RF-003

混合整数最適化による階層的変数選択を用いた店舗選択要因の分析

佐藤 俊樹† 高野 祐一‡ 中原 孝信♮

† 筑波大学大学院システム情報工学研究科
‡ 専修大学ネットワーク情報学部 ♮ 専修大学商学部

1 はじめに

回帰分析において，候補となる変数の中から有用な説明変数集合を選び出すことは変数選択と呼ばれる（他にも特徴選択，素性選択，部分集合選択などと呼ばれることもある）．変数選択には分析結果の解釈が容易になることに加えて，手元のデータへの過剰適合を抑制し予測性能が向上するなどの利点がある [10]．近年では扱うデータ量の増加を背景として，機械学習やデータマイニングなどの分野で変数選択手法が盛んに研究されている [6, 10, 16, 19]．

素朴な変数選択の方法としては，候補変数のすべての部分集合を調べる総当り法 [9] が考えられるが，候補変数が多い場合には現実的な方法とは言えない．より実用的な方法として，ステップワイズ法 [8]，罰則付き回帰 [27]，メタ戦略 [30] などがあるが，これらは発見的解法であり，適合度指標に関して最良の変数集合を選択できる保証は無い．一方で，近年では計算機とアルゴリズムの性能向上を背景として，混合整数最適化を用いた変数選択手法が注目を集めている [2, 3, 17, 20, 21, 24, 25]．

本研究では混合整数最適化による変数選択手法を用いて，消費者の店舗選択要因を分析する．消費者の店舗選択行動に関する研究は数多く存在し [1, 5, 7, 18, 22, 23]，特に Sato et al. [26] では本研究と同様に混合整数最適化による変数選択手法が利用されている．Sato et al. [26] は消費者の複数店舗における購買状況を把握できるスキャンパネルデータを対象に，購入商品の商品分類によって選択店舗を説明するモデルを提案しており，各店舗を選択する際に想起される商品群を明らかにできる点に特徴がある．一方で，各店舗の売上と強く関連する購買点数については分析しておらず，変数選択の際には大分類・中分類・小分類などの商品分類間の階層構造も考慮していない．

説明変数間の構造情報のような，データに内在する事前知識をモデル構築に活用する方法は構造正則化と呼ばれる [14]．例として，関連する変数をまとめて選択／削除するグループ変数選択や，変数間の優先関係に従って変数選択を行なう階層的変数選択などがある [4, 11, 12, 13, 15, 28, 29, 31]．これらの研究ではステップワイズ法や罰則付き回帰に基づく解法が提案されているが，我々の知る限り，混合整数最適化による階層的変数選択の研究は存在しない．

以上を考慮して，本研究の目的は以下の 3 点とする：

- 対象店舗における購買点数に着目した店舗選択要因の分析モデルを提案する．
- 変数間の階層構造を考慮した混合整数最適化による変数選択手法を提案する．
- 実際のスキャンパネルデータを用いて提案手法の有効性を検証し，顧客の店舗選択要因を分析する．

本論文の構成は以下のようになる．2 節では店舗選択要因の分析モデルを説明する．3 節では混合整数最適化による変数選択モデルを説明する．4 節で数値実験の結果を報告し，5 節で本論文のまとめを述べる．

2 店舗選択モデル

店舗選択の分析対象とする 2 店舗を「店舗 A（正例）」「店舗 B（負例）」とし，それらの店舗への顧客の来店をサンプル $i = 1, 2, \ldots, n$ とする．Sato et al. [26] では，顧客が選択する店舗を予測する 2 値分類問題として店舗選択モデルを作成している．一方で本研究では，各店舗における購買点数を考慮して分析を行なうために，来店 $i = 1, 2, \ldots, n$ に対して以下のように被説明

Analysis of store choice factors using hierarchical variable selection based on mixed integer optimization
Toshiki SATO†, Yuichi TAKANO‡, and Takanobu NAKAHARA♮
† Graduate School of Systems and Information Engineering, University of Tsukuba
‡ School of Network and Information, Senshu University
♮ School of Commerce, Senshu University

