

評点付きレビュー文書を対象とした トピックモデルの構築に関する検討

田村 一樹^{1,a)} 吉川 大弘¹ 古橋 武¹

受付日 2014年6月5日, 採録日 2014年10月8日

概要: 多くの企業にとって, 商品 (アイテム) に対するユーザレビューの解析は, 商品開発やマーケティングの面で重要な役割を占めている. ユーザが自由に感想を記述することのできるレビューには, 商品の長所や魅力だけでなく, 不満や改善点などの声が豊富に含まれている. それらは商品開発における重要な情報源であり, 大量のレビューに対し, 自動で解析を行うことができれば, 従来見逃していた有益な知見を得られるようになると期待できる. 推薦システムにおいては, レビューに含まれる評点情報を利用して, ユーザやアイテムの関係性をとらえる協調フィルタリングが成果をあげている. 一方, 一般的なテキストの解析においては, 潜在的なトピックを推定する手法である, トピックモデルが注目されている. 本論文では, これらを発展させ, 評点付きのレビューテキストに特化したトピックモデルを構築する. そこでまず, レビューテキスト中の単語の出現に対して, ユーザ・アイテムのトピックを仮定したトピックモデル SBi-LDA (Switching Bi-LDA) を提案する. 続いて, ユーザ・アイテムのトピックに加え, 評点情報も組み込んだ新たなトピックモデル MBi-LDA (Multi Bi-LDA) を構築し, 提案する. 実際のレビューテキストを用いた実験を行い, 提案するモデルが従来のモデルに比べてパープレキシティの面で優れていることを示す.

キーワード: レビュー分析, トピックモデル, 潜在的ディリクレ配分法, パープレキシティ, 情報推薦

A Study on Construction of Topic Model Intended for Review Texts with Rating

KAZUKI TAMURA^{1,a)} TOMOHIRO YOSHIKAWA¹ TAKESHI FURUHASHI¹

Received: June 5, 2014, Accepted: October 8, 2014

Abstract: In a lot of companies, analysis of users' reviews on items plays an important roll in the aspect of product development and marketing. Reviews are described by users freely so that they contain many opinions which are not only advantages of items but also disadvantages and improvements. They are important source of information. If huge number of reviews can be analyzed automatically, it is expected to obtain useful knowledge which is missed conventionally. In recommender systems, the collaborative filtering which can capture relationships among users and items using ratings of reviews achieves good results. In text analyses, topic models which can infer latent topics in texts are focused. This study develops these methods and constructs new topic models which are specific to review texts with rating. First, a topic model called SBi-LDA (Switching Bi-LDA) which assumes user-topics and item-topics on appearance of words is proposed. Then, a new topic model called MBi-LDA (Multi Bi-LDA) which incorporates ratings of reviews in addition to above topics is constructed and proposed. Experiments using real-life reviews are carried out and it is shown that the proposed models are superior to conventional models in terms of perplexity.

Keywords: review analysis, topic models, latent Dirichlet allocation, perplexity, information recommendation

1. 背景と目的

多くの企業にとって、商品（アイテム）に対するユーザーレビューの解析は、商品開発やマーケティングの面で重要な役割を占めている。また近年は、Amazon.com^{*1}や楽天市場^{*2}などのオンラインショッピングサイトの普及により、容易に大量のレビューを収集することが可能となっている。ユーザーが自由に感想を記述することのできるレビューには、商品の長所や魅力だけでなく、不満や改善点などの声が豊富に含まれる。それらは商品開発における重要な情報源であるものの、大量のレビューをすべて読むことは時間や労力の面で困難であり、自動でレビューの解析を行うことができれば、従来見逃していた有益な知見を得られるようになる期待できる [1], [2].

また、大量のレビューの中から、ユーザーやアイテムに関する情報を抽出することは、マーケティングの面でも有用である。特にユーザー間・アイテム間の嗜好の類似性を自動的にとらえる協調フィルタリングは、推薦システムを実現するための主要な方法となっている [3], [4]. これまでの協調フィルタリングに基づく手法では、主に評点や閲覧履歴の情報のみからユーザーの嗜好を推定していた。しかし近年、レビュー中のテキストを解析することにより、嗜好の推定性能を向上させられることが報告されている [5], [6]. 本論文では、このユーザー間・アイテム間の類似性に注目した、レビューテキストの解析手法についての検討を行う。

テキスト解析の手法として、確率的潜在意味解析 (PLSA: Probabilistic Latent Semantic Analysis) [7] や潜在的ディリクレ配分法 (LDA: Latent Dirichlet Allocation) [8] などのトピックモデルが注目を集めている。トピックモデルは、文書に出現する単語とその出現回数の情報から、それぞれの文書に潜在的に存在するトピックを、精度良く推定することができる手法として知られている。しかし、これらの手法は一般のテキストを想定しており、レビューテキスト特有のユーザー、アイテム、評点の情報を効果的に利用できていない。PLSA や LDA を発展させ、レビューテキストの解析に特化したトピックモデルも提案されている [9], [10] もの、ユーザー、アイテム、評点の情報に注目し、それらを同時にモデル化した研究はまだ報告されていない。

一方で、評点予測の手法では、LDA を拡張し、レビューにおけるユーザー間・アイテム間の類似性から評点を予測する、Bi-LDA [11] が提案されている。Bi-LDA は、ユーザー・アイテムそれぞれについてのトピックを仮定した階層ベイズモデルであり、各トピックの組合せを考慮することにより、精度良く評点を推定することができる。しかし、この

手法は評点の予測を目的としており、レビューテキストの解析に用いることは想定していない。

そこで本論文では、Bi-LDA を拡張し、ユーザー・アイテムのトピックを仮定した、レビューテキストのトピックモデル SBi-LDA (Switching Bi-LDA) を提案する。SBi-LDA では、各レビューはユーザートピック・アイテムトピックを保持し、レビューテキスト中の各単語は、いずれかのトピックから出現すると仮定する。各単語がユーザートピック/アイテムトピックどちらのトピックから出現するかは確率的に決定され、モデルの学習時にギブスサンプリングを用いてその確率が推定される。

また本論文では、SBi-LDA をさらに発展させ、評点の情報を組み込んだトピックモデル MBi-LDA (Multi Bi-LDA) を構築し、提案する。MBi-LDA は、評点生成部分と単語生成部分に分けられる。評点生成部分では、Bi-LDA と同様の過程で評点を予測する。一方、単語生成部分では、レビューテキスト中の各単語が、ユーザートピック、アイテムトピック、評点のいずれかから出現すると仮定する。各ユーザーが保持するユーザートピック、各アイテムが保持するアイテムトピックの情報に加え、評点生成部分で予測した評点の情報と、各トピック・評点から出現しやすい単語の情報を組み合わせることにより、未記述のレビューテキスト中に出現する単語の予測を可能とする。

実験では、提案モデルを実際のレビューデータに適用し、モデルの性能を定量的/定性的に評価する。定量評価実験では、言語モデルの性能評価に広く用いられる指標であるパープレキシティ [8] に基づき、従来モデルとの比較を行う。初めに、従来モデルである LDA, Bi-LDA と、提案モデルである SBi-LDA の性能比較実験を行い、SBi-LDA が優れていることを示す。さらに、SBi-LDA と MBi-LDA の2つの提案モデル間の比較も行い、MBi-LDA の性能がより優れていることを示す。次に定性評価実験では、MBi-LDA がユーザートピック・アイテムトピックを仮定どおり推定できていることを目視で確認する。また、評点から出現する単語について、評点の大小による評価語の出現傾向について検証する。

2. 従来研究

トピックモデルの代表的な手法には、確率的潜在意味解析 (PLSA) [7] や潜在的ディリクレ配分法 (LDA) [8] があげられる。特に LDA は、一般的なテキストのトピック解析において、高性能にトピックを推定できる手法として知られている。しかし、基礎的なモデルであるため、対象とするテキストによっては、モデルがうまく働かない場合や、仮定が不十分な場合が存在する [9]. LDA のグラフィカルモデルを図 1 に示す。なお、図中のパラメータについては付録 A.1 を参照されたい。

また、レビューテキストから得られる知見の応用を

¹ 名古屋大学
Nagoya University, Nagoya, Aichi 464-8603, Japan

^{a)} tamura@cmplx.cse.nagoya-u.ac.jp

^{*1} <http://www.amazon.com/>

^{*2} <http://www.rakuten.co.jp/>

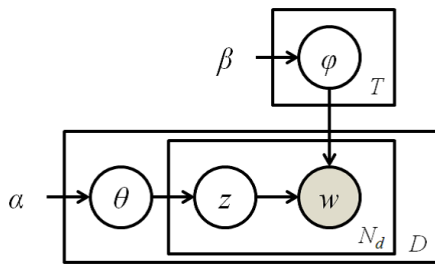


図 1 LDA のグラフィカルモデル
Fig. 1 Graphical model of LDA.

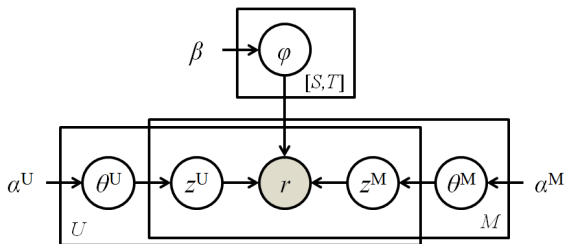


図 2 Bi-LDA のグラフィカルモデル
Fig. 2 Graphical model of Bi-LDA.

