

量子アニーリングを用いた非負二値行列因子分解による レコメンデーション手法の検討

香月 諒大^{1,a)} 村上 公平^{1,b)} 矢実 貴志¹ 稲葉 陽子¹ 生田目 崇²

受付日 2020年2月3日, 再受付日 2020年5月8日,
採録日 2020年5月20日

概要: 近年, 組合せ最適化問題を高速に解くためのアーキテクチャとして, 量子アニーリングが注目されている. 機械学習などの分野で量子アニーリングの活用方法について多数の研究がなされ始めているが, その多くがアルゴリズムなどの手法の検討であり, 課題解決的な実問題への検証が少ない. 本論文では, 実際のテレビの閲覧履歴データに対して, 量子アニーリングを利用した非負二値行列因子分解を行い, 視聴者の閲覧傾向の推定を行う手法の検討を行った. この推定手法はレコメンデーションなどに対する応用が考えられ, その手法もあわせて提案を行う. 数値実験を行った結果, 推定を行うための学習データとなる閲覧情報が少ない状況下において, 提案手法による推定性能が, 既存手法による推定性能を上回っていることを確認した.

キーワード: 量子アニーリング, 非負二値行列因子分解, 組合せ最適化, レコメンデーション

A Recommendation Method by Non-negative/Binary Matrix Factorization Using Quantum Annealing

RYOTA KATSUKI^{1,a)} KOHEI MURAKAMI^{1,b)} TAKASHI YAZANE¹ YOKO INABA¹ TAKASHI NAMATAME²

Received: February 3, 2020, Revised: May 8, 2020,
Accepted: May 20, 2020

Abstract: In recent years, quantum annealing has been attracting attention as an architecture for solving combinatorial optimization problems at high speed. A lot of studies start to seek the applications of quantum annealing in fields such as machine learning. However, most of them are limited to research field of theoretical analysis of algorithms or concept proposal, and there are few studies that target problems based on real data. In this research, we propose a method of estimating browsing tendency of viewers on actual TV browsing history data by performing non-negative binary matrix factorization using quantum annealing. We also propose a method to apply this estimation method to recommendations. As a result of the numerical experiment, we confirmed that the estimation performance by the proposed method was higher than that by the exist methods under the condition where there was little browsing information as learning data for estimation.

Keywords: quantum annealing, non-negative/binary matrix factorization, combinatorial optimization problem, recommendation

1. はじめに

近年, 組合せ最適化問題を計算するためのアーキテクチャとして, 量子アニーリングが注目されている. 組合せ最適化問題は, 実応用上の様々な数理モデルや機械学習の領域における基盤となっているが, 問題の構造によっては NP-Complete や NP-Hard と呼ばれる問題が存在する. こ

¹ 株式会社 NTT データ
NTT DATA Corporation, Koto, Tokyo 135-8671, Japan
² 中央大学理工学部
Faculty of Science and Engineering, Chuo University,
Bunkyo, Tokyo 112-8551, Japan
^{a)} Ryota.Katsuki@nttdata.com
^{b)} Kohei.Murakmai@nttdata.com

これらの問題は多項式時間での計算が困難であり、計算量の観点から扱いにくいことが知られている。そのため、これらのクラスに属する大規模な問題に対しては、厳密な最適解の求解を1度諦め、様々なHeuristicsを加えるなどして計算の負荷を軽減しながらも、できるだけ良質な解を導くといったアプローチがしばしばとられている。たとえば焼きなまし法 (Simulated Annealing, SA) [1], [2], [3] や、遺伝的アルゴリズム (Genetic Algorithm, GA) [4], [5], [6] などがその例としてあげられる。

量子アニーリングはそれらの手法と同様に、計算が難しいNP-Hardな組合せ問題に対して、Heuristicに解を導く手法として知られている[7]。現在、量子アニーリングを計算する専用マシンとして、D-Wave Systems Inc. のD-Wave 2000Q (以下、「D-Wave」とする)[8], [9]がある。D-Waveへの問題の実装には特徴があり、具体的には、解きたい対象の組合せ最適化問題と同様の構造を持つ物理系を構成し、その物理系に対するハミルトニアンが物理現象として最小化された結果を取得して、組合せ最適化問題に対する解として利用する構成となっている。

量子アニーリングでは物理実験のハミルトニアンへのマッピングを行うので、求解対象となる組合せ最適化問題の構造にも限りがあり、イジングモデルが入力可能な形式となる。 $J \in \mathbb{R}^{N \times N}$, $\mathbf{h} \in \mathbb{R}^N$ に対して、イジングモデルの最小化問題は次のような式で表される。

$$\underset{\mathbf{x} \in \{-1, 1\}^N}{\text{minimize}} \quad \mathbf{x}^\top J \mathbf{x} + \mathbf{h}^\top \mathbf{x} + \text{const.} \quad (1)$$

イジングモデルの最小化は有名なNP-Hardな問題であるが、他の組合せ最適化問題群を多項式時間での帰着を施すことによってイジングモデルとして解釈するための手法も提案されており[10]、組合せ最適化問題のソルバとしての活用手法も研究されている。