変数 y_i を定義する：

$$y_i = \text{（店舗 A における商品購入個数）} - \text{（店舗 B における商品購入個数）}.$$

したがって被説明変数 y_i は，来店 i で店舗 A が選択された場合には正の値，店舗 B が選択された場合には負の値となる．また，店舗 A と店舗 B が同時に選択されることはない．このように被説明変数を定義することで，店舗の選択要因に加えてそれらが顧客の購買点数に与える影響も分析することができる．

店舗選択の要因として，本研究では顧客属性（性別や年齢など）と購入商品属性（商品分類）を考える．まず顧客属性を表す説明変数の集合を G_0 とし，来店 $i = 1, 2, \ldots, n$ に該当する顧客の属性 $j \in G_0$ の値を x_{ij} とする．また商品の大分類の集合を G_1，中分類の集合を G_2，小分類の集合を G_3 とする．説明変数 $j \in G_1 \cup G_2 \cup G_3$ は，来店 $i = 1, 2, \ldots, n$ において分類 j の商品が購入された場合は $x_{ij} = 1$，購入されなかった場合は $x_{ij} = 0$ となるダミー変数とする．以降ではこれらの説明変数をまとめて $G = \bigcup_{k=0}^{3} G_k$ と表記する．

3 混合整数最適化による変数選択

本節では混合整数最適化による変数選択の基本モデルを紹介した後，本研究で提案する階層的変数選択モデルを提示する．

3.1 基本モデル

切片を表す決定変数を $b \in \mathbb{R}$ とし，偏回帰係数を表す決定変数のベクトルを $\boldsymbol{a} = (a_j;\ j \in G) \in \mathbb{R}^p$ とする．また，$\boldsymbol{z} = (z_j;\ j \in G) \in \{0,1\}^p$ は 0-1 決定変数のベクトルとし，$z_j = 1$ の場合には説明変数 j を選択することを表す．ここで $z_j = 0$ の場合には，対応する偏回帰係数を $a_j = 0$ とすることで説明変数 j を回帰式から削除する．選択変数の上限数を表すパラメータを s とし，事前に値を指定しておく．このとき，残差 2 乗和が最小となるように s 個以下の説明変数を選択する問題は以下のように定式化できる [20, 21]：

$$\text{最小化}_{\boldsymbol{a},\, b,\, \boldsymbol{z}} \quad \sum_{i=1}^{n} \left(y_i - \left(b + \sum_{j \in G} a_j x_{ij} \right) \right)^2 \tag{1}$$

$$\text{制約条件} \quad z_j = 0 \ \Rightarrow\ a_j = 0 \quad (j \in G), \tag{2}$$

$$\sum_{j \in G} z_j \le s, \tag{3}$$

$$z_j \in \{0,1\} \quad (j \in G). \tag{4}$$

制約条件 (2) は，タイプ 1 の特殊順序集合（SOS1）制約による表現が可能である．SOS1 制約は，集合の要素のうち非ゼロの値をとるのは高々一つであることを意味する．よって，集合 $\{1 - z_j, a_j\}$ に SOS1 制約を課すことで制約条件 (2) を表現できる．SOS1 制約は標準的な最適化ソルバーに実装されている．

3.2 階層的変数選択モデル

商品分類間には階層構造（包含関係）が存在する．例えば小分類「調味料」は中分類「加工食品」に含まれ，中分類「加工食品」は大分類「食品」に含まれる．このような包含関係が成り立つ商品分類の組 $(j_1, j_2, j_3) \in G_1 \times G_2 \times G_3$ の集合を H とする．

