

## 科学研究費助成事業（学術研究助成基金助成金）研究成果報告書

平成25年 5月 24日現在

機関番号：17102

研究種目：若手研究（B）

研究期間：2011～2012

課題番号：23700178

研究課題名（和文）ランキング問題のためのブースティング・オンライン予測手法とその応用

研究課題名（英文）Boosting and Online Learning Techniques for Ranking Problems and Their Applications

## 研究代表者

畑埜 晃平（HATANO KOHEI）

九州大学・システム情報科学研究院・助教

研究者番号：60404026

## 研究成果の概要（和文）：

ランキング問題とは、データ間の順位を予測する問題であり、情報検索を始め、推薦タスク、倒産リスク予測、バイオインフォマティクス、自然言語処理など様々な応用分野をもつ。本研究では、オンライン予測理論に基づき、効率的なランキング予測基盤技術を開発した。また、本提案手法はオフライン最適なランキングにほぼ匹敵するランキングのオンライン予測を行う事を示した。さらに、本研究では提案手法を離散構造のオンライン予測手法に拡張した。

## 研究成果の概要（英文）：

Ranking is the problem to predict orderings or permutations over data, and it appears in many application tasks such as information retrieval, recommendation, risk analysis, bioinformatics, natural language processing, and so on. In this project, we developed efficient ranking prediction methods based on the online learning theory. Our methods predict almost as well as the best ranking in hindsight. Further, we generalize our methods to online prediction methods for some classes of combinatorial concepts.

## 交付決定額

（金額単位：円）

	直接経費	間接経費	合計
交付決定額	3,400,000	1,020,000	4,420,000

## 研究分野：機械学習・計算学習理論

科研費の分科・細目：情報学・知能情報学

キーワード：ランキング，オンライン予測，ブースティング，劣モジュラ関数，離散構造機械学習，計算学習理論

## 1. 研究開始当初の背景

ランキング問題に対する従来手法のアプローチは、主に分類問題や回帰問題に帰着させ、機械学習手法を適用する方法が主流であった。例えば、ランキング問題の一種であるAUC最適化問題は正例・負例のペアをどちらが正例か？分類する問題に帰着できる。あとは、SVM やブースティングなど分類手法

を適用すればよい。これらのアプローチはある程度の成功を収めている。しかしながら、従来のアプローチはデータ間の（半）順序関係の情報を活かしてきれていないという問題点を持つ。

そこで近年、データ間のランキング（もしくは順列）そのもの直接モデル化し、予測する試みがなされている。このアプローチの利

点は、自然であることや情報を余すことなく使える点にある。しかし、データ間のランキングの候補（仮説空間）は膨大であり、計算量の問題をどう克服するか？が課題である。また、離散最適化問題を扱うため、従来、機械学習などで用いられる連続凸最適化問題の知見を活かすことが難しい。

## 2. 研究の目的

本研究では申請者らの研究をさらに発展させ、効率的なランキング予測基盤技術を開発する。

## 3. 研究の方法

本研究方法の概要：まず、(1) 順位、上位  $k$  リストをベクトルなどで表現することにより、(2) 問題を離散最適化問題から連続凸最適化問題に帰着させる。その上で連続凸最適化問題を解き、(3) 得られた連続解をランダムサンプリング・ラウンディングなどの手法を用いて再度離散化し、ランキングを得る。その際、収束性やノイズに対する頑健性などを保ったまま、離散化を行う。

## 4. 研究成果

### (1) 順位のオンライン予測

順位のオンライン予測問題はスケジューリングやランキング予測など多くの問題に現れる。本研究では、オンラインスケジューリングに対応する順位の予測問題を考える。本研究では、対応する順位のベクトル表現が置換多面体 (Permutahedron) と呼ばれる特殊な多面体の構造を持つことに着目し、多面体の幾何学的構造を活かした効率的な順位のオンライン予測手法を提案した。また、最

