i^B) を得た．また，対象を提示する順序による影響を避けるため，提示順序も被験者ごとにランダムに変更した．表示順と O_i^A や O_i^B との順位相関はそれぞれ 0.004 と 0.010 で，無相関なので，解答がアイテムの表示順に影響されていないことが分かる．SD 法と順位法による回答を比較する予備調査を行った． S_i^B と O_i^B に矛盾がある，すなわち，順序 O_i^B 中でアイテム x_a が x_b より前にあるが， S_i^B 中の x_b のスコアは x_a のそれより高い場合，もしくは，その逆が生じる場合について調査した．このような矛盾は 68.3% の利用者について観測された．このことから，順位法と SD 法では異なる側面から嗜好が計測されていることが分かる．

4.2 評価基準と実験手順

各手法の評価のため，10 分割の交叉確認を行った．すなわち，データ全体 D を 10 個のブロックに分け，一つをテスト集合 D' とし，残りをまとめて訓練集合 \bar{D}' とする．テスト集合を 10 個のブロックそれぞれに変更してその平均で評価する．各データは三つ組 (O_i^A, O_i^B, S_i^B) で構成されている． S_i^B と O_i^B とは，それぞれ，従来の SD 法に基づく手法と，順位法に基づく提案手法で用いた． \bar{D}' 中の全ての $O_i^B (S_i^B)$ で構成される集合を利用者 DB D_S として扱った． D' 中の各 $O_i^B (S_i^B)$ を順に取り出し，活動利用者の応答 $O_a^B (S_a^B)$ として扱った．与えられた D_S と $O_a^B (S_a^B)$ から，集合 X^A 中のアイテムの嗜好の順序を予測し，その予測順序を \hat{O}_i^A とした． D' 中の応答順序 O_i^A を真の嗜好順序と考えて，予測順序 \hat{O}_i^A と比較する．予測精度は， \hat{O}_i^A と O_i^A の間の Spearman の順位相関 ρ [15] によって行った．これは二つの順序が一致する場合に 1 となり，逆順の場合に -1 となる．

利用者 DB の大きさの変化に伴う影響を調査するために，大きさ $|D|$ を 5000, 3000, 1000, 500, 300, 100 の 6 種類に替えたデータを生成した．同様に， X_i^B 中のアイテムをランダムに取り除き，アイテムの順位をつけ直すことで，アイテム集合 X_i^B の大きさ $|X^B|$ を 10, 7, 5, 3, 2 と変えたデータも生成した．

4.3 実験結果

前節の評価手続きを各手法について適用した．これらの手法は以下の記号で表す：SCR (2.1 節), RCR (3.1 節), CLS (3.2 節) HCR (3.3 節)．

4.1 節で述べたように，テスト用の X^A と訓練用の X_i^B には重複があるので，ソートするアイテムの順位やスコアが活動利用者によって与えられている場合がある．だが，アイテムをソートして推薦リストを作成する場合に，これらの順位やスコアを無視して，全て式 (1) や (3) によって予測されたスコアで整列する方が結果が良かったので，この方法で評価した．CLS と HCR 法では，クラスタ数を決定する必要がある．この数の決定手法を開発できていないので，クラスタ数を 2 から 10 まで変化させて，最も良い結果を示した．また，クラスタリングの出力結果は初期分割に依存するので，ランダムな初期分割による 20 回の試行で，クラスタ内分散の総和を最小にする分割を採用した．

データ集合の大きさ $|D|$ が 5000 で，アイテム集合の

Table 1: ρ の平均と標準偏差 ($|X^B|=10, |D|=5000$)

