

人の選択行動を学習する制限ボルツマンマシン

恐神 貴行, 大塚 誠

人が選択肢集合の中から最も魅力的な選択肢を選ぶときに、各選択肢の魅力は他の選択肢に何があるかに強く影響を受ける。特に、魅力効果、妥協効果、類似効果は、人の選択に影響を与える代表的な効果として知られている。ところが、標準的な選択モデルである多項ロジットモデルは、これらの効果を表現することはできず、人の選択を予測する目的において、その予測精度が問題になることがある。これらの効果を定量的にデータから学習するために、制限ボルツマンマシンを用いた選択モデル（RBM 選択モデル）が最近提案された。本稿では、この RBM 選択モデルをわかりやすく解説する。

キーワード：選択モデル，認知バイアス，魅力効果，妥協効果，類似効果，ボルツマンマシン

1. はじめに

候補が複数あるときに、人は何を選ぶのか？ 人の選択行動をモデル化し、そのパラメタをデータから学習し、学習したモデルを予測に生かす技術は、オペレーションズ・リサーチにおいて広く応用がある。例えば、新商品のマーケティングには、価格や仕様などに応じた消費者の購買行動の予測が有用である [1]。また、都市計画や交通管理には、所要時間や運転間隔などに応じた住民の交通手段の選択の予測が求められる [2]。

ところが、人が何を選択するかは、複雑で非合理的な現象を示すことが認知心理学や行動経済学における実験によって示されており、そのような現象が人の選択行動のモデル化や予測を難しくしている [3]。人の選択行動に系統的に現れる典型的な現象には、魅力効果や妥協効果などが知られている。

魅力効果はマーケティングの実務においても活用されている。図 1 は、過去のある時点におけるエコノミスト誌の購読プランを示している [4]。オンラインでの購読が年間 59 ドル、冊子での購読が年間 125 ドル、そして冊子とオンラインの両方の購読が年間 125 ドルである。エコノミスト誌は国ごとに異なる購読プランを用意しているが、現在でも図 1 と類似した購読プランがみられる国もある。図 1 に示す購読プランの驚くべき点は、冊子だけの購読と、冊子とオンラインの両方の購読が同じ価格であるところにある。同じ価格であれば、両方の購読が明らかに得であると感じられるだ

Welcome to

The Economist Subscription Center

Pick the type of subscription you want to buy or renew.

Economist.com subscription - US \$59.00
One-year subscription to Economist.com. Includes online access to all articles from *The Economist* since 1997.

Print subscription - US \$125.00
One-year subscription to the print edition of *The Economist*.

Print & web subscription - US \$125.00
One-year subscription to the print edition of *The Economist* and online access to all articles from *The Economist* since 1997.

図 1 エコノミストの購読プラン [4]

ろう。

実際に、マサチューセッツ工科大学 (MIT) の学生 100 人に図 1 を見せて 1 つの購読プランを選んでもらったところ、16 人がオンラインでの購読を選択し、84 人が冊子とオンラインの両方の購読を選択し、冊子だけの購読を選択した学生はいなかったという Ariely の実験結果が知られている [4]。選択する人がいないのに、冊子だけの購読プランを用意しておく意味はあるのだろうか？ Ariely の実験によると、図 1 から 2 番目の選択肢（冊子だけの購読）を削除し、MIT の別の学生 100 人に 1 つの購読プランを選んでもらったところ、今度は、68 人がオンラインでの購読を選択し、32 人が冊子とオンラインの両方の購読を選択したという（表 1）。

すなわち、図 1 においては、2 番目の選択肢が「お

おそがみ たかゆき
IBM 東京基礎研究所
〒135-8511 東京都江東区豊洲 5-6-5
osogami@jp.ibm.com
おおつか まこと
フリーランス
motsuka@ucla.edu

表 1 エコノミストの購読プランの選択肢集合と各選択肢を選んだ人の数 [4]

選択肢集合	選んだ人の数		
	オンライン	冊子	両方
オンライン, 冊子, 両方	16	0	84
オンライン, 両方	68		32

表 2 各 PC の仕様と選択肢集合および各選択肢を選んだ人の数 [5]

	仕様				
	A	B	C	D	E
CPU [MHz]	250	300	350	400	450
主記憶容量 [MB]	192	160	128	96	64
選択肢集合	選んだ人の数				
	A	B	C	D	E
A, B, C	36	176	144		
B, C, D		56	177	115	
C, D, E			94	181	109