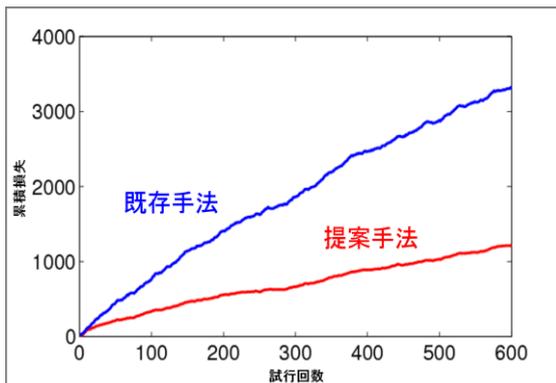
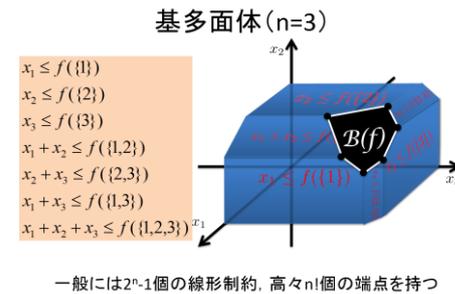


図1：人工データに対する累積損失

適な順位に対する相対的な累積損失（リグレット）の上界を示した。提案手法のリグレットは既存手法よりも定数倍小さい（実験結果、図1）。また、各試行における計算時間は既存手法よりも遥かに効率的である[7]。

### (2) オンラインランク統合問題

オンラインランク統合問題とは、予測した順位と与えられる真の順位間の距離が損失であるような順位のオンライン予測問題である。この問題は、ランク統合問題と呼ばれる複数のランキングを統合する問題の一般化である。本研究ではオンラインランク統合問題に対するオンライン予測手法を開発し、最適な順位に対する相対的な累積損失（リグレット）の上界を示した。また、オンラインランク統合問題に対するリグレットの下界



一般には $2^n-1$ 個の線形制約、高々 $n!$ 個の端点を持つ

図2：劣モジュラ基多面体（黒部分）

を導出することにより、提案手法がほぼ最適である事を示した[1]。

### (3) 劣モジュラ基多面体上のオンライン予測

本研究では、順位、全域木、森といったいくつかの離散構造のクラスが劣モジュラ基多面体と呼ばれる多面体の端点集合として特徴づけられる事に着目した。ここで、劣モジュラ基多面体とは、劣モジュラ関数によって決まる多面体である(図2)。そして、劣モジュラ基多面体で表現できる離散構造のクラスに対する統一的なオンライン予測手法を開発し、最適な離散構造に対するリグレットの上界を示した。さらに、特定のモジュラ関数に対してはより効率的なオンライン予測が可能である事を示した[2]。

### (4) ランキング問題から分類問題への帰着

ランキング問題の1つであるAUC最適化問題に対して、元の問題の最適解の性質を保持したまま、より簡単な分類問題に帰着させる手法を開発した[4]。この成果により、最適解の性質を近似的に保持したまま、問題のサイズを事例の2乗から線形個まで減少させることが可能となり、計算効率の大幅な向上に寄与した。また、共著者の学生である末廣氏は本研究により、IBIS2011若手研究者奨励賞のHonorable Mentionにノミネートされた。

(5) その他

楽曲間の類似性指標に基づく楽曲分類手法の開発, モンテカルロ木探索問題に対するサンプリング手法, 重み付き窓を用いた時系列予測アルゴリズム [5]や, 音楽の自動ジャンル分類手法の開発[6, 3]を行った.

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文, 査読あり] (計 7 件)

- [1] S. Yasutake, K. Hatano, E. Takimoto, and M. Takeda, “Online Rank Aggregation,” in *Proceedings of 4th Asian Conference on Machine Learning (ACML2012)*, 2012, pp. 539-553.
- [2] D. Suehiro, K. Hatano, S. Kijima, E. Takimoto, and K. Nagano, “Online Prediction under Submodular Constraints,” in *Proceedings of 23th Annual Conference on Learning Theory (ALT 2012)*, 2012, pp. 260-274.
- [3] Y. Anan, K. Hatano, H. Bannai, M. Takeda, and K. Satoh, “Polypohic Music Classification on Symbol Data Using Dissimilarity Functions,” in *Proceedings of the 13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, 2012, pp. 229-234.
- [4] D. Suehiro, K. Hatano, and E. Takimoto, “Approximate Reduction from AUC Maximization to 1-norm Soft Margin Optimization,” in *Proceedings of the 22nd*

*International Conference on Algorithmic Learning Theory (ALT 2011)*, 2011, pp. 324-337.

