

順序例からの学習 — 比較判断の法則の導入と嗜好調査データへの適用

Learning from Order Examples —

Adopting the Law of Comparative Judgement and Application to Questionnaire Survey Data

神島 敏弘 (産業技術総合研究所)

E-mail: mail@kamishima.net Homepage: <http://www.kamishima.net/>

Abstract: 本論文では, 学習タスク『順序例からの学習』(Learning from Order Examples; LOE) について述べる. LOE タスクの目的は, 未整列のアイテム集合の順序を推定する規則を獲得することである. この規則を, アイテム集合とこの集合の真の順序の対である学習事例の集合から獲得する. 著者の以前の LOE 手法に加え, Thurstone の比較判断の法則を用いた手法を導入し, これらの手法を人工データと嗜好調査の実データに適用し, 各手法の特性を調査する.

1 はじめに

本論文では, 文献 [3] で提案した学習タスク『順序例からの学習』(Learning from Order Examples; LOE) について述べる.

順序とは, 嗜好の強さ, 価格, 大きさなどの基準で整列したアイテムの系列である. 例えば, 三つの料理: a, b, c を, ある人が好んでいる順で整列した $c \succ a \succ b$ は順序である (料理 c が一番好きで, b が嫌い). LOE タスクの目的は, 未整列のアイテム集合の順序を推定する規則を獲得することである. この規則を, アイテム集合とこの集合の真の順序の対である学習事例の集合から, 獲得する.

この学習タスクは嗜好の調査などに適用できる. この種の調査では, 『好き』から『嫌い』までを何段階かに区切り, 被験者にそのうちの一つを指定するよう依頼する Semantic Differential (SD) 法 [8] が広く用いられている. SD 法の尺度は原点と各段階の間隔が等しいこと, すなわち, どの被験者も感覚の尺度を共有している, その尺度に基づき判断を行っていることを仮定している. この仮定は非現実的だが, 多くの被験者に対する平均では実用的には問題を生じない. だが, 被験者の数が減少するに従い, 個々の尺度の相違が問題となる. それに対し, 入力に順序尺度である場合にはこのような強い仮定はない.

文献 [3] の LOE 手法に加え, Thurstone の比較判断の法則 [5] を用いた手法を導入し, これらの

手法を人工データと嗜好調査の実データに適用し, 各手法の特性を調査する.

2 節では, 関連研究について述べ, 3 節では, LOE タスクを形式的に定義する. 4 節では LOE 手法について述べ, 5 節と 6 節でこれらの手法を適用した実験結果を示す. 最後に, 7 節でまとめを述べる.

2 関連研究

順序尺度を扱う研究について述べる. 20 世紀の初頭には, 順位の類似性を測定する順位相関係数 (Rank Correlation Coefficient) [4] が考案された. 順位相関係数には Spearman の ρ 係数や Kendall の τ 係数があり, 順位の一致性の判定や, 各種のノンパラメトリック検定に利用されている. 最近の研究では, 画像の照合 [9] に利用したのもある.

心理学の分野では, 一对のアイテムの順序関係から, その順序関係が保存されるような実数上の尺度を与える Thurstone の一对比較法 [5, 11] などの研究がある. 本研究とは, アイテムに属性がなく, 未知のアイテムの順序の決定を扱わない点で異なっている.

順序の決定に関する研究には, Cohen ら [1] の研究がある. タスクへの入力は, 本研究のように複数のアイテムの順序ではなく, 一对比較法と同じく, アイテム対の順序を示したものである. この入力から, アイテムの属性に基づいて, アイテ

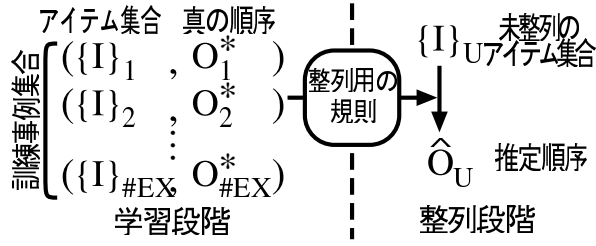


図 1: 順序例からの学習の概要

ム I^x が I^y より前にある可能性を表す評価関数 $\text{PREF}(I^x, I^y)$ を、独自のアルゴリズムにより求める。その後、次式を最大にする順序を求める。

$$\sum_{x,y:I^x \succ I^y} \text{PREF}(I^x, I^y) \quad (1)$$

しかし、この式を最大にする順序を求めることは NP 困難である証明も Cohen らは示している。そこで、欲張り法による探索手法と、部分的な網羅探索と欲張り法を組み合わせた探索手法を提案している。