	SCR	RCR	CLS	HCR
平均	0.433	0.489	0.428	0.489
標準偏差	0.3009	0.2920	0.3397	0.3050

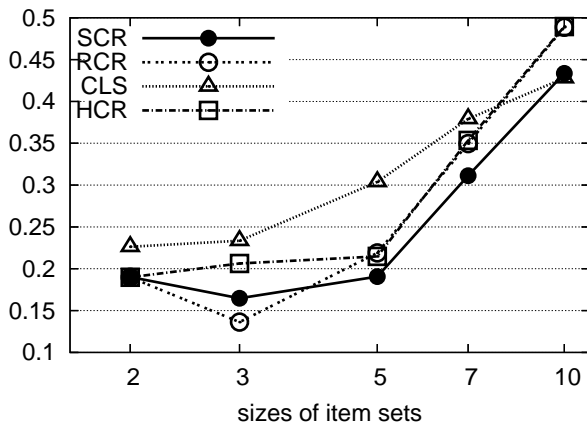


Fig. 3: 応答アイテム集合の大きさ $|X^B|$ に伴う ρ の変化

大きさ $|X^B|$ が 10 であるデータ集合に各手法を適用した。Spearman の順位相関 ρ の平均と標準偏差を表 1 に示す。SD 法を用いた手法は、順序を用いた手法より明らかに悪かった。この結果は、1.1 節で論じた SD 法を CF に用いた場合の、不適切な特性によるものだと考える。図 1 のようなデータの偏りに対しては、各利用者ごとに、スコアの最大値と最小値をそろえるといった正規化が有効ではないかと考えるかもしれない。このような正規化を SCR で行くと ρ の平均は 0.443 となり、元の SCR 法より良くなるが、依然として RCR より有意に劣る。

図 3 に、 $|D|$ を 5000 に固定して、応答アイテム集合の大きさ $|X^B|$ を変化させた場合の性能変化を示す。SCR 法と順序に基づく他の手法を比較すると、 $|X^B|$ が大きくなると順序に基づく手法が優位になることが分かる。順位に基づく手法の中で CLS 法はやや特徴的である。この方法は、たとえ $|X^B|$ が小さくても比較的良いが、 $|X^B|$ が大きいと他の手法より劣る。これは、他の手法は純粹に個人ごとに嗜好を予測するのに対し、CLS では活動利用者と嗜好が類似したグループの嗜好を予測している。そのため、個人情報が少ない場合には、CLS 法はやや一般化した嗜好に基づき、相対的に良い推薦ができたと考える。

図 4 に、 $|X^B|$ を 10 に固定したときの、データ集合の大きさ $|D|$ の変化に伴う ρ の平均の変化を示す。ここでも、順序に基づく手法はより多くの情報がある場

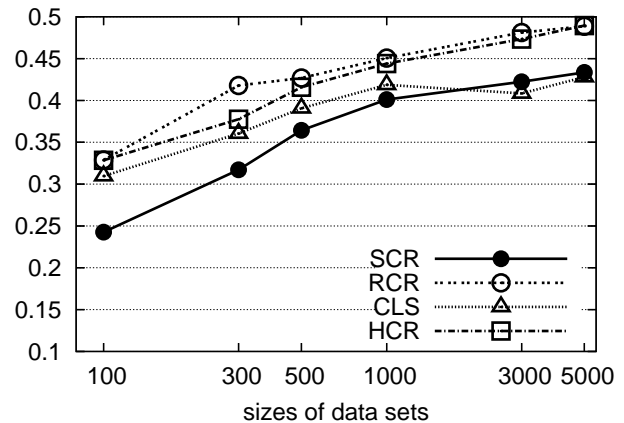


Fig. 4: データ集合の大きさ $|D|$ に伴う ρ の変化

Table 2: SCR と他の手法で得られた ρ の間の t -値

(a) アイテム集合の大きさ $|X^B|$ を変えた場合

$ X^B $	10	7	5	3	2
RCR	14.5725	8.6067	5.5445	-5.9460	-0.2259
CLS	-1.0383	11.0442	17.0685	10.2969	6.1435
HCR	13.0015	8.4744	4.3518	9.0968	-0.2304

(b) データ集合の大きさ $|D|$ を変えた場合

$ D $	100	300	500	1000	3000
RCR	2.6002	5.6695	4.5660	5.4847	11.4172
CLS	1.5243	1.9414	1.4077	1.4135	-1.9201
HCR	2.3917	3.2183	3.4867	4.1855	8.8708