とり」の役割を果たして、この「おとり」が3番目の選択肢の相対的な魅力を大きくしていると考えられることができる。この「おとり」が人の選択に与える影響が魅力効果として知られている。

妥協効果もマーケティングの実務において利用されている現象である。Kivetzらは、空港のロビーにおいて、表2に示す5つのパソコン(PC)の中から3つを旅行者に提示して、最も好ましいPCを選んでもらうという実験を行った[5]。プロセッサ(CPU)は、Aが最も低速で、Eが最も高速である。主記憶容量は、Aが最も大きく、Eが最も小さい。選択肢としてAとBとCを提示したときには、大多数の176人がBを選択したが、BとCとDが選択肢集合であるときには、大多数の177人がCを選択した。すなわち、BとCのどちらが好ましいかは、他の選択肢にAがあるかDがあるかに依存しており、BとCの嗜好関係が他の選択肢によって逆転することが示されている。特に、BやCが他の選択肢と比べて「極端でない」場合において、選択される確率が大きくなっている。CとDとEが選択肢集合である場合にも、「極端でない」Dが最もよく選ばれている。このように、人が極端な選択肢を避ける現象は、妥協効果として知られている。

魅力効果や妥協効果は人の選択行動に広く系統的に見られる現象であり[4]、適切な価格付けや需要予測のためには、データに基づいて定量的にこれらの効果の評価や予測を行う必要がある。ところが、これまでは魅力効果や妥協効果が定量的な理解に基づいて活用されることはほとんどなかった。

この定量的な理解や活用の欠如は、魅力効果や妥協効果が現れる人の選択行動をデータから学習できる選択モデルの欠如に一因があると言える。例えば、広く用いられている選択モデルのクラスは、ランダム効用最大化の原理に基づいているが、これらの選択モデルでは魅力効果を表現できないことが知られている[6]。また、魅力効果や妥協効果を表現できるように、人の認知プロセスを模倣した選択モデルも提案されてきているが[7]、そのパラメタをデータから学習する手法は考えられていない。

魅力効果や妥協効果に加えて、人の選択に現れるといわれている現象に類似効果がある。Rieskampらはこれらの3つの効果を人の選択行動によく見られる重要な現象としてあげている[6]。類似効果は「赤バス青バス」として知られる例に代表される現象である[8]。交通手段に車と青いバスの2つがある場合に、それぞれ半分のシェアがあるとする。選択肢に赤いバスが加わった場合には、赤いバスは青いバスのシェアを奪い、車と青バスと赤バスのシェアは、(1:1:1ではなく)2:1:1になるのが「赤バス青バス」の例である。魅力効果や妥協効果が認知バイアスによる非合理的な選択を引き起こすのに対して、類似効果は合理的な選択にも現れる現象である。

本稿では、制限ボルツマンマシン(restricted Boltzmann machine; RBM)を用いたモデル(RBM選択モデル)によって、魅力効果や妥協効果や類似効果が見られる人の選択行動をデータから学習する新しい試み[9]を紹介する。RBM選択モデルは、標準的な選択モデルである多項ロジットモデル[2]を拡張したモデルとなっている。多項ロジットモデルは、魅力効果、妥協効果、類似効果のいずれも表現できないことが知られている。本稿では、多項ロジットモデルと対比させることで、RBM選択モデルの概要を紹介する。興味を持たれた読者は、[9]にある詳細を参照されたい。

2. ロジットモデルの問題点

複数の候補の中から一つを選ぶという選択行動をモデル化するとき、最も広く使われているモデルに多項ロジットモデルがある。多項ロジットモデルはLuceらにより50年代に提唱され、McFaddenにより70年代に完成されたモデルであり、現在では、マーケティングをはじめ様々な分野で広く用いられている[2]。

多項ロジットモデルでは、選択肢集合 \mathcal{A} に含まれる各選択肢が独立に効用(魅力を数値化したもの)を持ち、選択確率はこれらの効用により決まると考える。例