この Cohen ら研究では、与えられた 2 項間の順序関係をできるだけ保存した順序の推定を目標としている。それに対し、本論文の目標は全体として適切な順序の推定である。この二つの目標は非常に関連があるが異なるものであり、この相違は後の実験結果からも確かめられる。また、彼らの研究では、 PREF 関数のエラーについて考慮しているが、得られた順序のエラーについては考慮してない。関数のエラーを小さくすれば、順序のエラーも一般的には小さくなるが、関数のエラーが非常に小さい ϵ でも $|\text{PREF}(I^x, I^y) - \text{PREF}(I^y, I^x)| < 2\epsilon$ なら順序が入れ替わる可能性があり、これら二つのエラーの相違は明らかである。さらに、意思決定をする場合、最も重要な要素は関数の値ではなく、その結果得られた順序である。よって、本研究では上記の順位相関係数を用いて順序のエラーを直接評価する。

他に、多くの順序データからパターンを検出する Mannila と Meek の研究 [6] や、順序変数のデータから相関ルールを検出する Sai らの研究 [10] などがある。

3 LOE タスクの形式的定義

順序例からの学習 (LOE) タスクの形式的定義について述べる。このタスクは、図 1 のように、学習段階と整列段階の二段階に分けられる。図 1 左の学習段階では、整列用の規則を訓練事例集合から獲得し、右の整列段階では、獲得された規則を用いて、未整列のアイテム集合の推定順序を求める。

アイテム I^x とは整列される物や対象で、属性ベクトル $A(I^x) = (a^1(I^x), a^2(I^x), \dots, a^{\#A}(I^x))$ ($\#A$ は属性数) で記述される。ここでは、全ての属性がカテゴリ属性である場合を扱い、 s 番目の属性は $v_1^s, \dots, v_{\#a^s}^s$ ($\#a^s$ は属性値数) の中のいずれかの値をとる。アイテム全体の集合をアイテム全集合、 $\{I\}_{All}$ と呼び、その部分集合を $\{I\}_i$ で表し、単にアイテム集合と呼ぶ。アイテム集合 $\{I\}_i$ の要素数は $\#I_i$ で表す。

順序とは、大きさ、嗜好の強さ、価格などの何らかの特性に従ってアイテムを整列した系列である。アイテム集合 $\{I\}_i = \{I^x, I^y, \dots, I^z\}$ の順序を $O_i = I^x \succ I^y \succ \dots \succ I^z$ と記す。 $O_1 = I^9 \succ I^3 \succ I^7$ は、アイテム集合 $I_1 = \{I^3, I^7, I^9\}$ の順序の一例である。また、二つのアイテムの間の順序が $I^1 \succ I^2$ であるとき、 I^1 は I^2 の前であるという。

アイテム全集合には観測できない順序があり、この順序を絶対順序 O_{All}^* と呼ぶ。事例は、アイテム集合 $\{I\}_i$ とこの集合の真の順序 O_i^* の対である。この真の順序は、絶対順序に無矛盾な順序に、入れ替わりなどのノイズが加えられたものである。これは、回帰分析で、線形関数にノイズが付加された値が観測されるモデルを用いるのと似ている。事例集合 EX は $\#EX$ 個の事例を含む

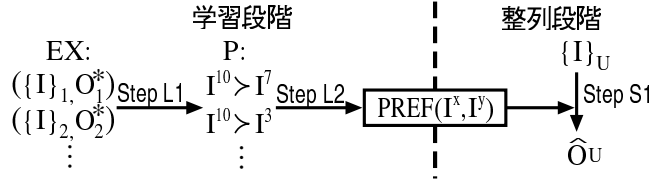


図 2: 分類手法を用いた LOE 解法の概要

集合である .

$$EX = \{(\{I\}_1, O_1^*), (\{I\}_2, O_2^*), \dots, (\{I\}_{\#EX}, O_{\#EX}^*)\}$$

アイテム全集合に含まれていても、事例集合のどのアイテム集合にも含まれないアイテムが存在しうることに注意されたい . これにより、アイテムの属性を変化させて未知のアイテムを生成し、そのアイテム絶対順序中の順位によって属性の影響を解析するといったことも可能になる .

LOE の学習段階での目的は、未整列のアイテム集合 $\{I\}_U$ の推定順序 \hat{O}_U を求める規則を、訓練用の事例集合から獲得することである . ただし、 $\{I\}_U$ は未整列だが、集合中のアイテムの属性値は既知とする .

真の順序と推定順序がどれだけ類似しているかの評価に、スピアマンの順位相関係数 (Spearman's Rank Correlation Coefficient, ρ 係数と略す) [4] を用いる . これは、各アイテムの二つの順序における順位の相関係数である . 特に同順位が無い場合、アイテム集合 $\{I\}$ の順序 O^1 と O^2 の ρ 係数は次式で簡単に計算できる .

$$1 - \frac{6 \times \sum_{I^x \in \{I\}} (r(O^1, x) - r(O^2, x))^2}{(\#I)^3 - \#I}$$

ただし、順位 $r(O, x)$ は、順序 O でアイテム I^x が $r(O, x)$ 番目に現れることを示す . この係数は二つの順序が完全に一致するときのみ 1 になり、完全に逆になるとき -1 になる .