本研究では商品分類の組 $(j_1, j_2, j_3) \in H$ に対して，小分類 j_3 が説明変数として選択される場合にはその上位の中分類 j_2 も必ず選択し，中分類 j_2 が説明変数として選択される場合にはその上位の大分類 j_1 も必ず選択するという以下の制約条件を課した変数選択モデルを提案する：

$$z_{j_1} \ge z_{j_2} \ge z_{j_3} \quad ((j_1, j_2, j_3) \in H).$$

このような階層的変数選択には以下のような利点がある：

- ある商品分類が店舗選択の要因となる場合には，その上位の商品分類も店舗選択に影響を与える可能性が高く，選択変数の信頼性が高まる．
- 上位の分類が説明変数として優先的に選択され，上位の分類ほど購入データが多いために偏回帰係数推定の信頼性が高まる．
- 探索すべき選択変数の組合せが制約条件により削減され，計算効率が向上する．

本研究で提案する階層的変数選択モデルは以下のように定式化できる：

$$\text{最小化}_{\boldsymbol{a},\, b,\, \boldsymbol{z}} \quad \sum_{i=1}^{n} \left(y_i - \left(b + \sum_{j \in G} a_j x_{ij} \right) \right)^2 \tag{5}$$

$$\text{制約条件} \quad z_j = 0 \ \Rightarrow\ a_j = 0 \quad (j \in G), \tag{6}$$

$$\sum_{j \in G} z_j \le s, \tag{7}$$

$$z_{j_1} \ge z_{j_2} \ge z_{j_3} \quad ((j_1, j_2, j_3) \in H), \tag{8}$$

$$z_j \in \{0,1\} \quad (j \in G). \tag{9}$$

4 数値実験

本節では数値実験により，本研究で提案する階層的変数選択モデルの有効性を検証し，顧客の店舗選択要

因を分析する．

4.1 分析データ

株式会社マクロミルより提供されたホームスキャン方式のスキャンパネルデータを利用した．データ期間は 2012 年 1 月からの 2 年間であり，業界動向サーチ（http://gyokai-search.com/）に基づき，コンビニエンスストア（以下，コンビニ）・ドラッグストア・スーパーマーケット（以下，スーパー）の各業態の上位 5 企業が展開する東京都内の店舗を分析対象とした．表 1 に対象企業の一覧を示す．

顧客の 1 回の来店を 1 サンプルとし，来店した店舗の業態に応じて正例・負例を定義し，表 2 の 3 種類のデータセットを作成した．ここで n はサンプル数を表す．

顧客属性に関する候補変数として，事前のアンケート調査に基づく表 3 の 37 種類のダミー変数を用いた．購入商品属性に関する候補変数は，JICFS 分類コードの大分類・中分類・小分類に対応するダミー変数とした．表 4 では大分類と中分類の一覧を示しており，さらに小分類は 214 種類ある．各データセットにおいて，すべてのサンプルに対して値がゼロとなる変数は事前に削除した．

以降では，以下の候補変数集合を用いて分析を行なう：

小分類なし集合 顧客属性と商品大分類・中分類からなる候補変数集合．

小分類あり集合 顧客属性と商品大分類・中分類・小分類からなる候補変数集合．

4.2 予測精度の検証

本節では 5 分割交差確認を実施して，以下の 4 種類の手法の予測精度を検証する：

SW$_s$ 説明変数無しのモデルから開始するステップワイズ法．

L1$_s$ L_1 罰則付き回帰 [27] による変数選択．

BM$_s$ 混合整数最適化による基本モデル（問題 (1)–(4)）．

HM$_s$ 混合整数最適化による階層的変数選択モデル（問題 (5)–(9)）．

ただし，s は選択変数の上限数を表す．小分類なし集合に対しては $s = 10,\ 20$，小分類あり集合に対しては $s = 20,\ 50$ と設定した．

ステップワイズ法と L_1 罰則付き回帰はそれぞれ統計解析ソフト R（http://www.r-project.org/）の step 関数と glmnet 関数を利用した．L_1 罰則付き回帰では，罰則項の重みパラメータを区間 $[0, 1]$ で 0.0001 刻みで動かし，偏回帰係数ベクトルの非ゼロ成分が s 個となった場合の非ゼロ成分に対応する変数を選択し，最小 2 乗法により偏回帰係数を推定した．基本モデルと階層的変数選択モデルの求解には最適化ソルバー Gurobi Optimizer（http://www.gurobi.com）を使用し，実用性を考慮して計算時間の上限は 1000 秒とした．