例えば、選択肢 A や X の効用を b_A や b_X と書くと、多項ロジットモデルでの「選択肢 A の選ばれやすさ」は

$$\lambda(A|\mathcal{X}) = \exp(b_A) \quad (1)$$

のように効用の単調増加関数で表される。そして、選択肢集合 \mathcal{X} から選択肢 A が選ばれる確率は、この「選択肢の選ばれやすさ」を用いて

$$P(A|\mathcal{X}) = \frac{\lambda(A|\mathcal{X})}{\sum_{X \in \mathcal{X}} \lambda(X|\mathcal{X})} \quad (2)$$

と書き表される。式 (2) の右辺は各選択肢の効用を引数として取るソフトマックス関数であり、大きな効用を持つものが高い確率で選ばれ、選択肢集合に含まれる各選択肢が選ばれる確率を足し合わせると 1 になることを表す。これは魅力的な（高い効用を持つ）選択肢は選ばれる確率が高いという、一見自明な仮定をモデルに反映したものである。通常、選択肢 X はいくつの特徴量（例えば価格や品質など）で記述され、この効用 b_X は特徴量の線形和で表される。

現在も多項ロジットモデルが広く用いられている理由に、シンプルさゆえのモデルの扱いやすさがある。実際、人の選択行動データを用いてパラメータを学習することができ、学習されたパラメータの意味を解釈することも容易である。しかし、多項ロジットモデルはそのシンプルさゆえに表現力が限られており、実際の人の選択行動に見られる現象（魅力効果、妥協効果、類似効果など）を表現できない。

例えば、魅力効果の実例として挙げたエコノミストの購読プランの選択行動（表 1）を多項ロジットモデルでモデル化しようとするときのような矛盾が生じる。この例では、選択肢に「おとり」が含まれるかどうかにより、最も人気のある選択肢が変化した。「オンライン」と「両方」の 2 つの購読方法だけが選択肢集合に含まれる場合、「オンライン」を選択する人が 68 人で「両方」は 32 人なので、「オンライン」の効用が「両方」の効用よりも高いはずである。しかしながら、選択肢集合に「冊子」という「おとり」を加えると、「両方」を選択する人は 84 人で、「オンライン」の 16 人を上回っており、今度は「両方」の効用のほうが「オンライン」の効用よりも高くなるなくてはならない。上の両方の条件を満たすような効用は存在しないことからわかるように、多項ロジットモデルは魅力効果が反映された選択行動を表現できない。

妥協効果の実例として挙げた Kivetz らの PC 選択実験の結果（表 2）を多項ロジットモデルでモデル化し

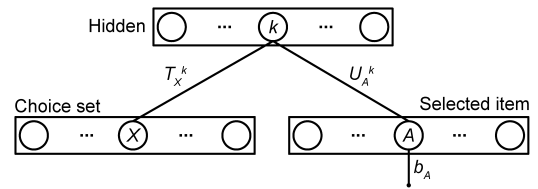


図 2 RBM 選択モデルの構造 [9]

ようになると同じような矛盾が起こる。この場合、選択肢集合が (A, B, C), (B, C, D), (C, D, E) のどれでも、真ん中の選択肢が最も多く選ばれるという結果になった。これは選択肢集合が (A, B, C) のときには B が C よりも高い効用を持つが、選択肢集合が (B, C, D) となると C が B よりも高い効用を持つことになり、多項ロジットモデルでは表現できない。

類似効果の例としてあげた「赤バス青バス」を多項ロジットモデルで表現しようすると、やはり以下の矛盾が起こる。車と青バスの 2 つだけが選択肢にある場合には、それぞれ半分のシェアがあるので、車の効用と青バスの効用は等しいと考えられる。ここに赤バスを導入してどのように赤バスの効用を決めても、車と青バスの効用が等しいため、車と青バスの選択確率は等しくなり、車と青バスと赤バスのシェアを 2:1:1 にすることはできない。

それでは、魅力効果、妥協効果、類似効果などが現れる、人の行動選択をモデル化するにはどうしたらよいただろうか？ その答えが RBM 選択モデルである。

3. RBM 選択モデル

RBM 選択モデルは、多項ロジットモデルの拡張として提案されたモデルであり、認知バイアスの乗った選択行動を表すのに十分な表現力と、行動データを使って学習できる扱いやすさを兼ね備えた選択モデルである。RBM 選択モデルは制限ボルツマンマシン (RBM) という確率モデルとして表現されており、RBM の持つ好ましい性質や学習則をそのまま受け継いでいるという利点を持つ。RBM 選択モデルが生まれた背景には、多項ロジットモデルをエネルギーベースなモデル（無向グラフ、マルコフ確率場）として捉え直したところにその発想の出発点がある。