活用先の拡大をテーマとして、機械学習などの基盤となる組合せ最適化問題の高速化を量子アニーリングによって図る研究も提案されている[11]。この先行研究では、D-Waveを用いた非負二値行列因子分解 (Non-negative/Binary Matrix Factorization, NBMF) のアルゴリズムを提案し、画像の基底となる特徴量の抽出タスクに対する性能の検証を行っている。 M 次元で表現されている N 枚の画像データ情報を保持した行列 $V \in \mathbb{R}^{N \times M}$ に対するNBMFは次のように表現される。

$$\underset{W \in \mathbb{R}_+^{N \times K}, H \in \{0, 1\}^{K \times M}}{\text{minimize}} \quad \|V - WH\|_F^2 \quad (2)$$

NBMFを計算する際のアルゴリズムとして、交互最小二乗法 (Alternating Least Squares, ALS) [14]が知られている。ALSは決定変数 W , H を各ステップでいずれかを定数と見なして固定しつつ、交互に最小化を行う手法であるが、 H に対する部分問題は組合せ最適化問題となり、この

求解にD-Waveを用いて計算を行っている。

一般に、 H に対して二値制約を置かない、非負値行列因子分解 (Non-negative Matrix Factorization, NMF) はより広く用いられており、NBMFと同様の画像の基底特徴量抽出[15], [16]だけでなく、次元圧縮やレコメンダ分野でも用いられている。特に、Netflix社が主催した、インターネット動画サイトにおける視聴者の閲覧データを用いた番組レコメンデーションを競うコンペティションでは、NMFを利用した手法が大きな成果を上げている[17], [18]。NBMFはこの派生形であり、分解後の行列 H の要素が0と1しか取りえないため目的関数(2)の最適値は劣化するが、少ない情報量でデータを表現することができるといったメリットがあると考えられる。行列 H に着目すると二値ラベリングをされたマルチラベル分類と見なすこともできる。また応用先として、画像の次元縮約による特徴抽出だけではなく、NMFと同様にレコメンダやクラスタリングでも活用できることが期待できる。

本研究では、量子アニーリングを利用したNBMFにより、テレビ視聴履歴データを対象とした番組レコメンデーションの手法の提案と、その性能の検証を行う。まず、テレビ閲覧履歴に関する情報から、視聴者と番組に関する閲覧傾向を表す情報を抽出し行列形式で表現する。この行列に対してNBMFを適用することで、元の行列が持つ情報から重要なサンプルのみ抽出を行う。真の視聴者の好みを反映した閲覧傾向があると仮定しそれを推定するとき、ノイズを含んだテレビの閲覧データに対して、NBMFで情報の圧縮を行うことにより、視聴者の好みの閲覧傾向を推定できると考える。このような推定手法は、 W , H の双方を非負の連続値行列とするNMFにおいても用いられているが[17]、 H が疎性であることを仮定することで、組合せ最適化問題の求解速度だけではなく、推薦タスクとしての観点からみた推定性能も変更されることが期待される。

数値実験の結果、テレビ閲覧履歴の情報の多くが欠落している状況下において、提案手法は二乗平均平方根誤差 (Root Mean Squared Error, RMSE) の観点からNMFと比較したとき、推薦タスクとしての高い推定性能を持っていることを確認した。また、NBMFの計算速度の観点では、 H の解を求めるために生じる部分問題の求解において、既存手法であるSAと比較したところ、計算速度の向上を確認することができた。

2. 非負二値行列因子分解 (NBMF) について

本章では、先行研究[11]で提案されている、D-WaveによるNBMFのアルゴリズムを紹介する。行列 $V \in \mathbb{R}^{N \times M}$ に対して、

$$\underset{W \in \mathbb{R}_+^{N \times K}, H \in \{0, 1\}^{K \times M}}{\text{minimize}} \quad \|V - WH\|_F^2 \quad (3)$$

を計算し、 W , H を求める。 W , H はいずれも決定変数

であるが、式 (3) に対し ALS で計算を行ううへでは、 W 、 H のうち片方のみを決定変数と見なし、もう片方を定数と見なした部分問題を交互に解く。

まず、 $W \in \mathbb{R}_+^{N \times K}$ を計算する際には NMF における部分問題の既存の更新式 [14] を用いて、

$$W \leftarrow W \circ (VH^\top \oslash WHH^\top) \quad (4)$$

という更新を施す。 \circ 、 \oslash はそれぞれアダマール積、商を表す。これは連続値に対する行列計算として、 $O(NMK)$ で実現可能である。一方、 H に対する更新は、次のような組合せ最適化問題の解で与えられる。

$$\begin{aligned} & \underset{H \in \{0,1\}^{K \times M}}{\text{minimize}} \|V - WH\|_F^2 \\ & = \sum_{i=1}^M (H_i^\top W^\top W H_i - 2V_i^\top W H_i + V_i^\top V_i) \end{aligned} \quad (5)$$

この部分問題は H の各要素が $\{0, 1\}$ であるような 2 次関数である。さらに決定変数の固定を行って、 $i = 1, \dots, M$ を各々別々に解くことにより、 $H_i \in \mathbb{R}^K$ に対する最適化問題とする。この決定変数に対してアフィン変換を施すことによって、 $\{-1, 1\}^K$ が探索を行う対象の空間となり、イジングモデルに帰着することができる。このイジングモデルの最小化問題に対して、D-Wave を用いて計算を行う。