目的とする研究は、これまでに数多く報告されている [12], [13], [14]. 近年は、トピックモデルを、レビューテキストの解析を目的として発展させた研究も報告されており、成果をあげている [9], [10]. 文献 [9] の手法は、レビューテキスト中のトピックの局所性を仮定することで、トピックの推定性能向上を試みている。しかし、ユーザ間・アイテム間の関係性には着目しておらず、評点の情報も用いていないため、本論文の想定とは異なる。また文献 [10] のモデルでは、レビューテキストの背景にある感情を仮定し、同時にモデル化することでレビューの感情推定に発展させているものの、文献 [9] と同様に、ユーザ間・アイテム間の関係性を表現していない。

一方、ユーザ間・アイテム間の関係性を利用して、評点の予測を行う研究には、Mnih ら [15] と Salakhutdinov ら [16] の研究がある。これらの研究では、評点行列をユーザ・アイテムの因子に確率的に分解する、確率的行列因子分解 (PMF: Probabilistic Matrix Factorization) を提案している。PMF では、評点を実数として学習・予測をしているため、テキスト中の単語の出現を離散数で扱う言語モデルとの親和性は低い。それに対し、評点を離散数として学習・予測をするモデルに、Bi-LDA がある [11]. Bi-LDA は、LDA を発展させた階層ベイズモデルで、評点の観測に多項分布を、その事前分布にはディリクレ分布を仮定している。Bi-LDA のグラフィカルモデルを図 2 に示す。なお、図中のパラメータについては付録 A.2 を参照されたい。

3. 提案モデル

本論文では、レビューテキストに出現する単語の背景に、ユーザトピック、アイテムトピック、評点の要因が存在す

ることを仮定し、それらをモデル化することを目指す。そこでまず、ユーザトピック・アイテムトピックを仮定した評点傾向のモデル化手法である Bi-LDA を、レビューテキストの解析手法に発展させる。Bi-LDA は、ユーザトピックとアイテムトピックの組合せから、評点を生成するモデルである。このモデルをそのままテキスト解析に適用することも可能ではあるが、テキスト中の単語は、ユーザトピック、アイテムトピックのいずれかから出現すると考えられるため、ユーザトピック・アイテムトピックの組合せから単語が出現するという仮定は適切ではないと考えられる。そこで初めに、Bi-LDA の上記問題点に対処したモデルである、Switching Bi-LDA (SBI-LDA) を構築する。次に、SBI-LDA をさらに発展させ、単語の出現の因子に評点も仮定したモデルである、Multi Bi-LDA (MBi-LDA) を構築し、提案する。

3.1 Switching Bi-LDA

本節では、テキスト解析を目的として、Bi-LDA を改良した提案モデルである、SBI-LDA について述べる。レビューテキスト中の単語に着目すると、ユーザに特徴づけられる単語とアイテムに特徴づけられる単語に分けられることが観察できる。たとえば、価格にこだわりを持つユーザは、「価格」「値段」「料金」といった単語をどのレビューにおいても多く用いる傾向がある。また、海産物に関連するアイテムでは、「魚」「新鮮」「冷凍」といった単語がどのレビューテキストにおいても多く出現する。これらの傾向から、SBI-LDA では、各レビューにはユーザトピックとアイテムトピックが存在し、レビューテキストは、それぞれのトピックから出現する単語の組合せで構成されていると仮定する。どちらのトピックから単語が出現するかは、スイッチ変数によって制御される。SBI-LDA では、スイッチ変数は二値 0, 1 を持ち、それぞれユーザトピック、アイテムトピックに対応する。ここで、0, 1 の出現分布は各ユーザに固有であると仮定する。単語の生成過程では、まずスイッチ変数が確率的に決定される。このスイッチ変数が 0 である場合はユーザトピックから、1 である場合はアイテムトピックから、それぞれ単語が確率的に決定される。これらを単語数分繰り返すことで、レビューテキスト中の単語が、ユーザトピック、アイテムトピックからそれぞれ出現する単語の組合せで構成されるよう制御できる。トピック、単語の出現には多項分布を仮定し、それらの事前分布としてディリクレ分布を導入する。また、スイッチ変数には二項分布を仮定し、事前分布にはベータ分布を導入する。SBI-LDA のグラフィカルモデルを図 3 に示す。なお、図中のパラメータについては付録 A.3 を参照されたい。

3.2 Multi Bi-LDA

本節では、SBI-LDA をさらに発展させ、レビューテキ

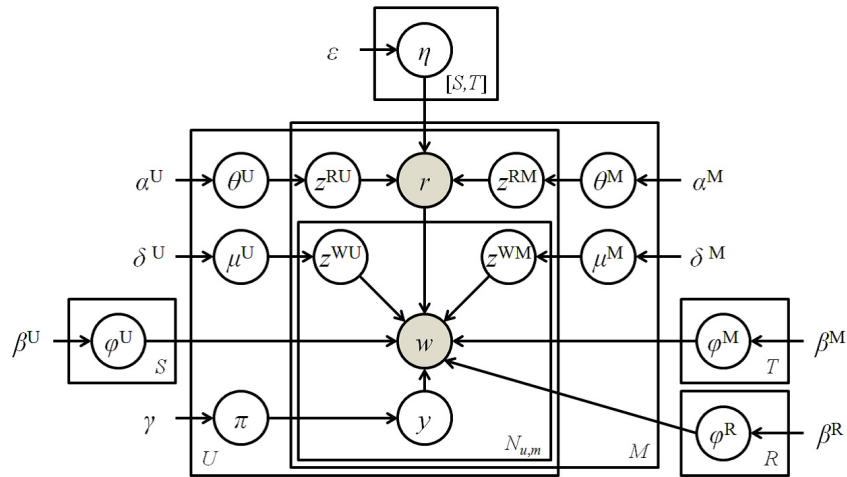


図 4 MBI-LDA のグラフィカルモデル

Fig. 4 Graphical model of MBI-LDA.

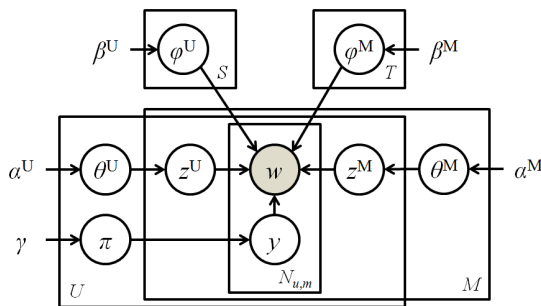


図 3 SBI-LDA のグラフィカルモデル

Fig. 3 Graphical model of SBI-LDA.

スト中の単語が、ユーザトピック、アイテムトピック、評点のいずれかから生成されると仮定したモデルである、MBi-LDA を提案する。レビューテキスト中の単語に着目すると、SBI-LDA で仮定したユーザに特徴づけられる単語とアイテムに特徴づけられる単語のほかに、レビューに付与される評点に関連の強い単語がレビュー中に多く存在することが確認できる。たとえば、高評価のレビューテキスト中には、「最高」「満足」「素晴らしい」といった肯定的な単語が多く用いられる一方で、低評価のレビューテキスト中には、「最低」「不満」「がっかり」といった否定的な単語が多く用いられる。これらの傾向から、評点情報を単語生成の潜在情報に組み込むことで、さらに高性能なレビューテキストのモデル化が実現できると考えられる。

そこで MBI-LDA では、レビューに付与される評点と、レビューテキスト中の単語の出現を同時にモデル化する。レビューの評点は Bi-LDA と同様に、評点ユーザトピックと評点アイテムトピックの組合せより、確率的に生成される。テキスト中の単語は、各レビューに存在するユーザトピック、アイテムトピック、評点のいずれかから生成され、それぞれから出現した単語の組合せでレビューテキストが構成されると仮定する。

いずれのトピック・評点から単語が出現するかは、ス

イッチ変数によって制御される。MBi-LDA では、スイッチ変数は三値 0, 1, 2 を持ち、それぞれユーザトピック、アイテムトピック、評点に対応する。ここで、0, 1, 2 の出現分布は各ユーザに固有であると仮定する。単語の生成過程では、まずスイッチ変数が確率的に決定される。0 である場合はユーザトピックから、1 である場合はアイテムトピックから、2 である場合は評点から、それぞれ単語が確率的に決定される。これらを単語数分繰り返すことで、レビューテキスト中の単語が、ユーザトピック、アイテムトピック、評点からそれぞれ出現する単語の組合せで構成されるよう制御できる。

提案モデルにおいて、ユーザトピックとアイテムトピックは、評点の生成と単語の生成に対して、それぞれ別のトピックを導入する。本論文では、評点生成に寄与するトピックを、評点ユーザトピック、評点アイテムトピックと表し、単語生成に寄与するトピックを、単語ユーザトピック、単語アイテムトピックと表す。トピック、評点、単語、スイッチ変数の出現には多項分布を仮定し、それらの事前分布としてディリクレ分布を導入する。MBi-LDA のグラフィカルモデルを図 4 に示す。なお、図中のパラメータと本モデルの詳細な導出過程については、付録 A.4 ~ A.6 を参照されたい。

4. 実験

提案したモデルの性能を評価するため、性能評価実験を行った。実験では、実際のレビューデータから構築した実験用データセットを用い、性能の定量評価と定性評価を行った。また、モデルの学習時間についての比較も行った。比較するモデルは、LDA, Bi-LDA, SBI-LDA, MBI-LDA とした。

定量評価実験では、データセットを訓練データとテストデータに分け、訓練データでモデルの学習を、テストデータで性能評価を行った。言語の予測性能を表す指標である

表 1 各データセットの統計情報

Table 1 Statistical information of each dataset.

| (a) データセット「食品」 | |
|----------------|--------|
| レビュー件数 | 34,305 |
| ユーザ数 | 1,918 |
| アイテム数 | 612 |
| 語彙数 | 6,183 |
| 評点の密度 | 2.92% |

| (b) データセット「CD・DVD・楽器」 | |
|-----------------------|--------|
| レビュー件数 | 11,580 |
| ユーザ数 | 1,653 |
| アイテム数 | 322 |
| 語彙数 | 4,109 |
| 評点の密度 | 2.18% |

パープレキシティ [8] を評価指標に用い、モデル間での性能比較を行った。各モデルでトピック数を変化させて実験を行い、各モデルでの最良値を用いて多重性を考慮した検定を行った。

定性評価実験では、各トピックから高確率で出現する単語を目視で確認し、提案モデルで仮定した、ユーザトピック・アイテムトピックを推定できていることを確認した。また、MBi-LDA において、評点から出現する単語について、既存の評価極性辞書を用い、評点の大小による評価語の出現傾向について検証した。

4.1 適用データ

提案モデルは、ユーザ、アイテム、評点の情報から、レビューテキストのモデル化を行う。本論文では、楽天技術研究所が公開^{*3}する、「みんなのレビュー・口コミ情報」のデータを用い、実験用データセットを構築した。公開されているデータベースには、楽天市場のアイテムに対する 1~5 の評点やレビューテキスト、投稿者の属性情報などが含まれている。データセットは、カテゴリ「食品」、「CD・DVD・楽器」のレビューを抜粋して構築した。その際、含まれるデータ数の偏りによる影響をなくすため、評点 1, 2 を評点 3 に統合したうえで、各評点が同数ずつになるよう、レビューを抽出した。また、データサイズに対してアイテムの種類数が膨大であるため、付与されている詳細カテゴリでアイテムを統合した。レビューテキスト中の単語としては、MeCab^{*4}による形態素解析で得られた「名詞」「動詞」「形容詞」を用いた。なお、全レビューでの合計出現回数が 5 回以下の単語についてはあらかじめ除去した。各データセットの統計的な情報を表 1 に示す。