表 1: 対象企業一覧

業態	企業
コンビニ	セブンイレブン，ローソン，ファミリーマート，ミニストップ，スリーエフ
ドラッグストア	マツモトキヨシ，サンドラッグ，ツルハホールディングス，コスモス薬品，スギホールディングス
スーパー	イオン，セブン&アイ，ユニーグループ，ダイエー，イズミ

表 2: データセット

略称	正例	負例	n
CvsD	コンビニ	ドラッグストア	225,639
CvsS	コンビニ	スーパー	252,513
DvsS	ドラッグストア	スーパー	139,188

表 3: 顧客属性の候補変数一覧

	候補変数（37 種類）
性別など	女性，既婚，子持ち
年代	20 代，30 代，40 代，50 代，60 代以上
収入階級	2 以下，3〜4，5〜6，7〜9，10 以上
交差項	{ 男性, 女性 } × { 既婚, 子持ち }，{ 男性, 女性 } × 年代，{ 男性, 女性 } × 収入階級

5 分割交差確認の結果を表 5，6 に示す．p は候補変数の数を表し，時間（秒）は変数選択に要した平均計算時間を表す．また，R^2 と RMSE はそれぞれ検証データにおける決定係数と平方平均 2 乗誤差の平均値であり，各手法の予測精度を表す．これらの予測精度指標に関して 4 種類の手法の最良値は太字で記載した．

表 5 の小分類なし集合の結果では，ステップワイズ法と階層的変数選択モデルの予測精度が高く，平均では僅かに階層的変数選択モデルが上回っていた．また，L_1 罰則付き回帰の予測精度は最も低かった．他の手法とは異なり，階層的変数選択モデルでは商品分類間の階層構造を考慮して変数選択を行なっており，このことによって高い予測精度が得られたと考えられる．

表 4: JICFS 分類コードの大分類と中分類

大分類（5 種類）	中分類（29 種類）
食品	加工食品，生鮮食品，菓子類，飲料・酒類，その他食品
日用品	日用雑貨，OTC 医薬品，化粧品，家庭用品，DIY 用品，ペット用品，その他日用品
文化用品	文具・事務用品・情報文具，玩具，書籍，楽器・音響ソフト，情報機器，その他文化用品
耐久消費財	家具，車両用品，時計・メガネ，光学・写真関連品，家電，その他耐久消費財
衣料・身の回り品・スポーツ用品	衣料・衣服，寝具・装飾品，身の回り品，靴・履物，スポーツ用品

表 5: 小分類なし集合に対する 5 分割交差確認の結果

	p	手法	時間（秒）	R^2	RMSE
CvsD	69	SW_{10}	106.4	**0.3367**	**2.7090**
		$L1_{10}$	10.0	0.3360	2.7104
		BM_{10}	43.3	**0.3367**	**2.7090**
		HM_{10}	14.9	0.3366	2.7091
		SW_{20}	311.1	**0.3384**	**2.7055**
		$L1_{20}$	6.8	0.3375	2.7073
		BM_{20}	1000.0	0.3381	2.7060
		HM_{20}	1000.0	0.3382	2.7059
CvsS	71	SW_{10}	121.9	0.2105	5.2533
		$L1_{10}$	24.7	0.2090	5.2583
		BM_{10}	43.7	0.2105	5.2533
		HM_{10}	14.4	**0.2106**	**5.2530**
		SW_{20}	360.5	0.2140	5.2416
		$L1_{20}$	14.5	0.2138	5.2426
		BM_{20}	1000.0	0.2138	5.2423
		HM_{20}	1000.0	**0.2143**	**5.2406**
DvsS	71	SW_{10}	66.4	**0.2098**	**6.8622**
		$L1_{10}$	28.6	0.2089	6.8663
		BM_{10}	18.0	0.2097	6.8626
		HM_{10}	3.9	0.2097	6.8628
		SW_{20}	195.8	0.2113	6.8555
		$L1_{20}$	13.1	0.2110	6.8570
		BM_{20}	1000.0	0.2114	6.8552
		HM_{20}	1000.0	**0.2116**	**6.8545**
平均		SW	193.7	0.2535	4.9379
		L1	16.3	0.2527	4.9403
		BM	517.5	0.2534	4.9381
		HM	505.5	**0.2535**	**4.9377**