それでは RBM 選択モデルの構造を見てみよう。RBM 選択モデルはバイナリノード (0 か 1 を出す確率変数) で構成される 3 つの層（出力層、入力層、隠れ層）とウエイトとバイアスで構成されるパラメータからなる（図 2 参照）。1 つ目の層は出力層で、この

層は選ばれた選択肢 (Selected item) を表現する。全ての可能な選択肢に対応するノードを持ち、選択肢集合の中で選択肢 X が選択された場合には、 X に対応するノードのみが 1 の値を取り、それ以外のノードは 0 となる。2 目目の層は入力層で、この層は選択肢集合 (Choice set) を表現する。この層も出力層と同じように全ての可能な選択肢に対応するノードを持つが、1 の値を取るのには選択肢集合に含まれる全ての選択肢で、それ以外のノードは 0 となる。例えば、Kivetz らの PC 選択実験 (表 2) の場合、全ての可能な選択肢は A, B, C, D, E なので、出力層と入力層にそれぞれ 5 つのノードを用意する。この場合「選択肢集合として (B, C, D) が与えられたときに D を選択した」という事例は、入力層を (0, 1, 1, 1, 0), 出力層を (0, 0, 0, 1, 0) とすることで表現できる。3 目目の層は隠れ (Hidden) 層で、この層はモデルの表現力を高める働きをする。多項ロジックモデルでは「選択肢集合が各選択肢の選ばれやすさに影響を与えない」という仮定のためにその表現力が限られていたが、RBM 選択モデルでは「選択肢集合が各選択肢の選ばれやすさに影響を与える」ことを表現できるように入力層と隠れ層を含んでいる。

学習の対象となるパラメータにはバイアスとウエイトの 2 種類がある。バイアスは各ノードが 1 となる確率を調整するものであり、RBM 選択モデルでは出力層のノードだけが持つ (選択肢集合から 2 つ以上を選択する場合は、隠れ層にもバイアスを入れる必要がある。詳細は [10] を参照されたい)。ウエイトは 2 ノードの間の統計的な関係性を決める。RBM 選択モデルでは、入力層と隠れ層の間 (T_{Xk}^k は選択肢 X と k 番目の隠れノードをつなぐウエイト)、また出力層と隠れ層の間 (U_A^k は選択肢 A と k 番目の隠れノードをつなぐウエイト) にウエイトが定義される。実はこのウエイトが RBM 選択モデルにおいて大きな役割を果たすのでもう少し詳しく見てみよう。

ノードが 0 か 1 の値を確率的に取るので、2 ノード間のウエイトが正のときには 2 つのノードが両方とも 1 になる確率が高くなり、ウエイトが負のときには 2 つのノードが両方とも 1 になる確率が低くなる。また、2 ノード間のウエイトが 0 のとき、もしくは 2 ノード間にウエイトが存在しないときは、2 ノード間に直接的な統計的な関係性はなくなる。ここで面白いのは、2 ノード間にはウエイトによる直接的なつながりがないのに、隠れノードを介して 2 つのノードがつながっている場合である。この場合、2 つのノードは統計的な

関係性を持つ。そして、この統計的な関係性の複雑さは、間に挟まる隠れノードの数によって決まる。RBM 選択モデルでは入力層と出力層の間に隠れ層が存在するため、入力層の状態が出力層の出方に影響を与えるのである。

RBM 選択モデルにおいて、選択肢集合 \mathcal{X} から A を選択する確率は式 (2) と同様の形で表すことができる。変更される点は「選択肢 X の選ばれやすさ」の定義のみで、RBM 選択モデルでは、隠れノードの集合を \mathcal{K} として、

$$\begin{aligned} \lambda(X|\mathcal{X}) &= \exp(b_X) \prod_{k \in \mathcal{K}} (1 + \exp(\sum_{Y \in \mathcal{X}} T_Y^k + U_X^k)) \\ &= \exp(-F(X, \mathcal{X})) \end{aligned} \quad (3)$$