^{*3} <http://rit.rakuten.co.jp/rdr/>

^{*4} <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

4.2 定量評価実験

4.2.1 実験方法

定量評価実験では、従来モデルである LDA, Bi-LDA と、本論文で提案する SBi-LDA, MBi-LDA についての性能評価を行った。評価指標には、言語モデルの一般的な定量指標である、パープレキシティを用いた。パープレキシティは、式 (1) で求められる。

$$Perplexity = \exp \left(-\frac{1}{T} \sum_{t=1}^T \log p(w_t) \right) \quad (1)$$

ただし、 T はテストデータ中の単語の総数であり、 $p(w_t)$ は単語分布の MAP (Maximum a posteriori: 最大事後確率) 推定値から計算される。パープレキシティは、テストデータに含まれる単語の平均出現確率の逆数を表しており、値が小さいほど性能が高いことを表す指標である。

実験では、事前に 100 件のテストデータをランダムに抽出し、残りを訓練データとしてモデルの学習を行い、パープレキシティを算出した。その操作を 1 試行とし、30 試行の平均を用いて比較を行った。なお、各モデルにおける変数について、すべてのハイパーパラメータを 0.1 に固定し、ギブスサンプリングのイテレーション回数は 100 回とした。トピック数を 10 刻みで 10~70 まで変化させ、各モデルで最良となる値を用いて比較を行った。ただし、複数の潜在トピックを仮定するモデルにおけるトピック数は、組合せ数の都合上、ユーザトピック・アイテムトピックを同数として変化させた。このとき、MBi-LDA の評点ユーザトピック数、評点アイテムトピック数は、事前実験で評点の予測性能が最良となった値を設定した。また、統計的に手法間の差異を示すため、各モデルでの最良値に対し、検定を行った。

4.2.2 結果と考察

各データセットにおいて、LDA, Bi-LDA, SBi-LDA, MBi-LDA の各モデルでトピック数を変化させてパープレキシティを算出した。なお、MBi-LDA の評点生成部分のトピック数は、事前実験の結果より、データセット「食品」で $(K, L) = (10, 10)$ 、データセット「CD・DVD・楽器」で $(K, L) = (5, 5)$ とし、このとき評点の予測精度はそれぞれ、89.1%, 72.3% だった。実験結果を表 2, 表 3 に示す。なお、 K, L, S, T は付録 A.4 と同様に、各トピックの総数を表し、表中に斜体で示した値は、各モデルでの最良値であることを表している。

表 2, 表 3 にあるとおり、どちらのデータセットにおいても、提案モデルである SBi-LDA と MBi-LDA が、性能の面で他の従来モデルを上回る結果となった。また、2 つの提案モデル間において、MBi-LDA が SBi-LDA を上回る結果を得ることができた。

ここで、提案モデルの有用性を統計的に示すための検定を行った。各データセットにおける各モデルの最良値の比

表 2 各モデルのパープレキシティの比較 (データセット「食品」)
 Table 2 Comparison of perplexities of each model (dataset “Food”).

| (a) LDA | | | (b) Bi-LDA | | | |
|---------|----------|-------|------------|-----|----------|-------|
| T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 |
| 10 | 897.6 | 97.7 | 10 | 10 | 740.0 | 58.0 |
| 20 | 866.6 | 103.8 | 20 | 20 | 873.7 | 80.4 |
| 30 | 826.9 | 102.9 | 30 | 30 | 997.4 | 70.9 |
| 40 | 887.9 | 117.3 | 40 | 40 | 1080.8 | 108.9 |
| 50 | 879.2 | 93.8 | 50 | 50 | 1134.8 | 73.6 |
| 60 | 887.2 | 85.8 | 60 | 60 | 1211.7 | 80.0 |
| 70 | 857.6 | 75.5 | 70 | 70 | 1237.3 | 125.6 |

| (c) SBi-LDA | | | | (d) MBi-LDA | | | | | |
|-------------|-----|----------|------|-------------|-----|-----|-----|----------|------|
| S | T | パープレキシティ | 標準偏差 | K | L | S | T | パープレキシティ | 標準偏差 |
| 10 | 10 | 650.4 | 46.8 | 10 | 10 | 10 | 10 | 635.9 | 52.1 |
| 20 | 20 | 632.5 | 61.8 | 10 | 10 | 20 | 20 | 617.9 | 42.6 |
| 30 | 30 | 632.3 | 39.5 | 10 | 10 | 30 | 30 | 607.1 | 47.3 |
| 40 | 40 | 637.0 | 49.1 | 10 | 10 | 40 | 40 | 594.7 | 47.4 |
| 50 | 50 | 653.7 | 47.1 | 10 | 10 | 50 | 50 | 624.1 | 59.9 |
| 60 | 60 | 642.1 | 50.2 | 10 | 10 | 60 | 60 | 618.5 | 56.8 |
| 70 | 70 | 665.2 | 49.0 | 10 | 10 | 70 | 70 | 625.5 | 54.2 |

表 3 各モデルのパープレキシティの比較 (データセット「CD・DVD・楽器」)
 Table 3 Comparison of perplexities of each model (dataset “CD DVD Instruments”).

| (a) LDA | | | (b) Bi-LDA | | | |
|---------|----------|-------|------------|-----|----------|-------|
| T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 |
| 10 | 891.8 | 84.2 | 10 | 10 | 791.7 | 59.8 |
| 20 | 937.9 | 82.4 | 20 | 20 | 887.0 | 77.0 |
| 30 | 922.1 | 75.2 | 30 | 30 | 999.8 | 110.0 |
| 40 | 942.2 | 64.7 | 40 | 40 | 1051.1 | 94.2 |
| 50 | 946.1 | 68.1 | 50 | 50 | 1071.7 | 57.1 |
| 60 | 968.1 | 79.2 | 60 | 60 | 1110.7 | 102.0 |
| 70 | 1003.1 | 139.8 | 70 | 70 | 1172.4 | 94.5 |

| (c) SBi-LDA | | | | (d) MBi-LDA | | | | | |
|-------------|-----|----------|------|-------------|-----|-----|-----|----------|------|
| S | T | パープレキシティ | 標準偏差 | K | L | S | T | パープレキシティ | 標準偏差 |
| 10 | 10 | 702.1 | 58.8 | 5 | 5 | 10 | 10 | 650.8 | 51.6 |
| 20 | 20 | 684.4 | 31.6 | 5 | 5 | 20 | 20 | 635.1 | 63.1 |
| 30 | 30 | 646.2 | 48.8 | 5 | 5 | 30 | 30 | 635.9 | 46.5 |
| 40 | 40 | 657.3 | 48.2 | 5 | 5 | 40 | 40 | 618.4 | 44.9 |
| 50 | 50 | 642.3 | 51.8 | 5 | 5 | 50 | 50 | 621.1 | 36.3 |
| 60 | 60 | 652.1 | 52.9 | 5 | 5 | 60 | 60 | 602.9 | 38.9 |
| 70 | 70 | 643.1 | 36.8 | 5 | 5 | 70 | 70 | 609.0 | 38.8 |

較を図 5 に示す. 初めに, 従来モデルの LDA, Bi-LDA に対し, 提案モデルの SBi-LDA が優れていることを統計的検定により検証した. 次に, 2つの提案モデル間で比較を行い, MBi-LDA の性能が SBi-LDA よりもさらに優れていることを検証した.

LDA, Bi-LDA と SBi-LDA の比較では, 多重性を考慮

した検定手法である, ダネットの片側検定を用いた. ここで, 本実験における検定全体の有意水準を 5% に保つため, ダネットの片側検定の名義水準は, シダックの方法により 2.53% に設定した. SBi-LDA を対照群とし, LDA, Bi-LDA を処理群として検定を行ったところ, データセット「食品」, 「CD・DVD・楽器」のいずれにおいても, LDA

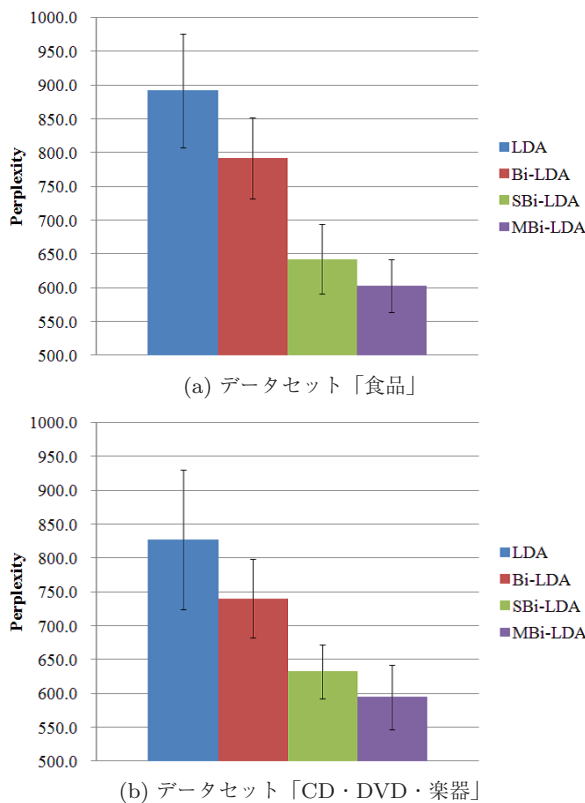


図 5 各データセットにおける各モデルの最良値の比較

Fig. 5 Comparison of the best values of each model.