表 6: 小分類あり集合に対する 5 分割交差確認の結果

	p	手法	時間（秒）	R^2	RMSE
CvsD	234	SW_{20}	1047.0	**0.3642**	**2.6522**
		$L1_{20}$	6.8	0.3612	2.6586
		BM_{20}	1000.0	0.3412	2.6997
		HM_{20}	1000.0	0.3618	2.6573
		SW_{50}	6018.8	**0.3701**	**2.6400**
		$L1_{50}$	17.5	0.3691	2.6422
		BM_{50}	1000.0	0.3552	2.6708
		HM_{50}	1000.0	0.3693	2.6417
CvsS	277	SW_{20}	1439.6	0.2440	5.1413
		$L1_{20}$	15.0	0.2369	5.1655
		BM_{20}	1000.0	0.2004	5.2865
		HM_{20}	1000.0	**0.2486**	**5.1256**
		SW_{50}	8331.7	0.2511	5.1172
		$L1_{50}$	29.9	0.2495	5.1228
		BM_{50}	1000.0	0.2314	5.1838
		HM_{50}	1000.0	**0.2511**	**5.1170**
DvsS	281	SW_{20}	786.7	**0.2256**	**6.7927**
		$L1_{20}$	10.1	0.2243	6.7984
		BM_{20}	1000.0	0.2088	6.8658
		HM_{20}	1000.0	0.2255	6.7930
		SW_{50}	4551.5	0.2286	6.7795
		$L1_{50}$	21.9	**0.2289**	**6.7782**
		BM_{50}	1000.0	0.2228	6.8045
		HM_{50}	1000.0	0.2289	6.7782
平均		SW	3695.9	0.2806	4.8538
		L1	16.2	0.2783	4.8610
		BM	1000.0	0.2600	4.9185
		HM	1000.0	**0.2809**	**4.8522**

表 6 の小分類あり集合の結果でも同様に，ステップワイズ法と階層的変数選択モデルの予測精度が高く，平均では僅かに階層的変数選択モデルが上回っていた．一方で基本モデルの予測精度は大きく悪化し，4 種類の手法の中で最も悪い値となった．このような結果となる理由としては，小分類あり集合は候補変数が多く，基本モデルの計算時間は 1000 秒では不十分であったことが挙げられる．一方で階層的変数選択モデルは，L_1 罰則付き回帰を上回る予測精度を得ることができた．階層的変数選択モデルでは，探索すべき選択変数の組合せが制約条件により削減されるために，計算効率が向上する．このため，候補変数が増えた場合にも良い性能を維持することができたと考えられる．

表 6 の CvsD データセットに対する予測精度では，階層的変数選択モデルはステップワイズ法に劣っているが，ステップワイズ法の平均計算時間は 1000 秒を超えており，特に選択変数の上限数を $s = 50$ とした場合の平均計算時間は 6000 秒を超えていた．ステップワイズ