4 LOE タスクの解法

LOE タスクの解法は次の二種類に分類できる .

分類手法を用いた解法 : Cohen の方法に類似した方法で、訓練事例をアイテムの対に分解し、評価関数 $\text{PREF}(I^x, I^y)$ を推定する . この評価関数を用いて、未整列のアイテム集合の推定順序を求める .

回帰手法を用いた解法 : 訓練事例中の順序を一つの全順序にまとめ、この全順序から回帰手法によってアイテムの順位を推定する関数を求める . この関数で推定した順位に従ってアイテムを整列する .

4.1 分類手法を用いた LOE の解法

図 2 に、分類手法を用いた LOE 解法の概要を示す . 学習段階は L1 と L2 で、整列段階は S1 で構成される .

ステップ L1 では、事例 $(\{I\}_i, O_i^*)$ のアイテム集合 $\{I\}_i$ から、順序 O_i^* で I^x が I^y より前にあるような全てのアイテム対 (I^x, I^y) を取り出す . 例えば、順序 $I^3 > I^1 > I^2$ からは、 (I^3, I^1) 、 (I^3, I^2) 、 (I^1, I^2) の三つのアイテム対を取り出す . EX 中の全ての事例からこのようなアイテム対を取り出して、その集合全体を P で表す .

ステップ L2 では、この P から評価関数 $\text{PREF}(I^x, I^y)$ を求める . この関数は、属性ベクトル $A(I^x)$ と $A(I^y)$ から、絶対順序で I^x が I^y より前にある可能性を示すもので、これを単純ベイズ分類器 [7] の手法を用いて求める .

$$\begin{aligned} \text{PREF}(I^x, I^y) &= \Pr[I^x > I^y | A(I^x), A(I^y)] \\ &= \frac{\Pr[A(I^x), A(I^y) | I^x > I^y]}{\Pr[A(I^x), A(I^y) | I^x > I^y] + \Pr[A(I^x), A(I^y) | I^y > I^x]} \\ \Pr[A(I^x), A(I^y) | I^x > I^y] &\approx \prod_{s=1}^{\#A} \Pr[a^s(I^x), a^s(I^y) | I^x > I^y] \end{aligned}$$

ただし、 $\Pr[I^x > I^y] = \Pr[I^y > I^x] = 1/2$ を仮定した . $\Pr[a^s(I^x), a^s(I^y) | I^x > I^y]$ には、Dirichlet 分布を

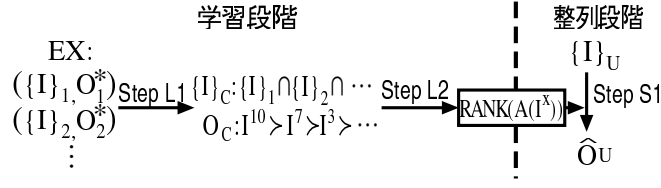


図 3: 回帰手法を用いた LOE 解法の概要

事前分布とした次式のベイズ推定量を用いた .

$$\frac{\#(a^s(I^x), a^s(I^y)) + 1 / (\#a^s)^2}{\#P + 1}$$

ただし, $\#(a^s(I^x), a^s(I^y))$ は, $a^s(I^x) = a^s(I^z)$ かつ $a^s(I^y) = a^s(I^w)$ を満たすような P 中のアイテム対 (I^z, I^w) の数で, $\#P$ は P 中の対の数 .

整列段階では, $\text{PREF}(I^x, I^y)$ を用いて $\{I\}_U$ の真の順序を推定する . ステップ S1 では次の 2 種類の方法を用いた .

PREF 関数の積を用いる方法 (PC 法): この方法は, 次の評価関数を最大化する順序を求める方法である .

$$\prod_{x,y:I^x \succ I^y} \text{PREF}(I^x, I^y) \quad (2)$$

$\#I_U$ が大きいときには計算量の問題のため, 最適解は求められないので, 欲張り法により, 最も前にあると推定されるものから一つずつ推定順序に加える次のアルゴリズムを用いる .

