順位相関係数を保証するノイズありソートアルゴリズム

On Incomplete Noisy Sorting

小宮山 純平 *1
Junpei Komiyama

*1東京大学生産技術研究所
The Institute of Industrial Science, the University of Tokyo.

This paper discusses the way to sort items by the noisy feedback of pairwise comparisons. An algorithm that partially sorts items are provided, and its empirical performance is shown.

1. はじめに

要素の良さに順序をつけたい場合は多い．経済学の 1 分野である社会選好理論 [Arr63] は個人の選好をまとめて統一する方法を研究する学問である．情報科学の分野でも、要素の順序付けは重要である．例えば、推薦システム、検索エンジン、オンライン広告などは、ユーザにどのような要素が好まれるかの選好の研究だと考えることができる．

要素に順序をつけるための最も簡単な方法は、ペア比較をくり返すことである．本研究では、動的に要素がクエリ可能な下での、ペア比較による要素の順序の決定を考える．つまり、順序付けのアルゴリズムが、順序を効率的に決定するために次に比較するペアを選べるというものである．これは、次のような応用を想定している．

- 推薦アルゴリズム、検索エンジンなどのエンジンの比較を、ユーザに複数の結果を見てもらうことにより比較する [CJRY12]
- 将棋や囲碁などのゲーム AI プログラムがいくつかあったときに、どのプログラムが強いかをプログラム同士の対戦の勝敗により決定する．

ペア数は要素数 K に対して $K(K-1)/2 = O(K^2)$ であり、K が大きい場合のランダムな比較はスケールしない．しかし、要素の間に順序関係を仮定し、効率の良い比較スケジュールを立てると、$O(K \log K)$ の比較で順序を決定することができる．ここで、ペア (i,j) の選好を $i \succ j$ と書いたときに、要素 i が要素 j より優れると呼び、この場合 $\mu_{i,j} > 1/2$ を要素 i の要素 j に対する勝率と言うことにする．

関連研究: 本研究は、要素間の順序をつけるという意味でソートアルゴリズムの研究である．クイックソート、マージソートなどの通常のソートアルゴリズムは、ペアを比較するとその順序関係が得られる状況を仮定している．一方で、我々のケースはフィードバックがノイズを含むということを仮定している．つまり、ペア (i,j) を比較した場合、より優れたほうが 0.5 を超える勝率を持つものの、フィードバックは確率的な要素を伴う．このような確率的フィードバック下での順序付けは、ノイズありソーティング (noisy sorting) として 1980-90 年代に研究された [FRPU94]．ノイズありソーティングでは、ある定数 $\lambda \in (0, 0.5)$ が存在し、$\mu_{i,j} > 0.5 + \lambda$ を仮定し、この λ に応じた比較回数を抑えることが多い．一方、近年の機械学習では、ノイズありソーティングとやや離れた文脈で要素間の順序付けを行っている．比較バンディット問題 [YBKJ12, KHKN15] は、検索エンジンの比較をモチベーションとし、オンラインの設定で複数の要素の優劣を探す問題をモデル化している．関連して、オンライン順序決定（online rank elicitation）や PAC 順序決定（Probability-Approximately-Correct rank elicitation）と呼ばれる分野で、ノイズのあるソーティングと同じく要素間の順序付けが行われている [BSH14, SBPH15]．

これらの関連研究は、それぞれの仮定下で順序の完全な復元を可能とするアルゴリズムを提案している．しかし、要素数が多い場合、順序の完全な復元は $O(K \log K)$ で終わるとはいえども、現実的には非常に高コストである．これは、ペア (i,j) の勝率が 0.5 に近い場合、どちらが良いかの決定に本質的に大きな比較回数がかかるからである．つまり、大偏差原理 [CT06] により、ペア (i,j) の順序付けは $O(1/(\mu_{i,j} - 0.5)^2)$ 回の比較回数が本質的に必要である．これは、例えば、非常に実力が拮抗した AI プログラム間の優劣をつけるためにはたくさんの対戦が必要だということから理解可能であろう．一方で、多くの応用では要素に対する大雑把な順序づけが十分であろう．例えば、テニスの試合で Top-100 の選手の勢力図を決めるためには、フェデラーやマレー、錦織などの Top-10 に入るであろうプレイヤー、Top-30 程度のプレイヤー、Top-50 以下のプレイヤー、ぐらいの大雑把な区分けは、多くの人にとって必要十分ではないだろうか．