法の計算を途中で終了すると選択変数が少なくなってしまい公平な比較とならないために，本研究の数値実験ではステップワイズ法には計算時間の上限を課さなかった．しかしながら CvsD データセットの場合には，SW_{20} の平均計算時間が 1047 秒であるのでステップワイズ法は 1000 秒の制限時間では 20 変数程度しか選択することができず，その場合の予測精度は HM_{50} の予測精度を大きく下回る．つまり選択変数が多く，計算時間が制限されているような状況では，階層的変数選択モデルの性能はステップワイズ法を上回ると考えることができる．L_1 罰則付き回帰に関しては，ステップワイズ法や階層的変数選択モデルと比較すると予測精度は劣ってはいるが，計算は非常に高速であった．

以上の結果をまとめると，本研究で提案する階層的変数選択モデルは，候補変数が少ない場合には他の手法と同等以上の予測精度を達成することができた．さらに候補変数が多い場合には，階層的変数選択モデルの予測精度は L_1 罰則付き回帰を上回り，基本モデルの予測精度を大きく改善していた．ステップワイズ法は階層的変数選択モデルと同程度の予測精度であったが，選択変数が多い場合に計算時間が長くなり，計算時間 1000 秒以内の性能では階層的変数選択モデルに劣っていた．

4.3 店舗選択要因の分析

本節では変数選択の結果を考察し，各業態の店舗が選択される要因を分析する．表 7–9 は小分類あり集合に対して，階層的変数選択モデルによって得られた 20 個の説明変数とそれらの偏回帰係数を示している．ここでは表 2 の各データセットの全サンプルを用いて変数選択を行ない，計算時間は 1 万秒とした．また商品の大分類，中分類，小分類をそれぞれ（大），(中)，(小）と記載している．

表 7 はコンビニとドラッグストアの選択モデルの結果であり，偏回帰係数の正値はコンビニ，負値はドラッグストアの選択に寄与することを表す．また偏回帰係数の絶対値は併売商品数を意味する．たとえば，説明変数「(中）その他日用品」は偏回帰係数の符号が正であることからコンビニを選択する要因と考えることができ，当該商品の購入によりコンビニにおける購買点数が 5.43 個増える傾向があると解釈できる．

階層的変数選択モデルは，選択された商品分類の上位分類も選択するという制約を加えたモデルであり，「（中）その他日用品」とその上位の「（大）日用品」や，「（小）その他食品」とその上位の「（中）加工食品」「(大）食品」などが実際に選択されている．ただし，偏回帰係数の符号には必ずしも共通性はなく，上位分類と偏回帰係数の符号が異なる場合もある．

店舗選択要因はこれまで実証的に「価格的要因」「利便性要因」「サービス要因」「商品要因」の 4 つに分類されており，これらの観点から考察を行なう．表 7 を見ると，「（中）その他日用品」はコンビニの選択要因であるが，その上位分類の「（大）日用品」はドラッグストアの選択要因となっている．「（大）日用品」はティッシュ，トイレットペーパー，洗剤などドラッグストアの目玉商品が多く含まれている分類であり，価格的要因によりコンビニではなくドラッグストアで購入されていると考えられる．一方でコンビニの「(中）その他日用品」は，Amazon や iTunes のギフトカードやゴミ収集券が大部分を占めており，偏回帰係数の値が 5.43 と最も大きく他の商品と同時購入されやすい分類である．これはチケット購入や公共料金の支払いなどと同じように，利便性要因によりコンビニで購入されていると考えられ，ついで買いを誘発しやすい商品分類であると解釈できる．それ以外にも「(小）その他食品」「（小）冷凍食品」「（小）アルコール飲料」「（小）惣菜類」などは偏回帰係数の絶対値がそれほど大きくなく，その商品の購入のみを目的としてコンビニが利用されている可能性が高く，利便性要因によりコンビニが選択されていると考えられる．