合に、SCR との性能差が広がった。

SCR 法と他の手法との差を厳密に検証するために、対応のある t -検定を適用した結果が表 2 である。正の t -値は順序に基づく手法の優位を示し、上線と下線はそれぞれ、順序に基づく手法が優位水準 1% で有意に優れていることと劣っていることを示す。表 2(a) について検証すると、前述のように、CLS 法を除いて、順序に基づく方法は SCR 法より $|X^B|$ が大きければ優れている。だが、 $|X^B| = 2$ や 3 のときには SCR 法は若干よい。これは、 $|X^B| = 2$ では順位法では二つのアイテムのどちらを上位にするか限られた選択しかないため、十分な情報を取得できないものとする。まとめると、5 個以上のアイテムを整列する場合、順位に基づく手法は既存手法より有意に優れていた。表 2(b) の場合、CLS を除いて、 $|D|$ が大きくなるに従い、順序に基づく手法が相対的に良くなった。

3.3 節の合成法の検証ため、合成した方法と元の手法

Table 3: 合成法とその元の手法との比較

(a) アイテム集合の大きさ $|X^B|$ を変えた場合

$ X^B $	10	7	5	3	2
RCR	0.2506	1.0983	-1.0522	15.7523	-0.0062
CLS	12.7900	-5.1430	-15.0477	-4.2395	-6.2567

(b) データ集合の大きさ $|D|$ を変えた場合

$ D $	100	300	500	1000	3000
RCR	-0.0096	-3.4678	-1.1130	-0.8846	-2.0710
CLS	0.4436	0.8570	1.5689	2.3934	9.7887

Table 4: アルゴリズムの計算量

	off-line	on-line
SCR	0	$ X^A D $
RCR	0	$ X^A D $
CLS	$ X^* ^2 S + k X^* S $	$k X^A $
HCR	$ X^* ^2 S + k X^* S $	$ X^A D /k$

とで得られた結果を比較する。表 3 は HCR と、その元となった RCR や CLS とによって得られた ρ の差の t -値である。正の t -値は合成法が良いことを示す。RCR との比較では、 $|D|$ や $|X^B|$ の変化に伴う明らかな特徴はみられず、 $|X^B|$ が 3 の場合を除いて有意な差はない。合成法では、事前に利用者 DB をクラスタリングしておくことで、推薦のときに類似度を計算しなければならない標本利用者の数を抑制できる。さらに、合成によって性能が低下しなかったというこの結果から、予測精度を下げずに合成法によってモデルベース法の高速度が可能になるといえる。CLS との比較では、合成によって CLS 法の特徴は失われるようである。よって、他のメモリベース法と同様に、合成法は $|X^B|$ が大きくなるに従って CLS 法より性能が良くなる。一方 $|D|$ の変化に対しては目だった特徴は見られなかった。

4.4 計算量に関する議論

アルゴリズムの計算量について、表 4 にまとめた。off-line と on-line の列は、それぞれ、活動利用者の応答を受け取る前と後に必要な時間を示した。メモリベース法は、モデルベースより on-line 時間をより多く必要とするので、モデルベースの方がより早く推薦ができる。しかし、モデルベース法では、利用者 DB の更新の度にモデルを再構築する必要がある。前述のように、

合成法では予測精度を低下させることなく on-line の計算量を削減できる。しかし、この合成法もモデルベースと同じように、モデルの再構築が必要になるという同じ欠点をもつ。

5 まとめ

順序応答に基づいてアイテムを推薦するなんとなく CF の枠組みを提案した。なんとなく CF の予測性能が従来手法を上回ることを示した。

利用者が 10 個以上のような多数のアイテムを整理するのは困難なので、現在の枠組みでは、 X^* が大きい場合に対応できない。よって、利用者の嗜好データをより集めるために、各利用者が複数アイテム集合についての順序を返す場合を取り扱える手法を開発したい。

謝辞： 科研費萌芽研究 (14658106) の助成を受けた

参考文献

- [1] A. Agresti. カテゴリカルデータ解析入門. サイエンス出版社, 2003. (渡邊 裕之 他訳).
- [2] J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Uncertainty in Artificial Intelligence 14*, pp. 43–52, 1998.
- [3] W. W. Cohen, R. E. Schapire, and Y. Singer. Learning to order things. *Journal of Artificial Intelligence Research*, Vol. 10, pp. 243–270, 1999.
- [4] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar. Rank aggregation methods for the Web. In *Proc. of The 10th Int'l World Wide Web Conf.*, pp. 613–622, 2001.
- [5] Y. Freund, R. Iyer, R. E. Schapire, and Y. Singer. An efficient boosting algorithm for combining preferences. *Journal of Machine Learning Research*, Vol. 4, pp. 933–969, 2003.
- [6] A. Gionis, T. Kujala, and H. Mannila. Fragments of order. In *Proc. of The 9th Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 129–136, 2003.