3. 提案手法

本章では多様な消費者行動であるテレビ視聴を対象に、NBMF によるクラスタリングとレコメンデーションを行う手法を提案する。

3.1 テレビ閲覧履歴行列の構成

レコメンデーションの対象とするテレビの視聴者数を N 、番組の数を M とする。各視聴者 i は、自身の選択に基づいて番組 j を閲覧し続ける行動をとる。テレビ閲覧履歴行列 $V \in \mathbb{R}^{N \times M}$ には、その (i, j) 成分に対して、視聴者 i が番組 j を閲覧した時間 (分) を実績として保存する。このように集計を行うことにより、テレビ閲覧履歴行列上には、各視聴者から各番組に対する嗜好度合の情報が保持されていると見なすことができる [12], [13]。

3.2 クラスタリング

NBMF は、視聴者のテレビ閲覧実績や傾向データに基づくが、実際にレコメンデーションを行う場合、視聴者や番組に付随する有益な情報が多々ある。たとえば、視聴者に対して性別や年齢、職業による偏りがあり、テレビ番組に対しては、放映されているテレビ局・時間帯などにより大別することができる。そこで、あらかじめこれらの静的な情報をもとにしてユーザや番組をクラスタリングしておく、それらのクラスタに対してテレビ閲覧履歴行列を構成

する。特に視聴者の番組視聴行動に対する集計結果より、男女の間で閲覧傾向に差があることが分かっており、あらかじめ性別で分けたいうでそれぞれクラスタリングを行う処理を施した。クラスタリングには非階層型クラスタ分析の k -means 法を用いており、具体的な処理と数値は数値実験および付録に記載する。

3.3 レコメンデーション

テレビ閲覧履歴行列は、過去の閲覧情報の実績を元に作成される行列であるが、番組数が非常に多く、視聴者はすべての番組情報を知らないまま番組閲覧している状況などから、テレビ閲覧履歴行列の値と、真の視聴者の好みを反映した閲覧傾向との間に乖離が生じると考えられる。たとえば、番組の広告戦略は、特定の視聴者層にターゲットを絞って配信されるが、視聴者の真の需要以上の広告が配信された場合には閲覧実績が過剰になっていたり、逆に広告が配信されていないために閲覧実績が過少評価されていたりすることが考えられる。その意味でノイズが付与されているテレビ閲覧履歴行列に対して、NBMF によって潜在的な情報を低次元行列へ抽出した後に、真の分布を推定し、閲覧履歴行列の値との差分を用いてレコメンドを行うアプローチをとる。

テレビ閲覧履歴行列 V_{record} に対して、2 章で記述した NBMF を用いて得られた行列を W_{estimate} 、 H_{estimate} とする。推定閲覧傾向行列 V_{estimate} を次のように定める。

$$V_{\text{estimate}} = W_{\text{estimate}} H_{\text{estimate}} \quad (6)$$

そして、差分行列 D を次のように定める。

$$D = V_{\text{estimate}} - V_{\text{record}} \quad (7)$$

行列 D の (i, j) 要素 $D_{i,j}$ について、 $D_{i,j} \geq 0$ の場合には、実際には閲覧傾向が高いにもかかわらず、閲覧実績が少なく見なすことができる。逆に、 $D_{i,j} < 0$ の場合には、過剰評価の状態であると判定を行う。これらの情報を使って、視聴者 i に対して $D_{i,j} \geq 0$ となるような番組群を算出し、大きさの降順に指定の個数をレコメンドする。

4. 数値実験

本章では、式 (3) に対する D-Wave の計算速度の評価、および NBMF のモデル評価として、推薦タスクとしての観点からみた推定性能の測定を行う。

4.1 対象とするデータ

本論文では、経営科学系研究部会連合協議会主催・平成 30 年度データ解析コンペティションで提供された株式会社ビデオリサーチの VR-CUBIC データを用いる [19], [20]。本データは、2016 年から 2017 年の、テレビ視聴データ、Web 接触データ、プロフィールデータとそれらに付随する

マスタデータによって構成される。本研究では、テレビ視聴データおよびプロフィールデータ中の視聴者のデモグラフィック属性を利用して分析を行う。3.2節で記述した方法によって、クラスタリングの前処理を施し、番組クラスタリングを行う際の特徴量には、番組ごとに割り振られているジャンル（分類コード）、放送分数、放送曜日、放送時間帯、視聴者の10歳ごとの年代構成比、視聴者の男女比、視聴者の未婚・既婚率を設定した。視聴者のクラスタリングを行う際の特徴量には、1日の平均閲覧時間、1日を6分割した時間帯ごとの閲覧割合を設定した。その際に必要であれば、各々 Dummy 変数へと変形し、正規化処理を適用した。また提供データでは、視聴者が分単位の各時点にどのチャンネルを視聴していたかが記録されており、本研究では、視聴者がどの番組を視聴したかに焦点を当て、視聴履歴データと番組マスタを突合し、番組の放送時間のうち3分以上の時間を視聴した場合に視聴したものとす。それらを番組に対する情報で統合し、要素に視聴時間を持つテレビ閲覧履歴行列を作成し、数値実験の入力 V_{record} とする。