- 1) $\hat{O}^{(0)} := \emptyset, \{I\}^{(0)} := \{I\}_U, t := 0$
- 2) $I^x := \text{argmax}_x \prod_{y:I^y \in \{I\}^{(t)}, x \neq y} \text{PREF}(I^x, I^y)$
- 3) $\hat{O}^{(t+1)} := \hat{O}^{(t)} \succ I^x, \{I\}^{(t+1)} := \{I\}^{(t)} - I^x$
- 4) **if** $\{I\}^{(t+1)} = \emptyset$ **then** output $\hat{O}^{(t+1)}$ **as** \hat{O}
else $t := t + 1$, goto step 2

この方法は, Cohen らの欲張り法による方法 [1] が式 (1) のように関数 PREF の和を最大化するのに対し, 式 (2) のように積を最大化している点のみが異なる . 文献 [3] の実験では, これら二つの性能・特性には差がなく, さらに, PC 法には理論的な優位性があるため, ここでは PC 法のみで実験を行う .

Thurstone の一対比較法を用いる方法 (TC 法):

この方法は, Thurstone の比較判断の法則 (Law of Comparative Judgement)[5] を用いた一対比較法 [11] によって, アイテムを整列する . この方法は, $\{I\}_U$ 中の各アイテム I^x について次式の値を求め, 降順に整列する .

$$\sum_{I^y \in \{I\}_U, I^x \neq I^y} \Phi^{-1}(\text{PREF}(I^x, I^y))$$

ただし, $\Phi(\cdot)$ は平均 0, 標準偏差 1 の正規分布の分布関数 .

4.2 回帰手法を用いた LOE の解法

図 3 は, 回帰手法を用いた LOE 解法の概要である . 学習段階は L1 と L2 で, 整列段階は S1 で構成される .

ステップ L1 では, 事例集合中の全てのアイテム集合を一つにまとめたアイテム集合 $\{I\}_C$ を生成する . そして, 事例中の順序 O_i とできるだけ整合性のある, 集合 $\{I\}_C$ の結合順序 O_C を求める . この順序を求めるために, 前節の方法で順序対集合 P を生成し, 次の評価関数を求める .

$$\begin{aligned} \text{PREF}'(I^x, I^y) &= \Pr[I^x \succ I^y] \\ &= \frac{\#(I^x, I^y) + 0.5}{\#(I^x, I^y) + \#(I^y, I^x) + 1} \end{aligned}$$

ただし, $\#(I^x, I^y)$ は, アイテム対 (I^x, I^y) の数 . この PREF' を, 前節の PREF の代りに用いて, PC 法と TC 法の整列段階の手法を用いて O_C を求める . PC 法の整列方法を用いる場合を PR 法, TC 法の方法を用いる場合を TR 法と呼ぶ . 前節の評価関数 PREF とは異なり, PREF' は属性値を参照していないことに注意されたい .

ステップ L2 では、アイテムを表す属性ベクトルから、絶対順序中でアイテムがどれだけ前にあるかを予測する順位関数 $\text{RANK}(A(I^x))$ を求める。この関数は、カテゴリ属性をダミー変数で表した回帰分析（数量化 I 類ともいう）を用いて求める。属性 $a^s(I^x)$ が $v_1^s \cdots v_{\#a^s}^s$ の値をとるとき、ダミー変数は、 $\#a^s - 1$ 個の変数で属性値を表す。第 1 の属性値 v_1^s は全てのダミー変数を 0 にして表し、第 2 以降の属性値 v_t^s はダミー変数の $t-1$ 番目のみが 1 で、他を全て 0 にして表す。 $A(I^x)$ の全ての要素 $a^1(I^x) \cdots a^{\#A}(I^x)$ をダミー変数で表し、これらを連結したベクトルを $d(A(I^x))$ と表記する。次に、順序 O_C 中で i 番目のアイテムの $d(A(I^i))$ を第 i 行とする行列 D を生成して次のベクトル X を計算する。

$$X^T = (D^T D)^{-1} D^T (1, \dots, \#I_C)^T$$

このベクトルを用いて、順位関数は次式で表される。

$$\text{RANK}(A(I^x)) = X^T d(A(I^x))$$

整列段階のステップ S1 では、未整列のアイテム集合 $\{I\}_U$ の各アイテム I^x について、 $\text{RANK}(A(I^x))$ を求め、この値の順に整列することで推定順序 \hat{O}_U が求められる。

5 人工データを用いた実験

ここでは、4 節の手法を人工データに適用して各手法の特徴を解析する。

5.1 人工データの生成方法

実験に用いた人工データについて述べる。データ型は属性数と属性値数で決まり、属性数を 3, 5, 7 と、属性値数を 3, 4, 5 と変化させ、9 種類のデータ型を用いた。アイテムの絶対順序は線形のスコア関数を用いて定めた。例えば、アイテム I^7 の属性値ベクトルが $A(I^7) = (v_1^1, v_2^2, v_3^3)$ であるとき、スコアは、重み $w(a^s)$ や $w(v_t^s)$ を用いて、 $w(a^1)w(v_1^1) + w(a^2)w(v_2^2) + w(a^3)w(v_3^3)$ となる。絶対順序は、このスコアの順にアイテムを