と SBi-LDA 間, Bi-LDA と SBi-LDA 間に有意差が見られた。なお, 各モデル間の p 値は, どちらのデータセットにおいても 1%未満となった。

次に, SBi-LDA と MBi-LDA の性能の比較を行った。検定には t 検定を用い, ここでの名義水準はダネットの片側検定の場合と同様に 2.53%に設定した。t 検定 (両側検定) の結果, いずれのデータセットにおいても, モデル間に p 値 1%未満となる有意差が見られた。

これらの結果から, 2つのデータセットいずれにおいても, 提案する SBi-LDA の性能が, 他の従来モデルを統計的に上回っており, レビューテキスト解析の手法として有用であることが確認できた。特に LDA は, トピックについて, ユーザやアイテム, 評点について明示的に区別をしないことから, レビューテキストを十分に表現できておらず, 提案するモデルと比較して大きく性能が劣る結果となった。また Bi-LDA は, トピック数が増えると性能が悪化する結果となった。これは, 単語がトピックの組合せから生成されるとする仮定において, トピックの組合せ数が二乗で増加するため, 十分にモデルの学習ができないことが原因と考えられる。それに対し, SBi-LDA では, 各単語はユーザトピックとアイテムトピックのいずれかから生成されると仮定しており, モデルの学習における Bi-LDA の問題点を回避することができた。これにより, レビューテキストにユーザトピック・アイテムトピックを仮定し, モデル化することの有用性を示すことができた。さらに, ユーザト

ピック・アイテムトピックのみを仮定した SBi-LDA に対し, 評点からも単語を生成する仮定を加えた MBi-LDA の性能が優れていることから, テキスト解析モデルに評点を組み込むことの有用性も示された。

また, ユーザトピック・アイテムトピックの数が異なる場合の傾向についても確認を行った。各トピックの数を 10, 20, 30 の 3 通りとし, それぞれの組合せで得られるパープレキシティの比較を行った。なお, 実験におけるパラメータは 4.2.1 項と同様のものを用い, 30 試行の平均値を比較に用いた。それぞれのデータセットから得られた結果を表 4, 表 5 に示す。

表 4, 表 5 から, どちらのデータセットにおいても, 全体的な傾向として, Bi-LDA, SBi-LDA, MBi-LDA の順に性能が向上するという, 先述の実験結果と同様の傾向が得られた。これにより, トピック数の設定方法によらず, 相対的に提案モデルの性能が高いことが確認できた。また, ユーザトピック・アイテムトピック数の組合せについては, 同数の組合せに限らず探索することで, さらに良い結果を得られる可能性があることが確認できた。しかし, トピック数の組合せは膨大なパターンが存在するため, それらを効率良く設定する方法を検討する必要があるといえる。

4.3 定性評価実験

4.3.1 実験方法

初めに, 提案する MBi-LDA において, 仮定したユーザトピック・アイテムトピックが適切に推定できていることを確認した。実験では, LDA と MBi-LDA の各トピックから出現する単語のうち, 確率上位 15 語をそれぞれ抽出した。なお, 単語の抽出においては, 意味を持たない単語や形態素解析の誤解析による単語を, ストップワードとしてあらかじめ除いた。抽出された単語からトピックを解釈し, トピックの推定結果に対する検討を行った。

次に, MBi-LDA において評点から出現する単語について, 各評点における評価語の出現傾向を確認した。この実験では, 各評点の単語上位 100 語を抽出したときの, 肯定的単語/否定的単語の数で評価を行った。肯定的単語/否定的単語の判定は, 日本語評価極性辞書 [13], [17] をそれぞれ参考にしながら, 3名の被験者による主観的判断により行った。辞書を用いて直接判定を行わなかったのは, たとえば, 一般的なテキストでは“早い”が中立的な意味であるのに対し, レビューにおいては“配送が早い”という肯定的な意味を持つ場合があり, 自動で判断することが困難であったためである。なお, 本実験における各モデルのトピック数について, データセット「食品」は, LDA を $T = 40$, MBi-LDA を $(K, L, S, T) = (10, 10, 20, 20)$ とし, データセット「CD・DVD・楽器」は, LDA を $T = 40$, MBi-LDA を $(K, L, S, T) = (5, 5, 20, 20)$ とした。なお, K, L, S, T は付録 A.4 と同様に, 各トピックの総数を表す。

表 4 トピック数が異なる場合の各モデルのパープレキシティの比較 (データセット「食品」)

Table 4 Comparison of perplexities of each model in case the numbers of topics differ (dataset “Food”).

| (a) Bi-LDA | | | | (b) SBi-LDA | | | | (c) MBi-LDA | | | |
|------------|----|----------|------|-------------|----|----------|------|-------------|----|----------|------|
| S | T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 |
| 10 | 20 | 832.4 | 72.5 | 10 | 20 | 658.4 | 46.7 | 10 | 20 | 624.5 | 48.0 |
| 10 | 30 | 869.4 | 68.8 | 10 | 30 | 650.3 | 66.5 | 10 | 30 | 627.5 | 59.3 |
| 20 | 10 | 857.2 | 54.3 | 20 | 10 | 639.9 | 62.2 | 20 | 10 | 644.5 | 45.0 |
| 20 | 30 | 950.3 | 76.2 | 20 | 30 | 643.7 | 41.9 | 20 | 30 | 614.0 | 47.7 |
| 30 | 10 | 936.6 | 55.2 | 30 | 10 | 639.3 | 47.2 | 30 | 10 | 644.3 | 66.8 |
| 30 | 20 | 962.3 | 75.0 | 30 | 20 | 625.8 | 60.1 | 30 | 20 | 612.8 | 43.8 |

表 5 トピック数が異なる場合の各モデルのパープレキシティの比較 (データセット「CD・DVD・楽器」)

Table 5 Comparison of perplexities of each model in case the numbers of topics differ (dataset “CD DVD Instruments”).

| (a) Bi-LDA | | | | (b) SBi-LDA | | | | (c) MBi-LDA | | | |
|------------|----|----------|-------|-------------|----|----------|------|-------------|----|----------|------|
| S | T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 | S | T | パープレキシティ | 標準偏差 |
| 10 | 20 | 877.8 | 90.4 | 10 | 20 | 680.2 | 57.8 | 10 | 20 | 641.3 | 65.9 |
| 10 | 30 | 893.5 | 104.2 | 10 | 30 | 673.4 | 53.8 | 10 | 30 | 665.9 | 54.2 |
| 20 | 10 | 855.4 | 83.7 | 20 | 10 | 675.0 | 57.2 | 20 | 10 | 622.7 | 66.0 |
| 20 | 30 | 968.8 | 92.5 | 20 | 30 | 667.5 | 55.5 | 20 | 30 | 626.4 | 74.6 |
| 30 | 10 | 957.5 | 100.1 | 30 | 10 | 669.5 | 58.0 | 30 | 10 | 643.7 | 64.1 |
| 30 | 20 | 1023.0 | 100.7 | 30 | 20 | 651.7 | 63.9 | 30 | 20 | 632 | 60.0 |

表 6 LDA で推定されたトピックの例 (データセット「食品」)

Table 6 Example of topics inferred by LDA (dataset “Food”).

| T6 | T10 | T13 | T18 | T34 |
|-----|-----|-----|-------|------|
| 届く | 食べる | 値段 | 米 | 袋 |
| 到着 | 家族 | 安い | もち | 送料 |
| 発送 | 魚 | 円 | 炊く | 無料 |
| かかる | 実家 | 送料 | 美味しい | 注文 |
| 配送 | 部分 | 高い | 食べる | 梱 |
| 注文 | 送る | 価格 | モチ | もらう |
| 今日 | られる | 買う | 新米 | 我が家 |
| 時 | 料 | 購入 | 玄米 | くれる |
| 商品 | 自宅 | 込み | 水 | 近く |
| 良い | 調味 | 味 | 産 | 処理 |
| 時間 | みんな | 考える | 県 | 夫婦 |
| 配達 | 刺身 | 量 | ご飯 | 下さる |
| 遅れる | 好評 | 得 | コシヒカリ | かかる |
| 業者 | でる | 買える | ミルク | ポイント |
| 遅い | リンゴ | セール | 粒 | まずい |

4.3.2 結果と考察

LDA と MBi-LDA で得られたトピックについて、トピックから得られる単語の確率上位 15 語を抽出し、トピックの解釈を行った。データセット「食品」で得られる結果について、LDA の結果を表 6 に、MBi-LDA の結果を表 7 にそれぞれ示す。LDA で推定されたトピックを確認する

と、T6 は配送、T13 は価格、T18 は米や餅についてのトピックであると解釈することができる。これらのトピックは、意味ごとにまとまってはいるものの、アイテムに依存しないトピックである T6 と T13、ユーザに依存しないトピックである T18 が混在する結果となった。また、T10 と T34 のように、解釈の難しいトピックも数多く得られた。

一方、MBi-LDA で推定されたトピックの単語を確認すると、ユーザトピックについては、S1 は料金、S5 は対象者、S10 は贈答、S12 は店の対応をそれぞれ表すトピックと解釈でき、ユーザがアイテムによらず保持する着眼点をトピックとして推定することができていた。また、アイテムトピックについても、T5 はキムチ、T8 は米、T18 は麺、T19 は肉製品をそれぞれ表すトピックと解釈でき、アイテム固有のトピックを推定することができていた。LDA の結果と比較して、解釈が容易なトピックが多く得られた点、ユーザ・アイテムそれぞれに特有のトピックを明示的に推定できている点から、提案する MBi-LDA が想定どおり働いており、良好な結果を示すことが確認できた。

データセット「CD・DVD・楽器」についても、同様の確認を行った。LDA の結果を表 8 に、MBi-LDA の結果を表 9 にそれぞれ示す。LDA で推定されたトピックを確認すると、T19 は購入手続き、T27 は映画、T30 は価格、T37 は CD アルバム、T39 はギターについてのトピックで

表 7 MBI-LDA で推定されたトピックの例 (データセット「食品」)
Table 7 Example of topics inferred by MBI-LDA (dataset “Food”).