このようなことを考慮に入れた上で、本研究は「大雑把な順序付け」を少ない比較回数、理論的な保証ありで行うことを目指す．節 2.で問題設定と「大雑把な順序付け」の基準であるケンドールの順位相関係数を行う．節 3.では順序づけのアルゴリズムを提案する．節 4.では数値実験により提案アルゴリズムの良さを評価する．おわりに，節 5.で本研究を総括する．

2. 問題設定

K 個のアイテム集合 $[K] := \{1, 2, \ldots, K\}$ をソートすることを考える．要素は全順序を持つと考える: つまり，アイテムのペア i, j に対し，$i \succ j$ ならペアを比較したときの i の勝率 $\mu_{i,j} \in (0, 1)$ が $\mu_{i,j} > 1/2$ である．一般性を失うことなく，アイテムの順序は $1 \succ 2 \succ \cdots \succ K$ であるとする．もちろん，アルゴリズムはこの順序を知らず，ペアの比較を通じて順序を推定するのがアルゴリズムの目的となる．各時間ステップ $t = 1, 2, \ldots$ に，アルゴリズムはアイテムのペアを選び，どちらが好まれるかというフィードバックを受け取る．各時間ステップの終了時に，アルゴリズムは比較の終了を選び，

連絡先: 小宮山 純平, 東京大学生産技術研究所, 〒 153-8505 東京都目黒区駒場 4-6-1, junpei@komiyama.info

予測順位[*1] $\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K)$ を出力して良い．そうでないなら，次の時間ステップに進む．本研究では，任意の順序と勝率 $\{\mu_{i,j}\}$ に対して有限時間で確率 1 で終了するアルゴリズムのみを扱う．

2.1 ケンドールの順位相関係数

ケンドールの順位相関係数は，2 つの順序がどの程度近いかを測る指標の一つであり，一様ランダムにペアを選択したときに，2 つの順序間でペアの順序が一致する割合で定義される．本研究では真の順序 $(1, 2, \ldots, K)$ とアルゴリズムの予想した順序 $(\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K))$ の間の順位相関係数：

$$\frac{1}{K(K-1)/2} \sum_{i<j} 2(\mathbf{1}(\boldsymbol{r}(j) - \boldsymbol{r}(i)) - 0.5).$$

を考える．ここで，$\mathbf{1}(X)$ は $X > 0$ なら 1，そうでないなら 0 とする．順位相関係数は，真の順序と予想した順序が完全に一致している場合に 1 になり，完全に一致していない場合に 0 になる．所与の $\delta \in (0,1)$, $\tau \in [0,1]$ に対し，$\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K)$ の順位相関係数が $\tau \in (0,1)$ を超える確率が $1-\delta$ 以上であるとき，アルゴリズムは δ-τ 調和であるという．本論文では，δ-τ 調和であるアルゴリズムのうち，終了までの平均時間ステップ（＝比較回数）が少ないものを探す．τ の値に関して制限は求めないが，実用的には 1 にある程度近い値を想定している．

3. アルゴリズム

本節では，最初に，境界通過確率に基づいた統計的優位性を持ったペアの比較を提案し，次にそれを利用した δ-τ 調和なソートアルゴリズムである Iterative Quick Sort (IQS) を提案する．

3.1 ペア順序の判定と統計検定

本小節は主に統計検定に親しい方に対して境界通過確率の導入について解説しているが，統計検定にそれほど興味がない読者は飛ばさしていただいても構わない．ペア (i,j) の順序の決定は統計検定と考えることができる．帰無仮説は

$$H_0 : \mu_{i,j} = 1/2$$

であり，対立仮説は

$$H_i : \mu_{i,j} > 1/2$$
$$H_j : \mu_{i,j} < 1/2$$

である．通常の仮設検定は，固定した回数の比較を行った後で帰無仮説を棄却することを目指すが，本研究の設定では動的に比較回数を増やすことが可能であり，このような場合は通常の仮設検定と異なる確率過程の考慮が必要となる．このことを考慮し，節 3.2では，境界通過確率に基づいた停止時刻を説明する．