並べたものである。9 種類のデータ型それぞれについて、0 と 1 の間の重みをランダムに生成して、10 セットの異なる絶対順序を定めた。すなわち、90 種類のアイテム全集合と絶対順序の対を生成した。さらに、これらの全ての対について、アイテム数 $\#I_i$ が、3, 5, 及び 10 に、事例数 $\#EX$ を 10, 30, 及び 50 にして、9 種類の事例集合を生成した。こうして、全部で 810 個の事例集合を生成した。

この 810 個の事例集合に、分類手法を用いた二つの解法 (PC と TC) と回帰手法を用いた二つの解法 (PR と TR) の計 4 種類の解法を適用した。評価方法には、 $\#EX$ 分割の交叉確認法である leave-one-out (LVO) 法を用いた。これは、最初の事例、 $(\{I\}_1, O_1^*)$ を事例集合から取り出し、残りの事例を用いて整列用の規則を獲得する。最初に取り出した事例のアイテム集合 $\{I\}_1$ に規則を適用して推定順序 \hat{O}_1 を求める。この推定順序の損失、すなわち、3 節の ρ 係数を求める。損失は推定順序と事例中の真の順序の間で求めるのが一般的だが、ここでは人工データに対する実験で絶対順序が分かっているので、絶対順序に対する損失を求める。この手続きを事例集合中の全ての事例について繰り返し、その平均をもってどれだけ適切に整列されているかを測る。

5.2 実験結果

ここでは、ノイズの無い場合、すなわち、事例中の真の順序は絶対順序と無矛盾である場合の実験を行った。

表 1 には、LVO で求めた ρ 係数の平均を、全ての事例の場合と、アイテム数が同じものごとに分けた場合、及び、事例数が同じものごとに分けた場合とで示した。比較判断の方法を用いた TC や TR 法でも、文献 [3] と同様にアイテム数の増加に伴い、また、事例集合が多くなるに伴い、 ρ 係数が 1 に近づき、よりよい推定がなされている。詳細な結果は省略するが、アイテム全集合の要素数が少ない方が推定精度の良くなる傾向があった。

表 1 についてより詳細に検討する。アイテム

表 1: 各手法による ρ 係数の平均

	ALL	アイテム集合の大きさ			事例集合の大きさ		
		3	5	10	10	30	50
PC	0.808	0.667	0.825	0.932	0.715	0.835	0.874
TC	0.810	0.668	0.831	0.931	0.718	0.836	0.876
PR	0.802	0.617	0.837	0.950	0.698	0.838	0.869
TR	0.807	0.616	0.847	0.958	0.705	0.845	0.872

数が $\#I$ の二つのランダムな順序の間の ρ 係数について、次の t 値は自由度 $\#I - 2$ の t 分布に近似的に従うことが知られている [4] .

$$t = \rho \sqrt{\frac{\#I - 2}{1 - \rho^2}}$$

この式から、 $\#I$ が 3, 5, 及び 10 のとき ρ 係数が 0.9995, 0.9343, 及び 0.7155 以上であれば危険率 1% で有意な相関がある。厳密には、標本分布を考慮する必要があるが、簡単に ρ 係数の平均値と比較すると、アイテム数が 10 の場合は絶対順序と無矛盾な順序が推定されている。アイテム数が 5 や 3 の場合は明確な相関があるとは断定できないが、アイテム数が 5 の場合は、危険率を 5% まで緩和すれば (0.8054 以上であればよい) 有意な相関がある。

各手法の推定精度を比較するため、対応のある t 検定を行い、その t 値を表 2 に示した。第 1 列の X-Y は手法 X による ρ 係数から手法 Y による ρ 係数を引いた場合の t 値を示し、 $<t_\alpha$ は ($>t_{1-\alpha}$) は、危険率 α で手法 X の ρ 係数の方が有意に小さい (大きい) ことを示す。

PREF 関数の積を用いる PC や PR と、一対比較法を用いる TC や TR との比較結果 (PC-TC と PR-TR の結果) について述べる。PC と TC の比較ではそれほど明確ではないが、PR と TR の比較では、一対比較法を用いる方が有意に優れている。これは Thurstone の一対比較法が根拠とする比較判断の法則も、Spearman の ρ 係数のどちらも、実数直線上の等間隔にアイテムを配置するという暗黙の仮定を用いて順序全体を評価する。それに対し、PREF 関数の積を用いる方法では、順序全体ではなく、アイテムの 2 項間関係という