4.2 数値実験諸元

まず、視聴者と番組に関して、3.2節で記述した方法によって、クラスタリングの前処理を施す。クラスタリングのサイズは、視聴者数が $N = 10$ 、番組の数が $M = 15$ となるように設定した。これは、D-Wave はクラウド上で実装されており求解を行う際の通信時間遅延が生じること、および現行の利用可能なビット数が不足しているといった実験遂行の都合上、大幅なサイズ縮約が必須であることをふまえて設定している。ただし、計算時間の評価においては、部分問題のみを取り出し収束の過程の測定を行う際などには、改めて個別の設定を定めている。また、対象データの入手が困難であることを考慮して、具体的に構成された数値実験対象の V_{record} の値を付録の図 A.1 にまとめた。ALS の反復時間は 10 回と設定し、いずれのパラメータ設定においても十分に収束していることを確認しており、その具体的な収束の挙動は 4.4.1 項で詳述する。

D-Wave のバージョンは DW_2000Q_VFYC_5 を指定しており、2,048 bit の Chimera 結合の構成となっている。Chimera 結合では、最適化問題として一部の決定変数にのみ値を持たせられる仕様となるため、密な問題構造の場合、全結合への変換を行う代償として、64 bit が求解サイズの上限となる。なお、以下の数値実験結果を測定する際には、モデルの計算に関わる本質的な時間のみを対象とするため、通信時間は除外の対象とし、秒数を測定した。投入する問題の係数幅が $[-1, 1]$ 区間となるよう定数倍を行ったうえで、chain_strength=1.2 とした。その他のパラメータ群はデフォルトのまま、annealing_time=10, num_reads=1, No postprocessing

と設定した [21]。また、比較対象とする SA は同 D-Wave Ocean SDK に付属する neal ライブラリを利用しており、パラメータは num_reads=1, num_sweeps=1000, 温度のスケジュールは beta_range=default_beta_range 区間に対する beta_schedule_type=geometric とし、初期値は 0-1 値を等確率でランダムに割り当てている [22]。クラスタリングや部分問題生成などの前処理および SA などの計算は、以下のマシンによって行った。

- OS : Windows Server 2016
- CPU : Intel Core Processor (Broadwell) 2.20 GHz (8 Processors)
- RAM : 16 GB

4.3 NBMF の推定性能の評価

性能評価にあたり、Netflix 社が行った一連のレコメンデーションコンテスト [18] の方法に倣い、 V_{record} に対して、恣意的に σ ($0 < \sigma < 1$) の割合のデータにマスキングを施した行列 V_{masked} を作成する。具体的には、 V_{record} の全要素のうち、ランダムに確率 σ でマスキング対象とする。マスキング対象の行列の要素には、同一行のマスキングされていない対象だけからなる行列要素の値の平均をとり平滑化を施す (σ が大きく、ある行のすべての要素がマスキング対象となった場合には、さらに列方向へ平滑化を施す)。この V_{masked} は行列の一部の値の情報を消す操作であり、閲覧履歴行列の分布からの乖離が生じる。本研究では、 V_{masked} にのみ基づいて $V_{\text{estimated}}$ を学習したうえで、NBMF の分布推定性能指標として、 V_{record} , $V_{\text{estimated}}$ 間の RMSE である

$$\sqrt{\frac{\|V_{\text{record}} - V_{\text{estimated}}\|_F^2}{NM}} \quad (8)$$

を利用し評価を行った。NBMF との比較のために、同様の実験設定にて NMF に対する RMSE も計算している。また結果の平均化のために、NMF, NBMF どちらについてもマスキング対象のランダム性に対して 10 回 RMSE を評価した平均値を計算し記載している。

4.4 数値実験結果

本節では、提案した手法の数値実験の結果と考察を、計算時間と推定性能の観点からまとめる。

4.4.1 計算時間の評価

NBMF における部分問題 (5) の計算手法を SA および D-Wave と変更した場合における、目的関数 (3) の最小化の速度を比較する。 V_{record} に対して、 $\sigma = 1$, $K = 8$ としたときの計算速度は図 1 のようになった。

各反復において、SA, D-Wave とともに類似の目的関数値をたどるが、SA はより短時間で反復が進んでいる。ALS の 1 回の反復あたり、両手法ともに部分問題 (5) が 15 回固定で呼び出されるため、SA と D-Wave の計算時間の差は部

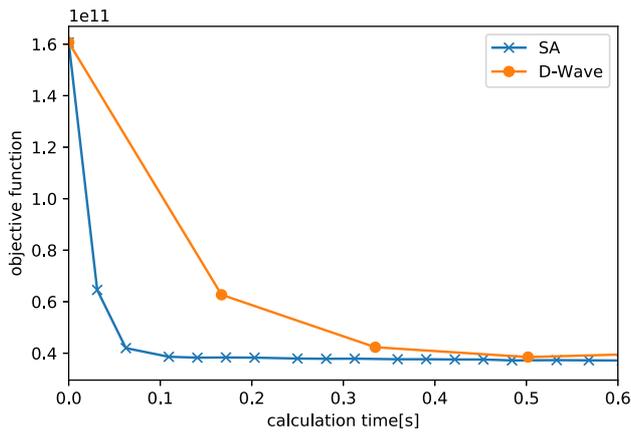


図 1 V_{record} に対し, $K = 8$ における目的関数 (3) の時間推移
Fig. 1 Time transition of objective function (3) at $K = 8$ for V_{record} .