- [7] R. Herbrich, T. Graepel, P. Bollmann-Sdorra, and K. Obermyer. Learning preference relations for information retrieval. In *ICML-98 Workshop: Text Categorization and Machine Learning*, pp. 80–84, 1998.
- [8] T. Joachims. Optimizing search engines using click-through data. In *Proc. of The 8th Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 133–142, 2002.
- [9] 神鷹敏弘. データマイニング分野のクラスタリング手法 (1) — クラスタリングを使ってみよう! —. 人工知能学会誌, Vol. 18, No. 1, pp. 59–65, 2003.
- [10] T. Kamishima. Nantonac collaborative filtering: Recommendation based on order responses. In *Proc. of The 9th Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 583–588, 2003.
- [11] T. Kamishima and S. Akaho. Learning from order examples. In *Proc. of The IEEE Int'l Conf. on Data Mining*, pp. 645–648, 2002.
- [12] T. Kamishima and J. Fujiki. Clustering orders. In *Proc. of The 6th Int'l Conf. on Discovery Science*, pp. 194–207, 2003. [LNAI 2843].
- [13] 神鷹敏弘, 藤木淳. 順序のクラスタリング — 順序平均の最適性について. 電子情報通信学会技術研究報告, PRMU 2003–83, 2003.
- [14] 賀沢秀人, 平尾努, 前田英作. Order SVM: 一般化順序統計量に基づく順位付け関数の推定. 電子情報通信学会論文誌 D-II, Vol. J86-D-II, No. 7, pp. 926–933, 2003.
- [15] M. Kendall and J. D. Gibbons. *Rank Correlation Methods*. Oxford University Press, fifth edition, 1990.
- [16] G. Lebanon and J. Lafferty. Conditional models on the ranking poset. In *Advances in Neural Information Processing Systems 15*, pp. 431–438, 2002.
- [17] G. Lebanon and J. Lafferty. Crankng: Combining rankings using conditional probability models on permutations. In *Proc. of The 19th Int'l Conf. on Machine Learning*, pp. 363–370, 2002.
- [18] H. Mannila and C. Meeck. Global partial orders from sequential data. In *Proc. of The 6th Int'l Conf. on Knowledge Discovery and Data Mining*, pp. 161–168, 2000.
- [19] J. I. Marden. *Analyzing and Modeling Rank Data*, Vol. 64 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, 1995.
- [20] 中森義輝. 感性データ解析 — 感性情報処理のためのファジィ数量分析手法. 森北出版, 2000.
- [21] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum. *The Measurement of Meaning*. University of Illinois Press, 1957.
- [22] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. GroupLens: An open architecture for collaborative filtering of Netnews. In *Proc. of The Conf. on Computer Supported Cooperative Work*, pp. 175–186, 1994.
- [23] P. Resnick and H. R. Varian. Recommender systems. *Communications of The ACM*, Vol. 40, No. 3, pp. 56–58, 1997.
- [24] Y. Sai, Y. Y. Yao, and N. Zhong. Data analysis and mining in ordered information tables. In *Proc. of The IEEE Int'l Conf. on Data Mining*, pp. 497–504, 2001.
- [25] 佐藤健, 岡本青史. 距離情報による類似度関数の重み学習. 人工知能学会誌, Vol. 11, No. 2, pp. 238–245, 1996.
- [26] J. B. Schafer, J. A. Konstan, and J. Riedl. E-commerce recommendation applications. *Data Mining and Knowledge Discovery*, Vol. 5, pp. 115–153, 2001.
- [27] S. S. Stevens. Mathematics, measurement, and psychophysics. In S. S. Stevens, editor, *Handbook of Experimental Psychology*. John Wiley & Sons, Inc., 1951.
- [28] L. L. Thurstone. A law of comparative judgment. *Psychological Review*, Vol. 34, pp. 273–286, 1927.
- [29] 鷲尾隆, 元田浩. スケールタイプ制約に基づく科学的法則式の発見. 人工知能学会誌, Vol. 15, No. 4, pp. 681–692, 2000.
- [30] 松原望. 計量社会科学. 東京大学出版会, 1997.