と定義される。この式をよく見ると、出力層のバイアス b_X が選択確率のベースラインを決め、選択肢集合 \mathcal{X} が隠れノードを介して最終的な X の選択確率に影響を与えることが見て取れる。また、 $F(X, \mathcal{X})$ は平衡自由エネルギーと呼ばれる量であり、RBM 選択モデルにおいてはこの量が大きくなるほど選択される確率が小さくなるので、平衡自由エネルギーは負の効用を表現していると言える。この一見シンプルな RBM 選択モデルが、任意の強さの魅力効果・妥協効果・類似効果を表現できることは理論的に証明されている [9]。

RBM 選択モデルは、その名が示すとおり RBM である。RBM は、同じ層に属するノード間にはウエイトが存在しない 2 層のボルツマンマシン (BM) であり、全結合型の BM に比べて、推定も学習も格段に簡単である [11, 12]。RBM 選択モデルは、可視層 (外側からデータが入られる層で、ここでは入力層と出力層を合わせたもの) と隠れ層の 2 層からなる。RBM は、片方の層にデータが与えられた際に他方の層内のノードが条件付き独立となり、事後分布を厳密かつ高速に求めることができるという性質を持つ。この性質により、RBM は、BM に起因する表現力を持ち、推定・学習ができるという扱いやすさを兼ね備えたモデルとなる。また、通常 RBM はコントラストティブ・ダイバージェンス (CD) 法などの生成学習則によって学習されるが、RBM 選択モデルでは可視層が入力層と出力層の 2 層からなるため、Larochelle と Bengio によって提唱された判別学習則を用いることができる [13]。これにより CD 法などを用いた近似手法とは異なり、学習の際に対数尤度勾配を厳密かつ高速に計算することができる。

RBM 選択モデルから入力層と隠れ層とウエイトを取り除き、出力層とバイアスのみを残すと、多項ロジック

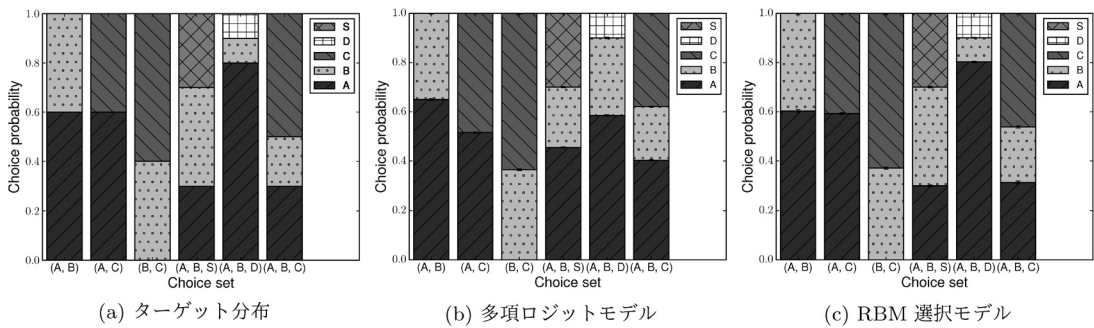


図3 ターゲット分布 (a) に基づいて学習された多項ロジットモデル (b) と RBM 選択モデル (c) の予測分布 [10]

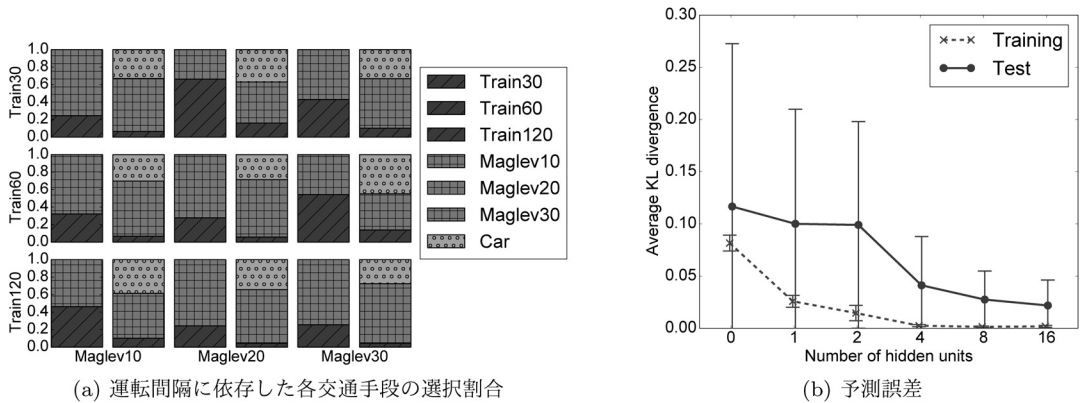


図4 スイスメトロのデータの概要 (a) と隠れ層のノード数を変化させた RBM 選択モデルによる予測誤差 (b) [9]