表 1 K に対する ALS の 1 反復目における部分問題 (5) の計算時間および目的関数 (3)

Table 1 Computation time and objective function (3) of sub-problems (5) at the first iteration of ALS per K .

	アルゴリズム	$K = 10$	$K = 20$	$K = 30$	$K = 40$
計算時間 (秒)	SA	0.1005	0.1880	0.7300	0.8000
	D-Wave	0.5000	0.5050	0.5050	0.5100
目的関数 (3)	SA	202.001	171.150	146.102	118.489
	D-Wave	223.813	222.377	256.819	231.729

分問題 (5) の求解時間にそのまま起因する。そこで、表 1 では、部分問題 (5) の部分問題のサイズを変更し、D-Wave と SA の計算時間の特性を比較した。計算時間測定のためのデータとして、 $N = 50, M = 50, [0, 1]$ 区間内で各値をランダムに与えた $V \in \mathbb{R}^{N \times M}$ を準備し、 K を変更しながら評価している。初期値における目的関数 (3) は 11,189.36 であり、表 1 には ALS 1 反復目に得られた目的関数 (3) もあわせて記載する。計算時間は、 $M = 50$ としたために生成される 50 個の部分問題 (5) の平均とした。

K のサイズが大きくなるにつれて、SA の計算時間は徐々に拡大している一方で、D-Wave の計算時間はほぼ一律であり、 $K = 40$ において、D-Wave は短時間で 1 反復を実現している。目的関数 (3) の観点では、SA に比べて劣化しているように見受けられる。

4.4.2 推定性能の評価

本項では、NBMF による欠損値推定の性能を NMF と比較しながら評価する。特に、NBMF の最適化アルゴリズムを SA と D-Wave とで変更した場合における、推定性能の評価もあわせて行う。表 2 は本論文で紹介した各手法に対して、主要なパラメータを変更した際の RMSE を測定したものである。D-Wave と SA の間の RMSE の差はほぼ認められない。これは両手法ともに NBMF の目的関数 (3) としており、また部分問題の呼び出し元となる ALS アルゴリズムが共通であるため、収束先の局所解が類似して

表 2 NMF と NBMF の RMSE の比較

Table 2 RMSE comparison between NMF and NBMF.

σ	モデル	$K = 2$	$K = 4$	$K = 6$	$K = 8$
0.2	NMF	7,780.00	7,898.99	7,996.51	7,782.71
	NBMF-SA	8,105.49	7,773.35	8,161.89	7,764.08
	NBMF-D-Wave	8,105.49	7,773.35	8,161.89	7,764.08
0.4	NMF	10,654.60	11,517.48	11,225.44	11,135.52
	NBMF-SA	10,163.96	10,018.71	10,085.45	10,171.57
	NBMF-D-Wave	10,163.96	10,018.71	10,085.45	10,173.75
0.6	NMF	15,422.50	15,928.64	15,660.74	16,441.19
	NBMF-SA	14,233.60	14,680.08	14,614.45	14,589.41
	NBMF-D-Wave	14,233.60	14,680.08	14,614.45	14,589.22
0.8	NMF	19,525.10	20,130.18	19,774.70	19,623.01
	NBMF-SA	18,315.79	18,655.92	18,768.22	18,656.49
	NBMF-D-Wave	18,315.79	18,655.92	18,768.22	18,656.49

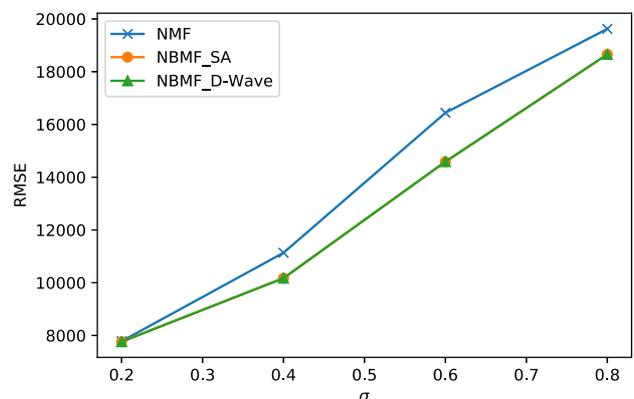


図 2 σ の変更に対する $K = 8$ での NMF と NBMF の RMSE
Fig. 2 RMSE of NMF and NBMF at $K = 8$ for changes of σ .

いるためであると考えられる。 K の大きさに対しても特筆すべき傾向は認められない。一方、元の行列に対してマスキングされる割合が大きい場合、すなわち σ が大きい場合においては、RMSE の観点から NBMF は NMF に比べて推定性能が優れている。図 2 は σ を変更した場合の各手法の RMSE を示しており、この傾向が確認できる。ところで、NBMF は NMF と比較した際に、分解対象となる H に対して 0-1 値の制約を課すため、最適化問題がとりうる決定変数の観点ではとりうる状態が狭まり、モデルとしての自由度が減少していると考えられる。本実験において RMSE 特性が NBMF において改善していることは、一見このモデルの性質に反するよう見受けられる。モデルの複雑性がデータに比べて過度に与えられた場合に、かえって推定性能が劣化する現象は過学習として一般的に知られているが、パラメータのとりうる範囲を正則化項により限定したり、パラメータの更新を限定的にしたりするなどのアプローチがしばしばとられる [23]。NBMF における 0-1 値制約、およびそれが及ぼす数値への改善もこの一種として解釈することができる。