局所的な情報をできるだけ保存しようとする。これらの違いのため、一対比較法の方が良い推定ができると考える。

この仮説は次の実験結果によっても確かめられる。4.1 節の方法は欲張り法なので、関数 PREF の積を最大にする順序が得られているわけではない。そこで、 $\#I$ がたかだか 10 と比較的小さく網羅的な探索ができるので、PREF の積を最大にする順序を求めた。 ρ 係数の平均は 0.806 と、欲張り法を用いた PC 法の結果より悪くなり、その差は危険率 1% で有意 (t 値は -2.9306) であった。このように、PREF の積という評価関数を最大化しても、 ρ 係数が必ず最大化されるわけではない。すなわち、2 項間の順序関係をできるだけ保存することと、全体として適切な順序を求めることとの相違を示している。また、Cohen らの方法と等価な PREF の和の最大化についても同じ実験をしたが、この実験でも ρ 係数は 0.808 から 0.805 に低下し、 t 値は -2.7915 でその差は有意だった。これらの結果は、2 節で述べた、Cohen らの研究と LOE タスクの目標の相違をいっそう明らかにするものである。

次に、分類手法を用いた PC や TC と、回帰手法を用いた PR や TR とを比較する (PC-PR と TC-TR について)。表 2 の PC-PR と TC-TR のいずれの結果でも全体では有意な差はない。しかし、 $\#I$ が小さい 3 の場合は分類手法による方が良く、 $\#I$ が増えると逆に回帰手法による方が良くなる。これは、事例集合の大きさが変化しても、分類手法と回帰手法の間で相対的な優劣があまり変化しなかったことと比べて特徴的である。文献 [3] の結果でもこの傾向は見られたが、今回

表 2: ρ 係数の平均の差の t 値

	ALL	アイテム集合の大きさ		
		3	5	10
PC-TC	-1.731	-0.227	-3.818 < $t_{0.01}$	1.238
PR-TR	-3.559 < $t_{0.01}$	0.264	-5.001 < $t_{0.01}$	-7.780 < $t_{0.01}$
PC-PR	1.463	4.425 > $t_{0.99}$	-2.355 < $t_{0.01}$	-8.502 < $t_{0.01}$
TC-TR	0.650	4.707 > $t_{0.99}$	-3.351 < $t_{0.01}$	-15.518 < $t_{0.01}$

	事例集合の大きさ		
	10	30	50
PC-TC	-1.011	-0.981	-1.563
PR-TR	-1.872 < $t_{0.05}$	-3.173 < $t_{0.01}$	-1.469
PC-PR	1.458	-0.629	1.557
TC-TR	1.175	-1.733 < $t_{0.05}$	1.215

表 3: 回帰手法の結合順序 O_C と絶対順序の間の ρ 係数の平均

	ALL	アイテム集合の大きさ			事例集合の大きさ		
		3	5	10	10	30	50
PR	0.792	0.689	0.803	0.883	0.796	0.787	0.793
TR	0.820	0.708	0.833	0.920	0.813	0.819	0.829

はこの理由について調査した。

回帰手法の途中で生成される結合順序 O_C と絶対順序の間の ρ 係数の平均値を、表 3 に示した。事例集合の数が増えてもこの ρ 係数は変わらないか微増する程度である。それに対して、アイテム集合の大きさ $\#I$ が増大するにつれて、 ρ 係数は急激に増加している。これは、違う順序の例にあるアイテム間では順序関係の推移性が保たれていないが、同じ順序の中で同時に与えられるアイテムの間では推移性が保たれている。同時に提示されないアイテムの順序の推移性は、同時に提示される場合よりも高くなることは、文献 [2] でも指摘されている。この推移性が保存されたアイテムの増加によって、精度の高い結合順序が得られるようになり、PR や TR 法の精度が急激に改善され、PC や TC 法を上回るようになると考えられる。

文献 [3] の実験と同じノイズがある場合につい

て実験した。表 4 には、アイテムの属性値が変化するノイズの影響を調査するため、訓練事例のアイテムの属性値が 0%–10% の確率で他の値に変わった場合の結果を示した。また、表 5 には、順序中で隣接するアイテムが入れ替わるノイズの影響を調査するため、訓練事例の順序で隣接するアイテムの 0%–10% が入れ替わった場合の結果を示した。どちらも、 ρ 係数の平均と、括弧内にノイズが無い場合に対する相対値を示した。全体的に、ノイズの影響による性能の低下の度合いは、手法によってあまり差がない。ただし、属性ノイズが 10% になると、回帰手法用いた場合の方が、分類手法を用いた場合と比べてやや大きな性能低下が見られた。ノイズの割合の意味はノイズ種類によって異なるため、正確な評価は難しいが、属性ノイズの方が影響が大きいようである。

表 4: 属性値が変化するノイズがある場合の ρ 係数の平均

	0%	1%	3%	5%	10%
PC	0.808	0.802 (0.9929)	0.795 (0.9846)	0.793 (0.9813)	0.778 (0.9630)
TC	0.810	0.805 (0.9939)	0.799 (0.9867)	0.796 (0.9822)	0.783 (0.9671)
PR	0.802	0.797 (0.9948)	0.794 (0.9910)	0.786 (0.9810)	0.765 (0.9541)
TR	0.807	0.802 (0.9932)	0.799 (0.9898)	0.790 (0.9789)	0.768 (0.9509)