3.2 境界通過確率

境界通過確率は，ランダムなサンプル系列が与えられたときに，その平均値[*2] がサンプル系列を通じて一貫して持つ性質である．今回の場合はあるペアの勝率が 0.5 より小さいのか大きいのかの決定が焦点になるが，各ラウンドでの平均勝率がどの程度であればこの決定ができるかは，まさに境界通過確率の一つの応用先である．ペア i,j に対し，$X_1, X_2, \ldots, X_n$ をペア間比較結果を示す確率変数とする．つまり，k 回目の比較でアイテム i が好まれた場合に $X_k = 1$，逆に j が好まれれば $X_k = 0$ である．また，$\hat{\mu}_{i,j} = (1/n)\sum_{k=1}^{n} X_k$ をその平均値とする．$\hat{\mu}_{i,j}$ が 0.5 より大幅に大きい場合は i のほうが良いアイテムだと想像できるが，これの厳密な境界を境界通過確率によって考えたい．つまり，$\hat{\mu}_{i,j} = \hat{\mu}_{i,j}(n)$ の関数としての停止時刻 τ で，誤って停止してしまう確率を δ 以下にするものを考えたい：

$$\mathbb{P}[\tau < +\infty | \mu_n = 1/2] \leq \delta \quad (1)$$

境界通過確率は，70 年代に Lai, Robbins, Siegmund らの著名な研究者によって調べられたが [Sie85]，彼らの手法は主に X_k が（今回と異なり）正規分布のときに正規分布を確率過程に埋め込む方向について研究されている．サンプル数が多い時に $\hat{\mu}_{i,j}$ は正規分布するものの，有限の値では今回の比較のような Bernoulli 分布の期待値を正規分布近似するのはリスクを伴う可能性がある．そのため，正規分布近似式の境界通過確率は今回扱わない．1990 年代において境界通過確率は一度研究が減速してしまったが，A/B テストやオンライン広告での応用 [JKPW17, Mai17] などを契機に近年再注目を浴びている概念である．本研究では，比較的使いやすい境界通過確率として以下のものを用いる．

$$\tau_{\mathrm{u}} = \tau_{\mathrm{u}}(\delta) = \min_{n \geq 10} \left\{ n d_{\mathrm{KL}}(\hat{\mu}_{i,j}(n), 1/2) \geq \log\left(\frac{n (\log n)^2 K^2}{4\delta}\right) \right\}.$$

Theorem 1. τ_{u} は境界通過確率の条件式 (1) を満たす．

証明は紙面の都合上省略する．

3.3 IQS アルゴリズム

境界通過確率を利用したアルゴリズムである，IQS (Iterative Quick Sort）アルゴリズム（Algorithm 1) を提案する．IQS

Algorithm 1 Iterative Quick Sort (IQS) algorithm.

順位相関係数 τ_0, 信頼度 δ, アイテム $[K]$.
初期ソート結果 $\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K)$.
for 各フェーズ $p = 1, 2, \ldots$ **do**
  $N_p \leftarrow 4^p$.
  $Q \leftarrow (\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K))$.
  while Q is not empty **do**
    Q の最初の要素 $(r_1, \ldots, r_l)$ を取り出す．
    中央要素をピボットにする：$p \leftarrow r_{(i+j)/2}$.
    ピボット以外の各要素 $k \in (r_1, \ldots, r_l) : k \neq p$ を p と比較する．このとき，このペアを $N_{k,p} := \min(N_p, \tau_{k,p}(\delta/(K(K-1)/2))$ 回比較する．
    $f_1, \ldots, f_{l_f} \leftarrow \{r_i : i \neq p, \hat{\mu}_{i,p} \geq 1/2\}$, $b_1, \ldots, b_{l_b} \leftarrow \{r_i : i \neq p, \hat{\mu}_{i,p} < 1/2\}$.
    ピボットとの比較で要素を並び替える：$(r_1, \ldots, r_l) \leftarrow (f_1, \ldots, f_{l_f}, r_p, b_1, \ldots, b_{l_b})$.
    $l_{\mathrm{fin}} \leftarrow |\{k \in \{f_1, \ldots, f_{l_f}\} : N_{k,p} = \tau_{k,p}\}|$.
    $r_{\mathrm{fin}} \leftarrow |\{k \in \{b_1, \ldots, b_{l_b}\} : N_{k,p} = \tau_{k,p}\}|$.
    $\hat{\tau}_{\mathrm{LB}} \leftarrow \hat{\tau}_{\mathrm{LB}} + \frac{(l_{\mathrm{fin}}+1)(r_{\mathrm{fin}}+1)-1}{(K(K-1)/2)}$.
    $\hat{\tau}_{\mathrm{LB}} \geq \tau_0$ ならば，すぐにアルゴリズムを終了させて予想順序 $(\boldsymbol{r}(1), \boldsymbol{r}(2), \ldots, \boldsymbol{r}(K))$ を出力する．
    If $l_f \geq 2$ then push $(f_1, \ldots, f_{l_f})$ at the end of Q.
    If $l_b \geq 2$ then push $(b_1, \ldots, b_{l_b})$ at the end of Q.
  end while
end for