5. 結論

本論文では、実際のテレビ閲覧のデータに対して、D-Wave を利用した NBMF により閲覧傾向の推定を行い、レコメンデーションを行う手法の提案を行った。視聴者や番組をその特性をふまえてクラスタリングを行った後に、視聴者の閲覧履歴となる実績データを集計し、テレビ閲覧履歴行列とした。このような閲覧履歴行列は多数の視聴者の番組の嗜好傾向など有用な情報を保持していると考えられるが、ノイズやテレビ番組の広告不足などによって有意に閲覧量が少なくなっていると考えられ、視聴者の閲覧傾向を NBMF によって推定するフローとなる。また、この NBMF の学習時のアルゴリズムとして組合せ最適化問題を D-Wave により解くような手法が知られているが、本レコメンデーション手法に対する性能、および計算速度の測定を行い、適用可能性の検証を行った。

実データに対してマスキングを行い、その数値を再現するような数値実験の結果、NBMF は NMF と比較して、マスキング前のテレビ閲覧履歴行列の値の推定性能が RMSE の観点で優れていることを示した。また、計算時間の観点では、 K のサイズが大きくなるにつれて、組合せ最適化問題の既存手法である SA は計算時間が拡大するのに対して、D-Wave の場合には固定の短時間での計算が可能である。

本研究の課題点として、D-Wave はハードウェアとして実装されているビット数の不足や、クラウドで公開されている関係上、通信遅延のために検証可能な問題規模に限界があることがあげられる。また、計算時間の観点では D-Wave は最適性が必ずしも保証されないアルゴリズムであるため、問題規模を大きく拡大した際にもスケールするか検討が必要である。得られた行列 H のスパースな構造をうまく活用して、対応する番組をマルチラベル分析するなど、得られた結果の活用方法についても検討の余地があると考えられる。

参考文献

- [1] Kirkpatrick, S., Gelatt Jr., C.D. and Vecchi, M.P.: Optimization by simulated annealing, *Science*, Vol.220, No.4598, pp.671-680 (1983).
- [2] Nahar, S., Sahni, S. and Shragowitz, E.: Simulated annealing and combinatorial optimization, *Proc. 23rd ACM/IEEE Design Automation Conference*, pp.293-299 (1986).
- [3] Aarts, E. and Korst, J.: *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*, John Wiley and Sons, Inc., New York (1989).
- [4] Goldberg, D.E. and Deb, K.: A comparative analysis of selection schemes used in genetic algorithms, *Foundations of Genetic Algorithms*, Vol.1, pp.69-93 (1991).
- [5] Muhlenbein, H.: Parallel genetic algorithms, population genetics and combinatorial optimization, *Workshop on Evolutionary Models and Strategies, Workshop on Par-*

- allel Processing: Logic, Organization, and Technology: Parallelism, Learning, Evolution*, pp.398-406, Springer-Verlag (1991).
- [6] 北野宏明: 遺伝的アルゴリズム, 人工知能学会誌, Vol.7, pp.26-37 (1992).
- [7] Kadowaki, T. and Nishimori, T.: Quantum annealing in the transverse Ising model, *Physical Review E*, Vol.58, No.5, pp.5355-5363 (1998).
- [8] Johnson, M.W., Amin, M.H.S., Gildert, S., Lanting, T., Hamze, F., Dickson, N., et al.: Quantum annealing with manufactured spins, *Nature*, Vol.473, pp.194-198 (2011).
- [9] Gibney, E.: D-Wave upgrade: How scientists are using the world's most controversial quantum computer, *Nature*, Vol.541, pp.447-448 (2017).
- [10] Lucas, A.: Ising formulations of many NP problems, *Frontiers in Physics*, Vol.2 (2014).
- [11] O'Malley, D., Vesselinov, V.V., Alexandrov, B.S. and Alexandrov, L.B.: Nonnegative/Binary matrix factorization with a D-Wave quantum annealer, *PLoS ONE*, Vol.13, pp.1-12 (2018).
- [12] 黒木修隆, 廣瀬裕二, 鈴木達也, 片岡亮照, 沼 昌宏, 山本啓輔: テレビ視聴者の選局行動に基づく番組嗜好度の推定, 映像情報メディア学会誌, Vol.60, No.3, pp.454-457 (2006).
- [13] 立川裕隆, 関亜紀子: 視聴状況に応じた番組推薦のための視聴傾向抽出方法の検討, 情報処理学会マルチメディア通信と分散処理ワークショップ平成 24 年 10 月, pp.96-98 (2006).
- [14] Lin, C.J.: Projected gradient methods for nonnegative matrix factorization, *Neural Computation*, Vol.19, No.10, pp.2756-2779 (2007).
- [15] Eee, D.D.L. and Seung, H.S.: Learning the parts of objects by non-negative matrix factorization, *Nature*, Vol.401, pp.788-791 (1999).
- [16] Eee, D.D.L. and Seung, H.S.: Algorithms for non-negative matrix factorization, *Proc. 13th International Conference on Neural Information Processing Systems (NIPS '00)*, pp.535-541 (2000).
- [17] Koren, Y., Bell, R.M. and Volinsky, C.: Matrix factorization techniques for recommender systems, *Computer*, Vol.42, pp.30-37 (2009).
- [18] Bennett, J. and Lanning, S.: The Netflix prize, *Proc. KDD Cup and Workshop 2007*, pp.3-6 (2007).
- [19] VR CUBIC, available from <https://www.videor.co.jp/service/communication/vrcubic.html> (accessed 2020-04-01).
- [20] 特集データ解析コンペティション: メディア接触データの解析, オペレーションズ・リサーチ: 経営の科学, Vol.65, No.2 (2020).
- [21] D-Wave Solver Properties and Parameters Reference, available from https://docs.dwavesys.com/docs/latest/_downloads/09-1169A-G.D-Wave.Solver-Properties_Parameters_Reference.pdf (accessed 2019-11-13).
- [22] dwave-neal, available from <https://docs.ocean.dwavesys.com/projects/neal/en/latest/> (accessed 2019-11-13).
- [23] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.: Dropout: A simple way to prevent neural networks from overfitting, *Journal of Machine Learning Research*, Vol.15, pp.1929-1958 (2014).