(a) ユーザトピック

| S1 | S5 | S10 | S12 |
|------|------|-------|-------|
| 円 | 子供 | 父 | 商品 |
| 送料 | 息子 | プレゼント | 購入 |
| 個 | 達 | 祖母 | 対応 |
| 買う | 娘 | おすすめ | 満足 |
| 込み | 家族 | ごはん | いただく |
| 価格 | 私 | ギフト | 大変 |
| 倍 | 妻 | 梱包 | 良い |
| スーパー | 我が家 | 年末 | 店 |
| セット | 歳 | 助かる | 丁寧 |
| 無料 | 点 | 仙台 | 価格 |
| ポイント | 全員 | 多い | 迅速 |
| 高い | 的 | 丁寧 | 他 |
| 袋 | メン | 敬老 | 的 |
| キロ | ショップ | 年寄り | いただける |
| 試し | 頂く | パック | お願い |

(b) アイテムトピック

| T5 | T8 | T18 | T19 |
|-----|-------|-----|-------|
| キムチ | 米 | うどん | ハム |
| 辛い | 炊く | 麵 | ベーコン |
| 漬け | 新米 | そば | ソーセージ |
| 冷 | 産 | つゆ | 豚 |
| 味 | 安い | 蕎麦 | 脂身 |
| 漬物 | 値段 | 茹でる | ウインナー |
| 大根 | コシヒカリ | コシ | 多い |
| 白菜 | 粒 | 讃岐 | 脂 |
| 酸味 | キロ | めん | 切り落とす |
| 漬ける | ミルク | こす | イペリコ |
| 韓国 | 美味しい | 人前 | サラダ |
| 好き | 購入 | 太い | 使う |
| 好み | ひかり | ゆでる | 得 |
| 甘い | 洗米 | 時間 | 炒める |
| 高菜 | 玄米 | 腰 | スライス |

あると解釈できる。ただし、T19, T30 はアイテムに依存せずユーザに関係のあるトピック、T27, T37, T39 はアイテムに依存するトピックであり、データセット「食品」と同様、それらが混在する結果となった。一方、MBI-LDA で推定されたトピックの単語において、ユーザトピックについては、S7 は対象者、S16 は質、S17 は買い得さ、S18 は時代をそれぞれ表すトピックと解釈でき、ユーザがアイテムによらず保持する着眼点をトピックとして推定することができていた。アイテムトピックについても、T8 は主題歌、T9 は特撮ヒーロー、T14 はライブ DVD、T18 はギターをそれぞれ表すアイテム固有のトピックとして推定することができていた。

次に、MBI-LDA において評点から出現する単語について、評点の大小による評価語の出現傾向を確認した。各

表 8 LDA で推定されたトピックの例 (データセット「CD・DVD・楽器」)

Table 8 Example of topics inferred by LDA (dataset “CD DVD Instruments”).

| T19 | T27 | T30 | T37 | T39 |
|-------|--------|------|------|-----|
| 届く | 作品 | 購入 | アルバム | ギター |
| 予約 | 映画 | 商品 | 楽曲 | 弦 |
| 発売 | 人 | 安い | 作品 | 使う |
| メール | ストーリー | 値段 | ベスト | 購入 |
| 楽天 | 者 | 内容 | 収録 | 使用 |
| 注文 | 原作 | 価格 | 時代 | ピック |
| 発送 | 出演 | 満足 | 曲 | 良い |
| 店 | ゲーム | 良い | 年 | 弾く |
| キャンセル | 好き | 対応 | 感じる | 楽器 |
| 到着 | アクション | 店 | ファン | 商品 |
| ボックス | 最後 | ショップ | 初期 | 音 |
| 早い | アンパンマン | 他 | 期待 | 自分 |
| 待つ | 展開 | 中古 | 購入 | 部分 |
| 初回 | 感じ | 残念 | アップ | 感 |
| 対応 | 映像 | 買う | CD | 感じ |

データセットの各評点について、上位 100 語に含まれる肯定的単語/否定的単語の数の一覧を、表 10 に示す。また、例として、被験者 1 が評価語として抽出した単語を表 11 に示す。表 10 の結果より、評点から出現する単語について、高い評点であるほど肯定的単語が多く出現し、低い評点であるほど否定的単語が多く出現していることが確認できた。これは本論文で提案した MBI-LDA によるモデルの構築が適切であったといえる結果となった。

以上の結果から、レビューテキストにユーザトピック・アイテムトピックの仮定をした提案モデルにおいて、それらをとらえたトピック推定が実際に実現できていることが、定性的にも確認できた。また、提案モデルで仮定した評点についても、高い評点からは肯定的な単語、低い評点からは否定的な単語が多く得られており、良好な結果を示すことができた。しかし一方、トピック・評点ともに、“購入”、“食べる”、“見る”などの各カテゴリで一般的に用いられる単語も、上位に一定数出現していることが確認できた。これら一般語はトピック・評点に依存しないため、それらが上位に出現することは望ましくない。そのため今後、トピック・評点に依存せず一般語が出現する過程をモデルに組み込むことで、一般語分布の推定を行うことができ、さらなる性能向上が可能となると考えられる。

4.4 各モデルの学習時間

4.2, 4.3 節において、提案モデルの性能が、定量的・定性的に従来手法を上回ることを示した。本節では、各モデルの学習時間についての比較・検討を行う。本論文では、64 bit 環境 (CPU: Intel®Core™i7-2600 CPU (3.40 GHz), メモリ (RAM): 16.0 GB) を用いて実験を行った。なお、イ

表 10 各評点上位 100 語中の肯定的単語/否定的単語の数

Table 10 Number of positive/negative words appeared in top 100 words of each rate.

(a) データセット「食品」

| | 被験者 1 | | 被験者 2 | | 被験者 3 | | 平均 | |
|------|-------|-----|-------|-----|-------|-----|------|-----|
| | 肯定的 | 否定的 | 肯定的 | 否定的 | 肯定的 | 否定的 | 肯定的 | 否定的 |
| 評点 3 | 10 | 4 | 11 | 7 | 6 | 9 | 9.0 | 6.7 |
| 評点 4 | 13 | 4 | 16 | 5 | 15 | 7 | 14.7 | 5.3 |
| 評点 5 | 23 | 0 | 25 | 1 | 22 | 2 | 23.3 | 1.0 |

(b) データセット「CD・DVD・楽器」

| | 被験者 1 | | 被験者 2 | | 被験者 3 | | 平均 | |
|------|-------|-----|-------|-----|-------|-----|------|------|
| | 肯定的 | 否定的 | 肯定的 | 否定的 | 肯定的 | 否定的 | 肯定的 | 否定的 |
| 評点 3 | 9 | 10 | 10 | 10 | 5 | 12 | 8.0 | 10.7 |
| 評点 4 | 16 | 2 | 17 | 2 | 10 | 1 | 14.3 | 1.7 |
| 評点 5 | 19 | 0 | 19 | 1 | 16 | 0 | 18.0 | 0.3 |

表 9 MBI-LDA で推定されたトピックの例 (データセット「CD・DVD・楽器」)

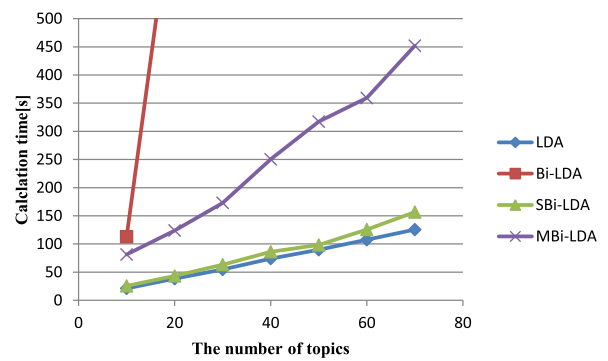
Table 9 Example of topics inferred by MBI-LDA (dataset “CD DVD Instruments”).

(a) ユーザトピック

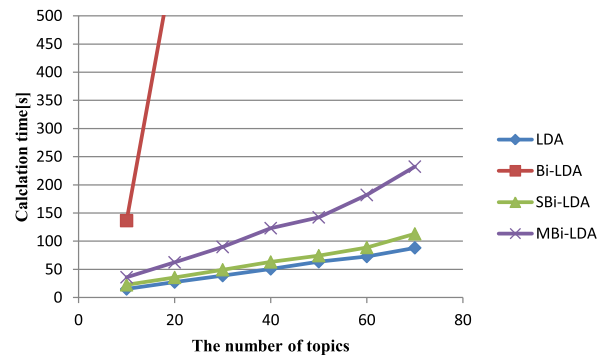
| S7 | S16 | S17 | S18 |
|-------|------|--------|------|
| 娘 | DVD | 送料 | 懐かしい |
| 子供 | 音質 | 無料 | 当時 |
| 購入 | 画面 | 商品 | 思い出す |
| 頼む | サイズ | 良い | 年代 |
| 歳 | CD | キャンペーン | 昔 |
| 気に入る | 画質 | ショップ | 出演 |
| 見る | 汗 | ポイント | 有る |
| 大好き | 私 | 価格 | 時代 |
| 息子 | ドラマ | 購入 | 行く |
| プレゼント | PV | 特典 | 中古 |
| ファン | ディスク | 倍 | TV |
| 主人 | BD | ケース | 解る |
| 家族 | 違う | 梱包 | 充分 |
| 親 | 今回 | お願い | 買い |
| 達 | 勧める | 問題ない | 青春 |

(b) アイテムトピック

| T8 | T9 | T14 | T18 |
|--------|------------|-------|------|
| アニメ | ライダー | ライブ | 弦 |
| サントラ | ウルトラマン | DVD | ギター |
| CD | 幕 | 発売 | ピック |
| 主題歌 | シンケンジャー | 行く | 弾く |
| 収録 | 短い | 嵐 | やすい |
| テーマ | ウルトラ | 初回 | 交換 |
| ドラマ | 続く | 予約 | ウクレレ |
| ゲーム | 昭和 | 楽しみ | アコギ |
| オープニング | 平成 | コンサート | 自分 |
| 版 | 対談 | 今回 | いい |
| エンディング | ディレクターズカット | 歌 | セット |
| 弾 | 登場 | 参加 | 形 |
| 言える | 王 | 今年 | 硬い |
| イベント | 本編 | 楽しい | 始める |
| 映画 | メビウス | ツアー | 感じる |



(a) データセット「食品」



(b) データセット「CD・DVD・楽器」

図 6 各データセットにおける各モデルの計算時間の比較

Fig. 6 Comparison of the best values of each model.

テレーション回数, トピック数などのパラメータは 4.2 節と同様の値を用いた。

各モデルの学習時間について, 各データセットの結果を図 6 に示す. 実験結果から, 最もシンプルなモデルである LDA が, 学習時間において最速であることが確認できた. また, SBI-LDA も LDA とほぼ変わらない学習時間であった. 一方, Bi-LDA は, トピック数の増加に対して急激に学習時間が増加していることが確認できる. これは, Bi-LDA において, 単語がユーザトピック・アイテムトピックのペアから生成されると仮定しているため, 各ペアについて単語の出現分布を学習しなければならない点, またトピック

表 11 被験者 1 が評価した肯定的単語/否定的単語の例
Table 11 Example of positive/negative words evaluated by subject 1.