括弧内はノイズがない場合に対する相対値

表 5: 順序が入れ替わるノイズがある場合の ρ 係数の平均

	0%	1%	3%	5%	10%
PC	0.808	0.806 (0.9980)	0.802 (0.9921)	0.799 (0.9895)	0.790 (0.9773)
TC	0.810	0.808 (0.9979)	0.804 (0.9928)	0.800 (0.9880)	0.791 (0.9761)
PR	0.802	0.801 (1.0000)	0.795 (0.9919)	0.794 (0.9904)	0.783 (0.9768)
TR	0.807	0.807 (1.0000)	0.804 (0.9963)	0.800 (0.9915)	0.790 (0.9782)

括弧内はノイズがない場合に対する相対値

6 嗜好調査データを用いた実験

ここでは、4 節の手法を、寿司ネタの嗜好について質問した小規模のアンケート調査の結果に適用した。

6.1 実験対象と質問項目

次の 10 種の寿司の嗜好を調査するアンケートを 52 人に行った。

とろ	いくら	うに	こはだ
はまち	いか	たい	サラダ巻
鉄火巻	かっぱ巻		

この調査対象を選んだ理由は、感覚的な質問であるため順序による提示の検証に適し、プライバシー上の問題が少なく被験者の収集が容易であるからである。前節の実験から、順序の差の有意性の検証には 10 個程度のアイテムが必要で、アイテム数が多いと被験者の負担が増えるためアイテム数は 10 個とした。図 4 のように Web ブラウザによって回答を得た。質問項目は以下のとおりである。

Figure 4 shows a web browser window with a survey titled "うちじゃ寿司ネタは10種類だ!". The instructions state: "質問は、このページを含めて6ページです。" and "お客さんが好きな順番に1~10番まで番号をつけてくんねえかな。" followed by "おっと、一番好きなネタが1番ってとこに気をつけておくんね。". Below the text is a table for ranking 10 types of sushi: ハマチ, コハダ, イカ, トロ, 鉄火巻, ウニ, サラダ巻, タイ, イクラ, カっぱ巻. Each row has 10 columns numbered 1 to 10. The 'Hamachi' row has a radio button selected in column 1. At the bottom, there is a button that says "終わったら戻してください".

図 4: 順序の入力インターフェース

順位法による嗜好の調査 好きなものから嫌いなものへ順に整列するよう依頼した。図 4 のように、10 種の寿司について好きなものから順に順位をつけるよう依頼した。同順位はつけられないようにした。アイテムの提示順序の影響を排除するため、被験者ごとに提示順序を変更した。参考までに、嗜好の順序と提示順序の ρ 係数は 0.0061 で無相関であった。

寿司自体についての質問 各寿司に対して、以下の三種類の質問に 3 段階で回答するよう依頼した。
 こってり感 あっさり — 中間 — こってり
 価格 安価 — 中間 — 高価
 希少性 珍しい — 中間 — 定番

SD 法による嗜好の調査 それぞれの寿司に対して、好き、中間、嫌いの三段階で回答するよう依頼した。

その他 性別、任意回答で年齢を質問したが、今回はこれらのデータは解析しなかった。回答時間は平均 325 秒（最短 126 秒，最長 2418 秒）であった。

寿司の嗜好について順位法と SD 法で質問した結果を比較した。同一被験者の回答について、全てのアイテム対のうち、順位法と SD 法による回答に矛盾がある割合について調べると、4.8% であった。矛盾とは、 $I^x \succ I^y$ と順位法では回答したが、SD 法では I^x を中間、 I^y を好きと答えた場合などである。全アイテム対について無矛盾な回答をした被験者の割合は $24/52 \approx 46.2\%$ で約半数であった。さらに、矛盾のある回答をした被験者のみの、全アイテム対についての矛盾の割合は 9.0% で、最大 37.8% もの矛盾のある被験者がいた。このように、順位法は、SD 法とは異なった側面の情報を得ることができる。ことがわかる。