ドラッグストアの「(小）食品包装」はラップやアルミホイル，ジップロックなどの分類であり，これらは価格的要因によりドラッグストアで購入されていると考えられる．また，「(大）衣料・身の回り品・スポーツ用品」は，ストッキング，タイツ，足指サラピーなど女性の美容グッズのシェアが高い分類であり，ドラッグストア特有の商品として商品要因により購入されていると考えられる．また，「女性×子持ち」の顧客層にドラッグストアが支持されていることも示唆されており，節約のために家族の生活必需品をコンビニではなくドラッグストアで購入していることが想像できる．以上の結果から，コンビニは利便性要因，ドラッグストアは価格的要因や商品要因により顧客に選択されていることが分かる．

表 8 はコンビニとスーパーの選択モデルの結果であり，偏回帰係数の正値はコンビニ，負値はスーパーの選択に寄与することを表す．ここでも「(中）その他日用品」の偏回帰係数は 6.60 と大きな値となっており，前述のギフトカードなどはスーパーとの比較においてもコンビニの選択要因として顕著に現れている．また調理の手間が少ない「(小）パン・シリアル類」「(小）スープ」「(小）麺類」「(中）菓子類」「(小）冷凍食品」

表 7: CvsD データセットで選択された変数
（正例：コンビニ，負例：ドラッグストア）

説明変数	偏回帰係数	
（中）その他日用品	5.43	***
（小）滋養強壮関連（指定医薬部外品）	3.83	***
切片	2.73	***
（小）その他食品	2.29	***
（中）家庭用品	1.47	***
（小）衛生医療用品・用具	1.40	***
（小）冷凍食品	1.14	***
（小）衛生紙用品・用具	1.10	***
（小）アルコール飲料	0.97	***
（小）惣菜類	0.80	***
（中）OTC 医薬品類	0.05	
（中）加工食品	−0.16	***
（中）飲料・酒類	−0.27	***
女性×子持ち	−0.45	***
（大）食品	−0.99	***
（中）日用雑貨	−1.24	***
（小）水物	−1.28	***
（中）その他食品	−1.87	***
（小）食品包装	−2.42	***
（大）衣料・身の回り品・スポーツ用品	−2.43	***
（大）日用品	−4.74	***

***：0.1%有意，**：1%有意，*：5%有意

表 8: CvsS データセットで選択された変数
（正例：コンビニ，負例：スーパー）

説明変数	偏回帰係数	
（中）その他日用品	6.60	***
（小）滋養強壮関連（指定医薬部外品）	4.66	***
（小）パン・シリアル類	4.59	***
（小）スープ	3.70	***
（小）惣菜類	3.27	***
切片	3.09	***
（中）その他食品	2.99	***
（小）麺類	2.53	***
（中）菓子類	2.16	***
（中）飲料・酒類	1.91	***
（中）OTC 医薬品類	1.80	***
（小）冷凍食品	1.73	***
男性× 60 代以上	−0.59	***
女性×既婚	−0.81	***
（小）デザート・ヨーグルト	−1.04	***
（小）乳飲料	−1.09	***
女性×子持ち	−1.55	***
（大）耐久消費財	−2.90	***
（中）加工食品	−3.73	***
（大）食品	−3.76	***
（大）日用品	−5.60	***

***：0.1%有意，**：1%有意，*：5%有意

はコンビニの選択要因として現れており，利便性要因によるコンビニの選択傾向が読み取れる．一方で,「男性× 60 代以上」「女性×既婚」「女性×子持ち」などの顧客属性と,「（小）デザート・ヨーグルト」「（小）乳飲料」がスーパーの選択要因として現れている．これらの結果から，コンビニと比較してスーパーの特徴的な利用者は高齢男性や既婚女性であること，それゆえ健康を意識した商品が求められていることが示唆されており，商品要因によりスーパーが選択されていると考えられる．それ以外にもスーパーの選択要因として「（大）耐久消費財」「（大）食品」「（大）日用品」などが現れており，これらは GMS などの総合スーパーや大型のショッピングモールなどの特徴を表している．