付 録

A.1 付録

図 A-1, 図 A-2, 図 A-3, 図 A-4 に, テレビ閲覧履歴行列 V_{record} と数値実験結果の具体例を示す. 視認性のために, 閲覧履歴行列の値が大きいものに対して濃く, 小さいものに対して薄くなるような値の色付けを施している.

	TV_0	TV_1	TV_2	TV_3	TV_4	TV_5	TV_6	TV_7	TV_8	TV_9	TV_10	TV_11	TV_12	TV_13	TV_14
Audience_0	44520	16380	46080	23160	77880	1620	28620	540	24060	78120	13500	12480	66780	78000	61860
Audience_1	37860	3900	21960	43200	85080	240	12120	2760	16020	42420	4140	840	41460	50100	49380
Audience_2	42360	1920	25500	31920	20700	11940	660	60	52680	37020	10080	4440	26280	10680	16320
Audience_3	55020	600	65820	12660	4560	2640	1620	8100	15360	25920	60	34860	24060	61140	3540
Audience_4	41520	11940	59760	29520	73020	2520	22260	900	9840	21900	2340	2820	40260	81000	22800
Audience_5	37140	8040	43200	40020	37740	2520	2640	9420	70680	16800	32340	4620	4680	17280	12240
Audience_6	25020	5040	58440	51540	21060	420	8880	240	58980	14040	11700	46920	53760	61080	57960
Audience_7	18480	15240	34260	16560	18480	9360	7860	0	22200	51840	180	25440	31560	2940	4380
Audience_8	18240	2520	1560	1560	17580	1080	0	10740	19200	8700	3360	1680	39480	12060	12360
Audience_9	29340	5280	48420	57360	11340	240	8040	4260	37680	38160	32880	5220	29340	62220	16080

図 A-1 テレビ閲覧履歴行列 V_{record}

Fig. A-1 TV browsing history matrix V_{record} .

	TV_0	TV_1	TV_2	TV_3	TV_4	TV_5	TV_6	TV_7	TV_8	TV_9	TV_10	TV_11	TV_12	TV_13	TV_14
Audience_0	44520	16380	44670	23160	21150	2708.57	8030	540	24060	78120	13500	12480	66780	78000	26380
Audience_1	37860	3900	21960	29366.7	21150	240	12120	2760	16020	42420	4140	16957.5	35982.9	46792.5	49380
Audience_2	37044	1920	25500	31920	20700	11940	660	60	26410	37020	10080	4440	26280	10680	26380
Audience_3	55020	600	65820	12660	21150	2708.57	1620	8100	15360	25920	9705	34860	35982.9	61140	3540
Audience_4	37044	11940	59760	29520	21150	2820	22260	900	9840	21900	2340	16957.5	35982.9	81000	22800
Audience_5	37044	8040	43200	40020	37740	2520	2640	9420	70980	16800	32340	4620	4680	17280	12240
Audience_6	37044	5040	58440	51540	21060	420	8880	240	26410	14040	11700	46920	53760	61080	57960
Audience_7	18480	15240	34260	16560	18480	2708.57	8030	495	22200	30615	180	25440	31560	2940	26380
Audience_8	37044	6637.5	44670	1560	17580	1080	8030	10740	26410	8700	3360	1680	39480	46792.5	12360
Audience_9	29340	5280	48420	57360	11340	240	8030	4095	26410	30615	9705	5220	29340	62220	26380

図 A-2 $\sigma = 0.2$ としたときの V_{masked}

Fig. A-2 V_{masked} when $\sigma = 0.2$.