- [5] S. Yoshida, K. Hatano, E. Takimoto, and M. Takeda, “Adaptive Online Prediction using Weighted Windows,” *IEICE Transactions on Information and Systems*, vol. 94-D, no. 10, pp. 1917-1923, 2011.
- [6] Y. Anan, K. Hatano, H. Bannai, and M. Takeda, “Music Genre Classification using Similarity Functions,” in *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, 2011, pp. 693-698.
- [7] S. Yasutake, K. Hatano, S. Kijima, E. Takimoto, and M. Takeda, “Online Linear Optimization over Permutations,” in *Proceedings of the 22nd International Symposium on Algorithms and Computation (ISAAC 2011)*, 2011, pp. 534-543.

[学会発表](計 16 件)

- [8] 藤田隆寛, 畑埜晃平, 瀧本英二, “オフラインアルゴリズムを用いた離散構造のオンライン予測,” in *冬のLAシンポジウム*, 2013.
- [9] 立石大悟, 畑埜晃平, 瀧本英二, “バイアス付き Passive Agressive アルゴリズム,” in *第12回 情報論*

- 的学習理論と機械学習研究会 (IBISML), 2013, pp. 1-8.
- [10] T. Fujita, K. Hatano, and E. Takimoto, “Combinatorial Online Prediction Using Oine Approximation Algorithms,” in *The sixth Annual Meeting of Asian Association for Algorithms and Computation (AAAC 2013)*, 2013.
- [11] 末廣大貴, 畑埜晃平, 来嶋秀治, 瀧本英二, 永野清仁, “基多面体上のオンライン予測,” in *夏のLAシンポジウム*, 2012.
- [12] 阿南陽子, 畑埜晃平, 坂内英夫, 竹田正幸, “類似度に基づくポリフォニックな楽曲の分類,” in *情報処理学会 第9回音楽情報科学研究会 (IPSJ-SIGMUS)*, 2012, pp. 1 - 6.
- [13] D. Suehiro, K. Hatano, and E. Takimoto, “Efficient AUC Maximization by Approximate Reduction of Ranking SVMs,” in *第15回情報論的学習理論ワークショップ (IBIS2012)*, 2012, pp. 243-250.
- [14] D. Suehiro, K. Hatano, and E. Takimoto, “Online Prediction over Base Polyhedra,” in *NIPS 2012 Workshop on Discrete Optimization in Machine Learning (DISCML)*, 2012.
- [15] 松本一成, 畑埜晃平, 瀧本英二, “Bradley-Terry モデルのオンライン予測,” in *夏のLAシンポジウム*, 2012.
- [16] D. Suehiro, K. Hatano, S. Kijima, E. Takimoto, and K. Nagano, “Online Prediction under Submodular Constraints,” in *第9回情報論的学習理論と機械学習研究会 (IBISML)*, 2012, pp. 15-22.
- [17] 寺岡和紀, 畑埜晃平, 瀧本英二, 竹田正幸, “モンテカルロ木探索問題に対する効率的サンプリング手法,” in *冬のLAシンポジウム*, 2012.
- [18] 奥山洋平, 畑埜晃平, 瀧本英二, 竹田正幸, “確率的評価値をもつゲーム木における最善手探索,” in *冬のLAシンポジウム*, 2011.
- [19] 安武翔太, 畑埜晃平, 瀧本英二, 竹田正幸, “トップ k リストのオンライン予測,” in *トップ k リストのオンライン予測*, 2011.
- [20] 金城瞬, 畑埜晃平, and 瀧本英二, “k-選択多腕バンディット問題,” in *夏のLAシンポジウム*, 2011.
- [21] S. Yasutake, K. Hatano, E. Takimoto, and M. Takeda, “Online Rank Aggregation,” in *NIPS 2011 Workshop on Computational Trade-offs in Statistical Learning (COST)*, 2011.
- [22] D. Suehiro, K. Hatano, and E. Takimoto, “Approximate Reduction from AUC Maximization to 1-norm Soft Margin Optimization,” in *NIPS 2011 Workshop on Computational*

*Trade-offs in Statistical Learning*  
(COST), 2011.

- [23] 阿南陽子, 畑埜晃平, 坂内英夫,  
竹田正幸, “類似性指標を用いた楽曲  
のジャンル分類,” in *人工知能学会  
第82回基本問題研究会 (SIG-FPAI)*,  
2011, pp. 25-30.

[その他]

ホームページ:

<http://www.i.kyushu-u.ac.jp/~hatano/>

## 6. 研究組織

### (1) 研究代表者

畑埜 晃平 (HATANO KOHEI)

九州大学大学院・システム情報科学研究院・

助教

研究者番号 : 60404026