---

*1 $[K]$ の置換である
*2 もしくは総和にしても等価である

は，大雑把に言うとフェーズごとに最大比較回数を増やしていく反復進化法とクイックソートを混ぜたアルゴリズムである．各フェーズごとにクイックソートが一度行われる．このクイックソートは幅優先探索（つまりキュー Q を利用）である．現在ソートしたい範囲の中央のアイテム r_p とそれ以外の要素との比較を行い，r_p より好まれるものを前に，r_p と比べ好まれないものを r_p より後にするということを再帰的に行う．IQS は境界通過確率に基づいた停止時刻を全フェーズで共有している：つまり，いったんペア (i,j) の比較回数が $\tau_{i,j}$ に達する（つまり，ペア間の順序が自信をもって確定される）と，以降のフェーズでもこの比較結果を活用し，無駄な追加比較はおこなわない．一方で，各フェーズ p で比較数の限界 $N_p := 4^p$ を設けており，停止時刻に達しない場合もこの回数までしか比較しない．停止時刻に達したペアのみ統計的有意性があるため，ピボット p より有意に好まれる要素数は l_{fin} 個，有意に好まれない要素数は r_{fin} 個であり，これらを利用し現在の順位相関係数の下界 $\hat{\tau}_{\text{LB}}$ を構成する．$\hat{\tau}_{\text{LB}}$ が所与の τ_0 に達したらアルゴリズムは現在の予想順序を出力し終了する．なぜこのようなフェーズごとのソートを行うのかというと，以下の理由がある：つまり，各ペアの比較がノイズを含むため，ペア間の比較を境界通過確率に基づいた停止時刻まで行わないとペアの統計的有意性を確保できない．一方，勝率が 0.5 に極めて近いペアを停止時刻まで評価してしまうと，そのペアの比較に大きな時間が食われてしまう．そのため，比較回数に上限を設けた幅優先でソートしやすそうな要素から比較し，そこから確定したペアで構成した順位相関係数の下界を持ち，勝率が 0.5 に近いペアの比較を確定させずに終わらせようというのが目的である．

以下，IQS の持つ理論的な性質を示す：これらの証明は省略する．まず最初に，IQS は順位相関係数に関する保証がある．

Theorem 2. IQS の順位相関係数に関する保証 *IQS (Algorithm 1)* は δ-τ-調和である．

次に気になるのは，IQS の比較回数と計算量である．簡単のため，以下の解析では線形効用モデルを仮定する．つまり，$\mu_{i,j} = 0.5 + O(|i-j|)$ とする．大雑把に言うと，順位が近いペアの勝率が 0.5 に線形に近づく．

Theorem 3. *(IQS の比較回数)* 任意の $\tau_0 < 1$ を考える．任意の δ-1-調和なアルゴリズム $\mathcal{A}$ の終了までの平均的な比較回数を $C_{\mathcal{A}}$ とする．*IQS (Algorithm 1)* アルゴリズムの終了までの比較回数を C_{IQS} とする．このとき，線形効用モデルを仮定すると，

$$\frac{C_{IQS}}{C_{\mathcal{A}}} \to 0$$

である．

つまり，完全にペアを並び替える任意のアルゴリズムに関して，「不完全な」比較アルゴリズムである IQS は，十分大きいアイテム数 K をソートすることを考えると，比較回数がいくらでも小さくなるというサンプル効率性がある．

次に，IQS の計算量について以下の定理を示す．

Theorem 4. *(IQS の計算量) IQS* の確定したペアのうち，最も比較回数が多かったものの比較数を N_c とする．*IQS (Algorithm 1)* は高々$O(K \log K N_c)$ 回の比較しか行わない．また，高々$O(K)$ 程度の空間計算量を持つ．

つまり，$O(K \log K)$ の計算量，$O(K)$ のメモリ消費量を持つ．これらの性質は，IQS が K が大きい場合にスケールするアルゴリズムであることを示している．

4. 実験

提案手法を他の順序付けアルゴリズムと比較するために，以下の実験を行った．提案手法である IQS アルゴリズムを，完全な順序を復元する MallowsMPR [BSW+14], PLPAC-AMPR [SBPH15], FPRU [FRPU94] と比較した．順位相関係数の τ_0 は 0.8, 信頼水準 $\delta = 0.05$ とした．