(a) データセット「食品」

| | 評価語 | 語数 | 単語例 |
|------|-----|----|--|
| 評点 3 | 肯定的 | 10 | 美味しい, 良い, 期待, おいしい, 安い, いい, 好き, できる, 好み, やすい |
| | 否定的 | 4 | 残念, 高い, 悪い, いらぬ |
| 評点 4 | 肯定的 | 13 | 美味しい, おいしい, 良い, 安い, いい, 好き, できる, 満足, やすい, 期待, 大好き, 嬉しい, 好み |
| | 否定的 | 4 | 残念, 高い, いらぬ, 大変 |
| 評点 5 | 肯定的 | 23 | 美味しい, おいしい, 良い, 安い, 便利, できる, 満足, いい, 好き, 楽しい, よい, 嬉しい, 大好き, やすい, 最高, 一番, 得, 期待, 安心, 助かる, 喜ぶ, すごい, 気に入る |
| | 否定的 | 0 | (なし) |

(b) データセット「CD・DVD・楽器」

| | 評価語 | 語数 | 単語例 |
|------|-----|----|---|
| 評点 3 | 肯定的 | 9 | 良い, 安い, 期待, よい, できる, いい, 好き, 楽しみ, 満足 |
| | 否定的 | 10 | 残念, 悪い, バカ, カス, いらぬ, 高い, 買わない, 遅い, 仕方ない, ショック |
| 評点 4 | 肯定的 | 16 | 良い, 安い, よい, できる, いい, 満足, 楽しみ, 好き, 嬉しい, 楽しめる, 気に入る, 喜ぶ, 面白い, 期待, 楽しむ, やすい |
| | 否定的 | 2 | 残念, いらぬ |
| 評点 5 | 肯定的 | 19 | よい, 安い, 満足, 楽しみ, 好き, 最高, 嬉しい, 早い, 喜ぶ, 良い, 大好き, いい, ♪, 素敵, 丁寧, 感動, すごい, 感謝, 気に入る |
| | 否定的 | 0 | (なし) |

の組合せ数が二乗で増加する点から、学習時間が増加していると考えられる。これに対し SBi-LDA は、ユーザトピック、アイテムトピックそれぞれの単語分布の学習であるため、Bi-LDA で生じる組合せ数増大の問題を回避している。また、MBi-LDA は、評点生成部分と単語生成部分の 2 つの生成機構を有すること、評点生成部分にトピックの組合せの機構が存在することから、LDA や SBi-LDA と比較して 2~3 倍程度の学習時間となった。しかしこれは定数倍の増加であるため、MBi-LDA は学習時間の面でそれらのモデルに大きくは劣っていないといえる。

5. まとめ

本論文では、レビューテキストの解析を目的とした新たなトピックモデルの構築を試みた。初めに、評点の予測モデルである Bi-LDA を、テキスト解析を想定した SBi-LDA に発展させた。SBi-LDA では、Bi-LDA と同様にユーザトピックとアイテムトピックを仮定し、それらが単語の生成過程において、確率的に選択されるようモデル化を行った。次に、SBi-LDA をさらに発展させ、ユーザトピック・アイテムトピックに加えて、評点も単語の生成過程に組み込んだ MBi-LDA を構築し、提案した。

実験では、提案モデルを実際のレビューデータに適用し、モデルの性能を定量的/定性的に評価した。定量評価実験

では、SBi-LDA が LDA, Bi-LDA と比較してパープレキシティが 10% 以上下がること、さらに MBi-LDA が SBi-LDA よりパープレキシティの面で優れていることを、統計的検定により示した。定性評価実験では、MBi-LDA が推定したユーザトピック・アイテムトピックを目視で確認し、モデルの仮定どおり、ユーザに関係するトピックとアイテムに関係するトピックを明示的に区別して推定できることを示した。また、評点から出現する単語について、評点の大小による評価語の出現傾向について検証した結果、高い評点からは肯定的な単語、低い評点からは否定的な単語が出現しやすいことが確認できた。また、学習時間においても、MBi-LDA はベースラインである LDA と比較して、2~3 倍程度の増加に抑えられていることが確認できた。これらの結果から、本論文で提案した MBi-LDA は、レビューテキストの解析手法として有用であることを示すことができた。

今後の課題として、一般語のモデル化、トピック数の自動決定、時系列への対応があげられる。トピック・評点に依存しないと考えられる一般語を独自にモデル化することで、モデルの性能向上を図ることができると考えられる。また、提案モデルではそれぞれのトピック数を事前に設定する必要があるが、近年階層ディリクレ過程 [18] を用いてトピック数を自動で決定することのできるモデルが多く提

案されており [19], [20], 本提案モデルにも適用することができると思われる。さらに, ユーザの嗜好は時間に依存して変化することも考えられるため [21], [22], それらに対応することで, いっそうのモデルの性能向上を図ることが課題である。

参考文献

[1] Nasukawa, T. and Yi, J.: Sentiment analysis: Capturing favorability using natural language processing, *Proc. 2nd International Conference on Knowledge Capture*, ACM, pp.70–77 (2003).

[2] Pang, B. and Lee, L.: Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, Vol.2, No.1-2, pp.1–135 (2008).

[3] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GroupLens: An open architecture for collaborative filtering of netnews, *Proc. 1994 ACM Conference on Computer Supported Cooperative Work, CSCW '94*, pp.175–186, ACM (1994).

[4] 土方嘉徳: 1. 嗜好抽出と情報推薦技術 (嗜好抽出・情報推薦の基礎理論, 特集: 利用者の好みをとらえ活かす—嗜好抽出技術の最前線), *情報処理*, Vol.48, No.9, pp.957–965 (2007).

[5] Ganu, G., Elhadad, N. and Marian, A.: Beyond the Stars: Improving Rating Predictions using Review Text Content., *WebDB* (2009).

[6] Wang, C. and Blei, D.M.: Collaborative topic modeling for recommending scientific articles, *Proc. 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, pp.448–456, ACM (2011).

[7] Hofmann, T.: Probabilistic latent semantic analysis, *Proc. 15th Conference on Uncertainty in Artificial Intelligence, UAI'99*, pp.289–296, Morgan Kaufmann Publishers Inc. (1999).

[8] Blei, D.M., Ng, A.Y. and Jordan, M.I.: Latent dirichlet allocation, *J. Mach. Learn. Res.*, Vol.3, pp.993–1022 (2003).

[9] Titov, I. and McDonald, R.: Modeling online reviews with multi-grain topic models, *Proc. 17th International Conference on World Wide Web*, pp.111–120, ACM (2008).

[10] Lin, C. and He, Y.: Joint sentiment/topic model for sentiment analysis, *Proc. 18th ACM Conference on Information and Knowledge Management*, pp.375–384, ACM (2009).

[11] Porteous, I., Bart, E. and Welling, M.: Multi-HDP: A Non Parametric Bayesian Model for Tensor Factorization, Vol.8, pp.1487–1490 (2008).

[12] Pang, B., Lee, L. and Vaithyanathan, S.: Thumbs up?: Sentiment classification using machine learning techniques, *Proc. ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, Association for Computational Linguistics, pp.79–86 (2002).

[13] 小林のぞみ, 乾健太郎, 松本裕治, 立石健二, 福島俊一: テキストマイニングによる評価現象の収集, *情報処理学会研究報告, 自然言語処理研究会報告*, Vol.2003, No.23, pp.77–84 (2003).

[14] 落合恵理香, 小林一郎: 商品の評価を対象としたレビュー文書の分析, *言語処理学会第 18 回年次大会発表論文集*, Vol.2012, pp.1176–1179, 言語処理学会 (2012).

[15] Mnih, A. and Salakhutdinov, R.: Probabilistic matrix

factorization, *Advances in Neural Information Processing Systems*, pp.1257–1264 (2007).

[16] Salakhutdinov, R. and Mnih, A.: Bayesian probabilistic matrix factorization using Markov chain Monte Carlo, *Proc. 25th International Conference on Machine Learning*, pp.880–887, ACM (2008).

[17] 東山昌彦: 述語の選択選好性に着目した名詞評価極性の獲得, *言語処理学会第 14 回年次大会論文集*, pp.584–587 (2008).

[18] Teh, Y.W., Jordan, M.I., Beal, M.J. and Blei, D.M.: Hierarchical dirichlet processes, *Journal of the American Statistical Association*, Vol.101, No.476 (2006).

[19] Ahmed, A. and Xing, E.P.: Timeline: A dynamic hierarchical Dirichlet process model for recovering birth/death and evolution of topics in text stream, *Proc. 26th Conference on Uncertainty in Artificial Intelligence*, pp.20–29 (2010).

[20] Li, W. and Blei, David, M.A.: Nonparametric bayes pachinko allocation, *Proc. 23rd Conference on Uncertainty in Artificial Intelligence* (2007).

[21] 岩田具治, 渡部晋治, 山田武士, 上田修功: 購買行動解析のためのトピック追跡モデル, *電子情報通信学会論文誌 D, 情報・システム*, Vol.J93-D, No.6, pp.978–987 (2010).

[22] 佐々木謙太郎, 吉川大弘, 古橋 武: Twitter におけるユーザの興味と話題の時間発展を考慮したオンライン学習可能なトピックモデルの提案, *情報処理学会論文誌 数理モデル化と応用*, Vol.7, No.1, pp.53–60 (2014).

付 録

A.1 LDA のグラフィカルモデルにおけるパラメータ

D は総文書数, N_d は文書 d に含まれる総単語数, T は総トピック数を表す. w は単語を, z は各単語に割り当てられるトピックを表す. θ は各文書のトピック分布, ϕ は各トピックの単語分布であり, α, β はそれぞれの分布のハイパーパラメータである.

A.2 Bi-LDA のグラフィカルモデルにおけるパラメータ

U は総ユーザ数, M は総アイテム数, S は総ユーザトピック数, T は総アイテムトピック数を表す. z^U はユーザトピック, z^M はアイテムトピック, r は評点を表す. θ^U は各ユーザのユーザトピック分布, θ^M は各アイテムのアイテムトピック分布, ϕ はユーザトピックとアイテムトピックの各ペアの評点分布であり, $\alpha^U, \alpha^M, \beta$ はそれぞれの分布のハイパーパラメータである.

A.3 SBi-LDA のグラフィカルモデルにおけるパラメータ

U は総ユーザ数, M は総アイテム数, S は総ユーザトピック数, T は総アイテムトピック数, $N_{u,m}$ はユーザ u がアイテム m に残したレビューテキスト中の総単語数を表す. z^U はユーザトピック, z^M はアイテムトピック, w は単語, y は各単語の出現を制御するスイッチ変数を表す.

θ^U は各ユーザのユーザトピック分布, θ^M は各アイテムのアイテムトピック分布, ϕ^U は各ユーザトピックの単語分布, ϕ^M は各アイテムトピックの単語分布, π は各ユーザのスイッチ変数分布を表す. $\alpha^U, \alpha^M, \beta^U, \beta^M, \gamma$ はそれぞれの分布のハイパーパラメータである.