6.2 LCE 手法の適用と実験結果

上記の調査結果に LCE 手法を適用した。アイテムの属性は 5 種類で、最初の 3 属性は寿司自体についての質問結果をもとに定めた。こってり

PC-TC	PR-TR	PC-PR	TC-TR
0.181	0.000	-0.100	-0.252

表 7: 各手法の ρ 係数の差の t 値

感、価格、及び、希少性の回答の分布を求め、この分布に基づいて、それぞれ、5、5、及び、3 種類の属性値をとるようにした。4 番目の属性は、にぎり、ぐんかん、巻といった形状で、5 番目の属性はネタの種類を魚、野菜、その他に分類した。5 種類の属性の全ての組み合わせ $31(=2^5 - 1)$ 通りについて、4 節の 4 種類の LCE 手法を適用した。各手法について、LVO テストによる ρ 係数の平均が最大であった結果を表 6 に、 t 検定の t 値を表 7 に示した。ただし、5 節と異なり、実データでは絶対順序は不明なので、事例中の真の順序との ρ 係数を求めた。 t 検定の結果、どの手法の間にも有意な差は見られず、前節の人工データに対する結果と異なっている。これは、アイテム総数が 10 個であるのに対して 52 個と十分な事例数があり、また、訓練事例中で観測されないアイテムが存在しないため、どの方法でも十分に学習ができたものと考えられる。このことは、各手法で獲得された順序には、隣接したアイテムの置換がある程度の差しかなく、ほぼ同じ解になっていることから確かめられる。

次に TR 法の結果の詳細を表 8 に示す。TR 法は 52 回の試行のうち、最も多く出力された順序は 1 行目のもので、2 と 3 行目は下線で示したネタに 1 行目の順序に対する置換がある。“とろ”が最も好まれ、“かっぱ巻”には人気がないといったことがわかる。さらに、この順序と各被験者の順序との ρ 係数を求めた。負の相関をもつ事例は 6/52 程度にとどまり、危険率 10% で同じ順序とみなせる事例は 27/52、さらに、1% でも 20/52 もあった。LOE は全体を一つの絶対順序で表す大まかなモデルではあるが、被験者間で共通している部分の情報をうまく獲得できたと考えられる。その他、使用した属性と ρ 係数の関係から、価格という属性が重要であることや、著者の順序と表 8 の 1 行目の順序の ρ 係数が 0.758 もあり、著者の嗜好が一般的なものであるといった解析なども可

表 6: 寿司の嗜好調査データに対する ρ 係数の平均値

	PC	TC	PR	TR
ρ 係数の平均	0.454	0.451	0.455	0.455
採用した属性	{1,5}	{2,4}	{2,4}	{2,3,4}

表 8: TR 法によって得られた推定順序の例

回数	推定順序
40	とろ > はまち > たい > いくら > うに > こはだ > 鉄火巻 > いか > サラダ巻 > かつぱ巻
11	とろ > はまち > たい > いくら > うに > 鉄火巻 > こはだ > いか > サラダ巻 > かつぱ巻
1	とろ > はまち > たい > いくら > うに > こはだ > <u>いか</u> > <u>鉄火巻</u> > サラダ巻 > かつぱ巻

能である .

7 まとめ

この論文では、著者が提案した学習タスクである順序例からの学習に、判断比較の法則を用いた方法を導入し、推定精度を改善した。さらに、嗜好調査データへの適用によって、実データにおいても LOE が有効な推定・解析手法であることを示した。

今後は、LOE をより大規模な実問題に適用し、 ρ 係数を直接扱う学習手法の開発を行う予定である。

参考文献

- [1] W. W. Cohen, R. E. Schapire, and Y. Singer. Learning to order things. *J. of Artificial Intelligence Research*, Vol. 10, pp. 243–270, 1999.
- [2] R. H. Hohle. An empirical evaluation and comparison of two models for discriminability scales. *J. of Mathematical Psychology*, Vol. 3, pp. 173–183, 1966.
- [3] 神島敏弘, 赤穂昭太郎. 順序例からの学習. 人工知能学会全国大会 (第 16 回) 論文集, 2002. (3C1-01).
- [4] M. Kendall and J. D. Gibbons. *Rank Correlation Methods*. Oxford University Press, fifth edition, 1990.
- [5] L.L.Thurstone. A law of comparative judgment. *Psychological Review*, Vol. 34, pp. 273–286, 1927.
- [6] H. Mannila and C. Meek. Global partial orders from sequential data. In *Proc. of The 6th Int. Conf. on Knowledge Discovery and Data Mining*, pp. 161–168, 2000.
- [7] T. M. Mitchell. *Machine Learning*. The McGraw-Hill Companies, 1997.
- [8] 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.
- [9] 流郷達人, 金子俊一, 五十嵐悟, 宮本敦, 亀和田俊一. 順位相関に基づくロバスト画像照合法とその地下透水係数推定への応用. 電子情報通信学会技術研究報告, Vol. PRMU 2001-26, pp. 47–52, 2001.
- [10] Y. Sai, Y. Y. Yao, and N. Zhong. Data analysis and mining in ordered information tables. In *Proc. of the IEEE Int. Conf. on Data Mining*, pp. 497–504, 2001.
- [11] 佐藤信. 統計的官能検査. 日科技連, 1985.