データセットとして，神嶌らによる 5,000 人規模の寿司の選好のデータセット [Kam03] から、代表的な 8 種の寿司ネタの間の選好 $\{\mu_{i,j}\}$ を作成した。各アルゴリズムの比較回数の一覧をテーブル 1に示す．提案手法は仮定を置かず順序を復元できるアルゴリズムである MallowsMPR と FPRU と比べ比較回数を 5 分の 1 以下に抑えることができた，同様の実験を将棋 AI などのデータセットでも行ったが、そちらでの結果は $K = 10$ 程度で IQS が MallowsMPR,FPRU に比べて 1 桁以上少ない比較回数で終了できることを示すものであった。

5. おわりに

要素間の順序付けを適応的に行う問題を扱った．完全な順序付けを行おうとすると，勝率が拮抗している場合に非常に大きな回数が必要になり，比較回数のほとんどをそれらのアイテムに使ってしまうことになる．実用的にはある程度不完全なソートができていれば良い場合も多い．これらの点を考慮し，所与の δ, τ に対し，確率 $1-\delta$ 以上で，順位相関係数を τ 以上にするアルゴリズムを提案した．完全なソートを行うアルゴリズムと比較し，比較回数を大きく減らすことができることを確かめた．

参考文献

[Arr63] Kenneth Joseph Arrow. *Social choice and individual values.* New York : Wiley, 2nd ed edition, 1963. Previous ed. (B51-9090), Chapman & Hall, 1951.

[BSH14] Róbert Busa-Fekete, Balázs Szörényi, and Eyke Hüllermeier. PAC rank elicitation through adaptive sampling of stochastic pairwise preferences. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada.*, pages 1701–1707, 2014.

表 1: 各アルゴリズムの停止までに要した比較回数．100 回のシミュレーションの平均を取っている．PLPAC-AMPR は順序のみならず，勝率が Prakett-Luce モデルに従うという仮定を置いている．そのため，PLPAC-AMPR は一般の順序のみを仮定した選好に対して順序を復元できるという保証はなく、実際このデータセットでは正しい順序を復元できなかった。それ以外の各アルゴリズムは勝率の順序をケンドールの順位相関係数を 0.8 以上にするという意味で復元できた．

アルゴリズム	比較回数
MallowsMPR	1.2×10^6
PLPAC-AMPR	9.6×10^5
FPRU	8.4×10^6
IQS	2.1×10^5

[BSW+14] Róbert Busa-Fekete, Balázs Szörényi, Paul Weng, Weiwei Cheng, and Eyke Hüllermeier. Preference-based reinforcement learning: evolutionary direct policy search using a preference-based racing algorithm. *Machine Learning*, 97(3):327–351, 2014.

[CJRY12] Olivier Chapelle, Thorsten Joachims, Filip Radlinski, and Yisong Yue. Large-scale validation and analysis of interleaved search evaluation. *ACM Trans. Inf. Syst.*, 30(1):6:1–6:41, March 2012.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, July 2006.

[FRPU94] Uriel Feige, Prabhakar Raghavan, David Peleg, and Eli Upfal. Computing with noisy information. *SIAM J. Comput.*, 23(5):1001–1018, 1994.

[JKPW17] Ramesh Johari, Pete Koomen, Leonid Pekelis, and David Walsh. Peeking at A/B tests: Why it matters, and what to do about it. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13 - 17, 2017*, pages 1517–1525, 2017.

[Kam03] Toshihiro Kamishima. Nantonac collaborative filtering: recommendation based on order responses. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, August 24 - 27, 2003*, pages 583–588, 2003.

[KHKN15] Junpei Komiyama, Junya Honda, Hisashi Kashima, and Hiroshi Nakagawa. Regret lower bound and optimal algorithm in dueling bandit problem. In *COLT*, pages 1141–1154, 2015.

[Mai17] Odalric-Ambrym Maillard. Boundary crossing for general exponential families. In *International Conference on Algorithmic Learning Theory, ALT 2017, 15-17 October 2017, Kyoto University, Kyoto, Japan*, pages 151–184, 2017.

[SBPH15] Balázs Szörényi, Róbert Busa-Fekete, Adil Paul, and Eyke Hüllermeier. Online rank elicitation for plackett-luce: A dueling bandits approach. In *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*, pages 604–612, 2015.

[Sie85] D. Siegmund. *Sequential Analysis: Tests and Confidence Intervals*. Springer Series in Statistics. Springer, 1985.

[YBKJ12] Yisong Yue, Josef Broder, Robert Kleinberg, and Thorsten Joachims. The k-armed dueling bandits problem. *J. Comput. Syst. Sci.*, 78(5):1538–1556, 2012.