トモデルとなる。この場合、出力層のバイアスは多項ロジットモデルの効用を表す。すなわち、出力層のみのエネルギーベースなモデルにおける負のエネルギー（隠れ層がないので平衡自由エネルギーとは呼ばない）は多項ロジットモデルにおける効用である。次に RBM 選択モデルと多項ロジットモデルの2つの対比実験をみてみよう。

1つ目の数値実験では、ターゲットとする分布に基づいて生成されたサンプルを使って、ターゲットの分布をモデルが学習できるかどうかを確かめた。ターゲットとする分布は、3つの効果を表すように (A, B), (A, C), (B, C), (A, B, S), (A, B, D), (A, B, C) という6つの選択肢集合に対して、適当に定めた (図3 (a))。まず、A と B が選択肢集合を構成するそれぞれの選択確率を $p_{A,B} = (0.6, 0.4)$ とした。次に、類似効果を表すために、選択肢 S を加えた際の選択確率を $p_{A,B,S} = (0.3, 0.4, 0.3)$ とした。これは、A に似た S という選択肢が A のシェアを奪うということを表す。同様に、魅力効果を表すために、 $p_{A,B,D} = (0.8, 0.1, 0.1)$ とした。これは、A に対する「おとり」となる D が A の魅力を大きくすることを表す。さらに、妥協効果

を表すために、 $p_{A,C} = (0.6, 0.4)$, $p_{B,C} = (0.4, 0.6)$, $p_{A,B,C} = (0.3, 0.2, 0.5)$ とした。これは、極端な選択肢の A と B が、C の魅力を上げることが表す。

多項ロジットモデルと、隠れノードを1つ持つ RBM 選択モデルが学習した分布を図3 (b)-(c) に示す。ここではモデルが3つの効果を表現できるかどうかを確かめるのが目的であるため、訓練データを用いて十分なトレーニングをし、モデルを訓練データに対して意図的にオーバーフィッティングさせた。図3 (b) を見ると、多項ロジットモデルが学習した分布はターゲットの分布とかけ離れていることが見て取れる。例えば、選択肢集合 (A, B) と選択肢集合 (A, B, D) に対して学習された分布には、魅力効果が示唆する選択確率の逆転は起こっておらず、選択肢集合が変化しても選択確率の比率は一定のままに保たれている。一方、RBM 選択モデルが学習した分布はターゲットの分布を忠実に再現している。図3 (c) は隠れノードが1つの場合であるが、隠れノードの数を増加させると、さらにターゲット分布に近づけることができる [10]。この実験により、十分な数の隠れノードを持つ RBM 選択モデルは、3つの効果が現れる人の選択の確率分布を、その人

の選択結果から学習できることがわかった。次に、スイスメトロの実データを用いて、RBM 選択モデルが学習の結果を汎化できることをみてみよう。

スイスメトロとはスイスで計画された長距離高速リニアモーターカー (Maglev) のことで、スイスメトロの実データセットはその計画の際に行った調査の結果である [14]。調査の対象者は、提示された複数の選択肢集合の中から、実際に利用すると思われる交通手段を選んだ。車を所有していない対象者の場合は、電車 (Train) と Maglev が選択肢となる。車を所有している対象者の場合は、これらに車 (Car) を加えた 3 つの交通手段が選択肢集合を構成する。電車と Maglev は、料金や所要時間など様々な条件も変えて提示されたが、ここではそれらの運転間隔だけを考慮する。すなわち、電車は 30 分間隔 (Train30)、60 分間隔 (Train60)、120 分間隔 (Train120) の 3 種類の中から 1 つを提示し、Maglev は 10 分間隔 (Maglev10)、20 分間隔 (Maglev20)、30 分間隔 (Maglev30) の 3 種類の中から 1 つを提示したとする。よって、可能な選択肢には Train30、Train60、Train120、Maglev10、Maglev20、Maglev30、Car の 7 種類があることになる。