	TV_0	TV_1	TV_2	TV_3	TV_4	TV_5	TV_6	TV_7	TV_8	TV_9	TV_10	TV_11	TV_12	TV_13	TV_14
Audience_0	42547.8	14149.9	49089.8	25919.8	21848.8	3440.81	7792.42	2680.3	20179	70670.8	12296.6	17073	64952.8	82810.3	29113.5
Audience_1	34253.9	5198.03	34391.6	29965.5	17474.8	3212.29	8266.55	1875.32	13759.3	39264.4	4475.45	19651.6	35096.6	51021.7	41182.1
Audience_2	31300.5	3621.78	29835.1	34129.6	19767.6	3549.5	3052.84	2242.87	23502.6	31202.4	10414.4	5799.58	19913.6	30903.6	19061
Audience_3	53751	6745.14	60968.2	19505.7	22054	3095.77	5537.67	4695.2	22109.1	30718.4	6961.12	32344.1	36763.4	58465.6	11609.9
Audience_4	43636.2	7680.89	57153.7	28319.4	18910.1	2790.22	12077.5	6088.05	24745.3	33157.5	4783.32	20885.1	40516.1	63831.4	27202
Audience_5	30679.5	6559.14	50853.8	48559.4	31037.7	2752.38	6968.99	6053.36	54651.4	15309.4	25788	11369.9	20481.8	24445.7	15536
Audience_6	42242.9	6893.84	54814.7	39675.8	27285.8	2140.39	12227.1	3982.83	34408.1	21151.2	15641.4	43888.6	51364.5	57641.6	60730.5
Audience_7	20546.4	3964.07	30860.5	22819.4	13773	1572.37	5930.67	2678.69	21486.2	16114.4	10127.7	18727.6	20059.6	25315.2	23938.4
Audience_8	31514.7	6246.58	34918	19617.6	12970.4	1289.34	9920.07	5751.44	23016.7	16406.9	5221.39	7481.65	39073.6	48160.8	16332.9
Audience_9	36032.2	7583.41	47937.2	41177.6	21638.2	3490.27	11959.8	6217.74	35069.5	38225.4	11638	6584.8	32824.7	50408	25772.7

図 A-3 $K = 8$ において NMF により計算した V_{estimate}

Fig. A-3 V_{estimate} calculated by NMF with $K = 8$.

	TV_0	TV_1	TV_2	TV_3	TV_4	TV_5	TV_6	TV_7	TV_8	TV_9	TV_10	TV_11	TV_12	TV_13	TV_14
Audience_0	55823.2	31226.8	55823.2	55823.2	51997.1	2850.41	30980.6	3426.51	50576.1	46207.2	27396.2	19719	58823.2	47725.7	50576.1
Audience_1	42444.2	20426	42444.2	42444.2	42195.9	8096.48	21511.5	6489.36	36274.5	37942.4	19423.1	19335.9	42444.2	28178.1	36274.5
Audience_2	29409.1	10755.4	29409.1	29409.1	26958.4	7601.73	10310.5	6383.71	28518.4	27087.6	9306.71	19113.3	29409.1	20916.7	28518.4
Audience_3	30552.7	15465.5	30552.7	30552.7	28688.8	5331.38	10701.7	2476.55	30187.4	20022.9	9884.34	20041.6	30552.7	24856	30187.4
Audience_4	41729.1	21141.8	41729.1	41729.1	41812.7	4793.39	20475.8	10419.1	38251.5	32983.5	15267.8	24074.6	41729.1	33458.1	38251.5
Audience_5	32963.7	18627.5	32963.7	32963.7	24987.4	3492.15	11506	3774.32	28729.4	27865.3	21477.8	20031.1	32963.7	25237.3	28729.4
Audience_6	47089.8	30455.4	47089.8	47089.8	33513.5	6812.98	5131.89	6468.62	39104.3	39505.5	25384.4	34442.2	47089.8	32291.3	39104.3
Audience_7	24851.7	16091.9	24851.7	24851.7	18973.1	4131.04	6609.24	2983.04	21625.5	20506.2	12522.2	17338.1	24851.7	17494.4	21625.5
Audience_8	14033.8	4050.11	14033.8	14033.8	11478.5	2806.13	4982.29	5699.92	13306.3	13839.2	7692.17	10955.9	14033.8	10800.2	13306.3
Audience_9	36623.3	14803.2	36623.3	36623.3	29244.1	1926.22	16805.3	12390.2	35702.1	35665.1	19531.1	22623.7	36623.3	33775.8	35702.1

図 A-4 $K = 8$ において NBMF により計算した V_{estimate}

Fig. A-4 V_{estimate} calculated by NBMF with $K = 8$.



香月 諒大

2018年京都大学大学院情報学研究科数理工学専攻修士課程修了. 同年株式会社エヌ・ティ・ティ・データ. 量子コンピュータ/量子アニーリング等マシンを用いた計算の研究開発に従事.



村上 公平

2018年京都大学大学院情報学研究科先端数理科学専攻修士課程修了. 同年株式会社エヌ・ティ・ティ・データ. 医療向け自然言語処理の研究開発に従事.



矢実 貴志

2011年京都大学大学院情報学研究科システム科学専攻修士課程修了. 同年株式会社エヌ・ティ・ティ・データ. 量子コンピュータ/量子アニーリング等マシンを用いた計算の研究開発に従事.



稲葉 陽子

2001年東京大学理学系研究科地球惑星科学専攻修士課程修了. 同年株式会社エヌ・ティ・ティ・データ. データアナリティクス・AI等の研究開発, および, 顧客向けのコンサルティングに従事.



生田目 崇

1999年東京理科大学大学院工学研究科経営工学専攻博士後期課程修了, 博士(工学). 2013年より中央大学理工学部経営システム工学科教授. マーケティング・サイエンス, 経営科学の研究に従事.