A.4 MBI-LDA における評点付きレビューテキストの生成過程

MBI-LDA における, 評点付きレビューテキストの生成過程は以下のようにモデル化される.

- (1) 評点ユーザトピック $k \in \{1, \dots, K\}$ と評点アイテムトピック $l \in \{1, \dots, L\}$ の各ペアについて, 評点分布 $\eta_{[k,l]}$ をディリクレ分布に従って生成する.

$$\eta_{[k,l]} \sim \text{Dir}(\epsilon)$$

- (2) 単語ユーザトピック $s \in \{1, \dots, S\}$ について, 単語分布 ϕ_s^U をディリクレ分布に従って生成する.

$$\phi_s^U \sim \text{Dir}(\beta^U)$$

- (3) 単語アイテムトピック $t \in \{1, \dots, T\}$ について, 単語分布 ϕ_t^M をディリクレ分布に従って生成する.

$$\phi_t^M \sim \text{Dir}(\beta^M)$$

- (4) 評点 $e \in \{1, \dots, R\}$ について, 単語分布 ϕ_e^R をディリクレ分布に従って生成する.

$$\phi_e^R \sim \text{Dir}(\beta^R)$$

- (5) 各ユーザ $u \in \{1, \dots, U\}$ について:

- (a) 評点ユーザトピック分布 θ_u^U をディリクレ分布に従って生成する.

$$\theta_u^U \sim \text{Dir}(\alpha^U)$$

- (b) 単語ユーザトピック分布 μ_u^U をディリクレ分布に従って生成する.

$$\mu_u^U \sim \text{Dir}(\delta^U)$$

- (c) スイッチ変数分布 π_u をディリクレ分布に従って生成する.

$$\pi_u \sim \text{Dir}(\gamma)$$

- (6) 各アイテム $m \in \{1, \dots, M\}$ について:

- (a) 評点アイテムトピック分布 θ_m^M をディリクレ分布に従って生成する.

$$\theta_m^M \sim \text{Dir}(\alpha^M)$$

- (b) 単語アイテムトピック分布 μ_m^M をディリクレ分布に従って生成する.

$$\mu_m^M \sim \text{Dir}(\delta^M)$$

- (7) ユーザ u , アイテム m のレビュー i について:

- (a) 評点ユーザトピック z_i^{RU} を多項分布に従って生成する.

$$z_i^{\text{RU}} \sim \text{Mult}(\theta_u^U)$$

- (b) 評点アイテムトピック z_i^{RM} をアイテムトピックに基づく多項分布に従って生成する.

$$z_i^{\text{RM}} \sim \text{Mult}(\theta_m^M)$$

- (c) 評点 r_i を評点に基づく多項分布に従って生成する.

$$r_i \sim \text{Mult}(\eta_{[z_i^{\text{RU}}, z_i^{\text{RM}}]})$$

- (d) レビューテキスト中に出現する単語 $j \in \{1, \dots, N_i\}$ について:

- (i) スイッチ変数 y_{ij} を多項分布に従って生成する.

$$y_{ij} \sim \text{Mult}(\pi_u)$$

- (ii) $y_{ij} = 0$ のとき:

- (A) 単語ユーザトピック分布に基づき, 単語ユーザトピック z_{ij}^{WU} を多項分布に従って生成する.

$$z_{ij}^{\text{WU}} \sim \text{Mult}(\mu_u^U)$$

- (B) 単語ユーザトピックの単語分布に基づき, 単語 w_{ij} を多項分布に従って生成する.

$$w_{ij} \sim \text{Mult}(\phi_{z_{ij}^{\text{WU}}}^U)$$

- (iii) $y_{ij} = 1$ のとき:

- (A) 単語アイテムトピック分布に基づき, 単語アイテムトピック z_{ij}^{WM} を多項分布に従って生成する.

$$z_{ij}^{\text{WM}} \sim \text{Mult}(\mu_m^M)$$

- (B) 単語アイテムトピックの単語分布に基づき, 単語 w_{ij} を多項分布に従って生成する.

$$w_{ij} \sim \text{Mult}(\phi_{z_{ij}^{\text{WM}}}^M)$$

- (iv) $y_{ij} = 2$ のとき:

- (A) 評点の単語分布に基づき, 単語 w_{ij} を多項分布に従って生成する.

$$w_{ij} \sim \text{Mult}(\phi_{r_i}^R)$$

ここで, $\text{Dir}(\cdot)$ はディリクレ分布, $\text{Mult}(\cdot)$ は多項分布を表し, $\alpha^U, \alpha^M, \delta^U, \delta^M, \beta^U, \beta^M, \beta^R, \gamma, \epsilon$ はそれぞれのディリクレ分布におけるハイパーパラメータである. また, K は総評点ユーザトピック数, L は総評点アイテムトピック数, S は総単語ユーザトピック数, T は総単語アイテムトピック数, U は総ユーザ数, M は総アイテム数, R は評点の種類数, N_i はレビュー i のテキストの総単語数を表す.

A.5 MBI-LDA における潜在変数の更新式の導出

データにおける評点集合を \mathbf{r} , それぞれの評点に対応する評点ユーザトピック, 評点アイテムトピックを $\mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}$ とし, データにおける単語集合を \mathbf{w} , それぞれの単語に対応する単語ユーザトピック, 単語アイテムトピック, スイッチ変数を $\mathbf{z}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y}$ とする. それらの結合分布は $p(\mathbf{r}, \mathbf{w}, \mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \mathbf{z}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y} | \alpha^U, \alpha^M, \beta^U, \beta^M, \beta^R, \delta^U, \delta^M, \gamma, \epsilon)$ で表される. ここで, トピック, 評点, 単語の独立性により, この結合分布は式 (A.1) のように表すこと

ができ、式 (A.1) の各項はそれぞれ式 (A.2)~式 (A.6) のように表せる。

$$\begin{aligned}
 & p(\mathbf{r}, \mathbf{w}, \mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \mathbf{z}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y} \\
 & \quad |\alpha^{\text{U}}, \alpha^{\text{M}}, \beta^{\text{U}}, \beta^{\text{M}}, \beta^{\text{R}}, \delta^{\text{U}}, \delta^{\text{M}}, \gamma, \epsilon) \\
 & = p(\mathbf{r}|\mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \epsilon) p(\mathbf{w}^{(\text{U})}|\mathbf{y}^{(\text{U})}, \mathbf{z}^{\text{WU}}, \beta^{\text{U}}) \\
 & \quad p(\mathbf{w}^{(\text{M})}|\mathbf{y}^{(\text{M})}, \mathbf{z}^{\text{WM}}, \beta^{\text{M}}) p(\mathbf{w}^{(\text{R})}|\mathbf{y}^{(\text{R})}, \mathbf{r}, \beta^{\text{R}}) \\
 & \quad p(\mathbf{z}^{\text{RU}}|\alpha^{\text{U}}) p(\mathbf{z}^{\text{RM}}|\alpha^{\text{M}}) p(\mathbf{z}^{\text{WU}}|\delta^{\text{U}}) p(\mathbf{z}^{\text{WM}}|\delta^{\text{M}}) \\
 & \quad p(\mathbf{y}|\gamma) \tag{A.1}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{r}|\mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \epsilon) \\
 & = \left(\frac{\Gamma(R\epsilon)}{\Gamma(\epsilon)^R} \right)^{K \cdot L} \prod_{[k,l]=[1,1]}^{[K,L]} \frac{\prod_{e=1}^R \Gamma(n_{[k,l]e} + \epsilon)}{\Gamma(n_{[k,l]} + R\epsilon)} \tag{A.2}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{w}^{(\text{U})}|\mathbf{y}^{(\text{U})}, \mathbf{z}^{\text{WU}}, \beta^{\text{U}}) \\
 & = \left(\frac{\Gamma(V\beta^{\text{U}})}{\Gamma(\beta^{\text{U}})^V} \right)^S \prod_{s=1}^S \frac{\prod_{v=1}^V \Gamma(n_{sv} + \beta^{\text{U}})}{\Gamma(n_s + V\beta^{\text{U}})} \tag{A.3}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{w}^{(\text{M})}|\mathbf{y}^{(\text{M})}, \mathbf{z}^{\text{WM}}, \beta^{\text{M}}) \\
 & = \left(\frac{\Gamma(V\beta^{\text{M}})}{\Gamma(\beta^{\text{M}})^V} \right)^T \prod_{t=1}^T \frac{\prod_{v=1}^V \Gamma(n_{tv} + \beta^{\text{M}})}{\Gamma(n_t + V\beta^{\text{M}})} \tag{A.4}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{w}^{(\text{R})}|\mathbf{y}^{(\text{R})}, \mathbf{r}, \beta^{\text{R}}) \\
 & = \left(\frac{\Gamma(V\beta^{\text{R}})}{\Gamma(\beta^{\text{R}})^V} \right)^R \prod_{e=1}^R \frac{\prod_{v=1}^V \Gamma(n_{ev} + \beta^{\text{R}})}{\Gamma(n_e + V\beta^{\text{R}})} \tag{A.5}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{z}^{\text{RU}}|\alpha^{\text{U}}) \\
 & = \left(\frac{\Gamma(K\alpha^{\text{U}})}{\Gamma(\alpha^{\text{U}})^K} \right)^U \prod_{u=1}^U \frac{\prod_{k=1}^K \Gamma(n_{uk} + \alpha^{\text{U}})}{\Gamma(n_u + K\alpha^{\text{U}})} \tag{A.6}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{z}^{\text{RM}}|\alpha^{\text{M}}) \\
 & = \left(\frac{\Gamma(L\alpha^{\text{M}})}{\Gamma(\alpha^{\text{M}})^L} \right)^M \prod_{m=1}^M \frac{\prod_{l=1}^L \Gamma(n_{ml} + \alpha^{\text{M}})}{\Gamma(n_m + L\alpha^{\text{M}})} \tag{A.7}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{z}^{\text{WU}}|\delta^{\text{U}}) \\
 & = \left(\frac{\Gamma(S\delta^{\text{U}})}{\Gamma(\delta^{\text{U}})^S} \right)^U \prod_{u=1}^U \frac{\prod_{s=1}^S \Gamma(n_{us} + \delta^{\text{U}})}{\Gamma(n_u + S\delta^{\text{U}})} \tag{A.8}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{z}^{\text{WM}}|\delta^{\text{M}}) \\
 & = \left(\frac{\Gamma(T\delta^{\text{M}})}{\Gamma(\delta^{\text{M}})^T} \right)^M \prod_{m=1}^M \frac{\prod_{t=1}^T \Gamma(n_{mt} + \delta^{\text{M}})}{\Gamma(n_m + T\delta^{\text{M}})} \tag{A.9}
 \end{aligned}$$

$$\begin{aligned}
 & p(\mathbf{y}|\gamma) \\
 & = \left(\frac{\Gamma(3\gamma)}{\Gamma(\gamma)^3} \right)^U \prod_{u=1}^U \frac{\prod_{a=1}^3 \Gamma(n_{ua} + \gamma)}{\Gamma(n_u + 3\gamma)} \tag{A.10}
 \end{aligned}$$