このスイスメトロのデータセットには、魅力効果と似た現象が現れている (図 4 (a))。例えば、(Train60, Maglev30) の組が選択肢集合である場合には、Maglev30 よりも Train60 を選択する確率が高くなっているのに、(Train60, Maglev30, Car) が選択肢集合であるときには、Train60 の選択確率が小さくなり、残ったシェアを Marlev30 と Car が取り合っている。

18 個の可能な選択肢集合のうち、17 個の選択肢集合を訓練データ、残りの 1 個をテストデータとして、学習された RBM 選択モデルの予測精度を評価した。この結果、隠れノードの数を増やすと、訓練データだけでなく、テストデータに対してもエラー (ターゲット分布と学習済み分布との間のカルバック・ライブラー擬距離) が減ることが示された (図 4 (b))。テストデータに現れる選択肢集合は、訓練データには一度も現れておらず、本実験により、RBM 選択モデルが汎化性能を持つことが示された。

4. おわりに

本稿では、[9] で提案された RBM 選択モデルをわかりやすく紹介した。特に、RBM 選択モデルの動機となった魅力効果、妥協効果、類似効果の 3 つの効果について、[9] にはない具体例を交えて丁寧に解説した。また、RBM 選択モデルについても、[9] とは異なる視

点の解説を交えた。本稿を [9] と合わせて読むことで、RBM 選択モデルとその周辺技術についてより深い理解が得られることを期待する。

謝辞 本稿の一部は独立行政法人科学技術振興機構 (JST)、CREST の助成を受けて執筆されました。

参考文献

- [1] P. E. Green, A. M. Krieger and Y. Wind, “Thirty years of conjoint analysis: Reflections and prospects,” *Interfaces*, **31**(3-supplement), S56–S73, 2001.
- [2] 三浦英俊, “ロジットモデルを用いたリニア中央新幹線の需要予測,” *オペレーションズ・リサーチ*, **54**, pp. 419–428, 2009.
- [3] D. McFadden, “Economic choices,” *The American Economic Review*, **91**, pp. 351–378, 2001.
- [4] D. Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*, revised and expanded edition, Harper Perennial, 2010.
- [5] R. Kivetz, O. Netzer and V. Srinivasan, “Alternative models for capturing the compromise effect,” *Journal of Marketing Research*, **41**, pp. 237–257, 2004.
- [6] J. Rieskamp, J. R. Busemeyer and B. A. Mellers, “Extending the bounds of rationality: Evidence and theories of preferential choice,” *Journal of Economic Literature*, **44**, pp. 631–661, 2006.
- [7] T. Otter, J. Johnson, J. Rieskamp, G. M. Allenby, J. D. Brazell, A. Diederich, J. W. Hutchinson, S. MacEachern, S. Ruan and J. Townsend, “Sequential sampling models of choice: Some recent advances,” *Marketing Letters*, **19**, pp. 255–267, 2008.
- [8] K. Train, *Discrete Choice Methods with Simulation*, 2nd ed., Cambridge University Press, 2009.
- [9] T. Osogami and M. Otsuka, “Restricted Boltzmann machines modeling human choice,” In *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence and K. Weinberger (eds.), pp. 73–81, 2014.
- [10] T. Osogami and M. Otsuka, Supplementary material for “Restricted Boltzmann machines modeling human choice,” <http://papers.nips.cc/paper/5280-restricted-boltzmann-machines-modeling-human-choice> (2014 年 12 月 8 日閲覧)
- [11] P. Smolensky, “Information processing in dynamical systems: Foundations of harmony theory,” In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1*, D. E. Rumelhart, J. L. McClelland and PDP Research Group (eds.), pp. 194–281, MIT Press, 1986.
- [12] G. E. Hinton, “Training products of experts by minimizing contrastive divergence,” *Neural Computation*, **14**, pp. 1771–1800, 2002.
- [13] H. Larochelle and Y. Bengio, “Classification using discriminative restricted Boltzmann machines,” In *Proceedings of the 25th International Conference on Machine Learning*, pp. 536–543, 2008.
- [14] M. Bierlaire, K. Axhausen and G. Abay, “The acceptance of modal innovation: The case of Swissmetro,” In *Proceedings of the First Swiss Transportation Research Conference*, 2001.