ただし、 $u, m, k, l, s, t, e, v, a$ はそれぞれユーザ、ア

イテム、評点ユーザトピック、評点アイテムトピック、単語ユーザトピック、単語アイテムトピック、評点、単語、スイッチ変数のインデックスを表しており、 n はインデックスについての合計出現回数を示している。 V は単語の種類数である。

トピックの推定には、ギブスサンプリングを用いる。ギブスサンプリングにおけるトピックの更新式は以下で表される。

$$\begin{aligned}
 & p(z_i^{\text{RU}} = k, z_i^{\text{RM}} = l | \mathbf{r}, \mathbf{w}, \mathbf{z}_{\setminus i}^{\text{RU}}, \mathbf{z}_{\setminus i}^{\text{RM}}, \mathbf{z}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y}) \\
 & \propto \frac{p(\mathbf{r}|\mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}})}{p(\mathbf{r}_{\setminus i}|\mathbf{z}_{\setminus i}^{\text{RU}}, \mathbf{z}_{\setminus i}^{\text{RM}})} \cdot \frac{p(\mathbf{z}^{\text{RU}})}{p(\mathbf{z}_{\setminus i}^{\text{RU}})} \cdot \frac{p(\mathbf{z}^{\text{RM}})}{p(\mathbf{z}_{\setminus i}^{\text{RM}})} \\
 & = \frac{n_{[k,l]e}^i + \epsilon}{n_{[k,l]}^i + R\epsilon} \cdot \frac{n_{uk}^i + \alpha^{\text{U}}}{n_u^i + K\alpha^{\text{U}}} \cdot \frac{n_{ml}^i + \alpha^{\text{M}}}{n_m^i + L\alpha^{\text{M}}} \tag{A.11}
 \end{aligned}$$

$$\begin{aligned}
 & p(z_{ij}^{\text{WU}} = s, y_{ij} = 0 | \mathbf{r}, \mathbf{w}, \mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \mathbf{z}_{\setminus ij}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y}_{\setminus ij}) \\
 & \propto \frac{p(\mathbf{w}^{(\text{U})}|\mathbf{y}^{(\text{U})}, \mathbf{z}^{\text{WU}})}{p(\mathbf{w}_{\setminus ij}^{(\text{U})}|\mathbf{y}_{\setminus ij}^{(\text{U})}, \mathbf{z}_{\setminus ij}^{\text{WU}})} \cdot \frac{p(\mathbf{z}^{\text{WU}})}{p(\mathbf{z}_{\setminus ij}^{\text{WU}})} \cdot \frac{p(\mathbf{y})}{p(\mathbf{y}_{\setminus ij})} \\
 & = \frac{n_{sv}^{ij} + \beta^{\text{U}}}{n_s^{ij} + V\beta^{\text{U}}} \cdot \frac{n_{us}^{ij} + \delta^{\text{U}}}{n_u^{ij} + S\delta^{\text{U}}} \cdot \frac{n_{u0}^{ij} + \gamma}{n_u^{ij} + 3\gamma} \tag{A.12}
 \end{aligned}$$

$$\begin{aligned}
 & p(z_{ij}^{\text{WM}} = t, y_{ij} = 1 | \mathbf{r}, \mathbf{w}, \mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \mathbf{z}^{\text{WU}}, \mathbf{z}_{\setminus ij}^{\text{WM}}, \mathbf{y}_{\setminus ij}) \\
 & \propto \frac{p(\mathbf{w}^{(\text{M})}|\mathbf{y}^{(\text{M})}, \mathbf{z}^{\text{WM}})}{p(\mathbf{w}_{\setminus ij}^{(\text{M})}|\mathbf{y}_{\setminus ij}^{(\text{M})}, \mathbf{z}_{\setminus ij}^{\text{WM}})} \cdot \frac{p(\mathbf{z}^{\text{WM}})}{p(\mathbf{z}_{\setminus ij}^{\text{WM}})} \cdot \frac{p(\mathbf{y})}{p(\mathbf{y}_{\setminus ij})} \\
 & = \frac{n_{tv}^{ij} + \beta^{\text{U}}}{n_t^{ij} + V\beta^{\text{M}}} \cdot \frac{n_{mt}^{ij} + \delta^{\text{M}}}{n_u^{ij} + T\delta^{\text{M}}} \cdot \frac{n_{u1}^{ij} + \gamma}{n_u^{ij} + 3\gamma} \tag{A.13}
 \end{aligned}$$

$$\begin{aligned}
 & p(y_{ij} = 2 | \mathbf{r}, \mathbf{w}, \mathbf{z}^{\text{RU}}, \mathbf{z}^{\text{RM}}, \mathbf{z}^{\text{WU}}, \mathbf{z}^{\text{WM}}, \mathbf{y}_{\setminus ij}) \\
 & \propto \frac{p(\mathbf{w}^{(\text{R})}|\mathbf{y}^{(\text{R})}, \mathbf{r})}{p(\mathbf{w}_{\setminus ij}^{(\text{R})}|\mathbf{y}_{\setminus ij}^{(\text{R})})} \cdot \frac{p(\mathbf{y})}{p(\mathbf{y}_{\setminus ij})} \\
 & = \frac{n_{ev}^{ij} + \beta^{\text{R}}}{n_e^{ij} + V\beta^{\text{R}}} \cdot \frac{n_{u2}^{ij} + \gamma}{n_u^{ij} + 3\gamma} \tag{A.14}
 \end{aligned}$$

A.6 MBI-LDA における各潜在変数分布の MAP 推定量

トピックを十分な回数更新することで得られたサンプルから、ユーザ u の評点ユーザトピック分布 θ_u^{U} 、単語ユーザトピック分布 μ_u^{U} 、アイテム m の評点アイテムトピック分布 θ_m^{M} 、単語アイテムトピック分布 μ_m^{M} 、ユーザトピック k とアイテムトピック l のペアの評点分布 $\eta_{[k,l]}$ 、単語ユーザトピック s の単語分布 ϕ_s^{U} 、単語アイテムトピック t の単語分布 ϕ_t^{M} 、評点 e の単語分布 ϕ_e^{R} 、ユーザ u のスイッチ変数分布 π_u についての MAP 推定量を得ることができる。ユーザ u から評点ユーザトピック k が生成される確率を θ_{uk}^{U} 、単語ユーザトピック s が生成される確率を μ_{us}^{U} 、アイテム m から評点アイテムトピック l が生成される確率を θ_{ml}^{M} 、単語アイテムトピック t が生成される確率を μ_{mt}^{M} 、

ユーザトピック k とアイテムトピック l のペアから評点 e が生成される確率を $\phi_{[k,l]e}$, 単語ユーザトピック s から単語 v が生成される確率を ϕ_{sv}^U , 単語アイテムトピック t から単語 v が生成される確率を ϕ_{tv}^M , 評点 e から単語 v が生成される確率を ϕ_{ev}^R , ユーザ u からスイッチ変数 a が生成される確率を π_{ua} とすると, それらは式 (A.15)~式 (A.23) でそれぞれ求めることができる.

$$\theta_{uk}^U = \frac{n_{uk} + \alpha^U}{n_u + K\alpha^U} \quad (\text{A.15})$$

$$\theta_{ml}^M = \frac{n_{ml} + \alpha^M}{n_m + L\alpha^M} \quad (\text{A.16})$$

$$\mu_{us}^U = \frac{n_{us} + \delta^U}{n_u + S\delta^U} \quad (\text{A.17})$$

$$\mu_{mt}^M = \frac{n_{mt} + \delta^M}{n_m + T\delta^M} \quad (\text{A.18})$$

$$\eta_{[k,l]e} = \frac{n_{[k,l]e} + \epsilon}{n_{[k,l]} + R\epsilon} \quad (\text{A.19})$$

$$\phi_{sv}^U = \frac{n_{sv} + \beta^U}{n_s + V\beta^U} \quad (\text{A.20})$$

$$\phi_{tv}^M = \frac{n_{tv} + \beta^M}{n_t + V\beta^M} \quad (\text{A.21})$$

$$\phi_{ev}^R = \frac{n_{ev} + \beta^R}{n_e + V\beta^R} \quad (\text{A.22})$$

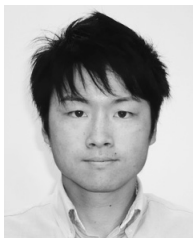
$$\pi_{ua} = \frac{n_{ua} + \gamma}{n_u + 3\gamma} \quad (\text{A.23})$$



古橋 武

1985 年名古屋大学大学院工学研究科博士後期課程電気系専攻修了. 工学博士. 2004 年名古屋大学大学院工学研究科計算理工学専攻教授, 現在に至る. ソフトコンピューティング, 感性工学に関する研究に従事. 1996 年日

本ファジイ学会論文賞受賞. IEEE, 日本知能情報ファジイ学会, 電気学会等の各会員.



田村 一樹

2012 年 3 月名古屋大学工学部電気電子・情報工学科卒業. 2014 年 3 月名古屋大学大学院工学研究科博士課程前期課程計算理工学専攻修了. 同年トヨタ自動車株式会社入社. 在学中は主として自然言語処理, 情報推薦に関する

研究に従事.



吉川 大弘 (正会員)

1997 年名古屋大学大学院博士課程修了. 同年カリフォルニア大学バークレー校ソフトコンピューティング研究所客員研究員. 1998 年三重大学工学部助手. 2005 年名古屋大学大学院工学研究科 COE 特任准教授. 2006 年

10 月同研究科准教授, 現在に至る. 主としてソフトコンピューティングとその応用に関する研究に従事. 博士 (工学). IEEE, 電子情報通信学会, 日本知能情報ファジイ学会, 進化計算学会各会員.