表 9: DvsS データセットで選択された変数
（正例：ドラッグストア，負例：スーパー）

説明変数	偏回帰係数	
（中）その他食品	5.43	***
（小）パン・シリアル類	3.48	***
（小）穀物	3.38	***
（小）菓子	2.35	***
（大）日用品	2.19	***
（小）スープ	2.13	***
（小）麺類	1.60	***
（小）水物	1.35	***
（小）清涼飲料	1.35	***
（小）調味料	1.31	***
（中）飲料・酒類	0.46	***
切片	0.16	**
（中）菓子類	−0.13	
女性×子持ち	−0.62	***
（中）家庭用品	−2.61	***
（小）その他家庭用品	−2.65	***
（大）文化用品	−2.72	***
（中）ペット用品	−2.76	***
（大）食品	−3.29	***
（小）その他食品	−4.51	***
（中）加工食品	−4.70	***

***：0.1%有意，**：1%有意，*：5%有意

表 9 はドラッグストアとスーパーの選択モデルの結果であり，偏回帰係数の正値はドラッグストア，負値はスーパーの選択に寄与することを表す．スーパーとの比較においてもドラッグストアの選択要因として，化粧品・医薬品などの複数種類の日用品を含んだ「（大）日用品」が現れており，商品要因がドラッグストアの選択要因となっていることが分かる．一方で表 8 でコンビニの特徴として現れていた，乳幼児食品・健康食品・たばこなどを含む「（中）その他食品」，調理の手間が少ない「（小）パン・シリアル類」「（小）スープ」「（小）麺類」，嗜好品である「（中）飲料・酒類」などがドラッグストアの選択要因として現れている．したがってスーパーとの比較において，ドラッグストアとコンビニの選択要因は類似しており，そこには利便性要因や価格的要因が含まれていると考えられる．またドラッグストアでは客単価を向上させる上で，顧客の衝動買いを誘発することが重要とされており，偏回帰係数の値が大きい「（中）その他食品」「（小）パン・シ

リアル類」「(小) 穀物」などは併売戦略の軸となりうる商品群だと考えられる．一方でスーパーは「(中) 家庭用品」や「(大) 食品」など家庭用品と食品を中心とした商品要因によって顧客から選択されている．とりわけ他の業態と比較して食品がスーパーの主要な選択要因となっており，主婦や健康を意識した商品展開が重要になると考えられる．

最後に，店舗選択モデルの切片の値は，選択変数の影響を除いた店舗の集客力を表していると解釈することができる．表 7–9 の結果から，コンビニの集客力が高く，立ち寄り型の購買を多く誘発していることが示唆される．

5 おわりに

本研究では，商品分類間の階層構造を考慮した混合整数最適化による変数選択手法を提案し，対象店舗における購買点数に着目した店舗選択要因の分析を行なった．提案手法は，商品分類間の階層構造を活用してモデル構築の信頼性を高めることができ，探索すべき選択変数の組合せが減少するために計算効率が向上するという利点もある．本研究では 5 分割交差確認によってステップワイズ法や罰則付き回帰と性能を比較し，提案手法の有効性を実証した．

本研究では，提案手法を用いてコンビニ，ドラッグストア，スーパーの 3 種類の業態の選択要因を分析した．分析結果から，コンビニは顧客から利便性要因により選択されていること，またドラッグストアは価格的要因や商品要因により選択されていることが分かった．一方で，スーパーは他の業態と比較して子持ちの女性に支持されており，食品が主な選択要因となっていることが分かった．

本研究では 3 種類の業態を対象として選択要因の分析を行なったが，今後の研究の方向性としては，同じ業態に属する店舗間の選択要因を分析することや，地域や季節ごとの店舗選択要因の差異を分析することなど，提案手法を用いた様々な分析方針が考えられる．現在は業態間競争が加速しており，企業レベルや店舗レベルで商品開発や商品展開，価格設定などを工夫し，来店客数を増やすことが重要である．したがって本研究で示したような分析結果は，各業態の優位性や劣位性を理解し，各店舗の販売戦略を検討していく上で有用な情報であると言える．

謝辞

本研究の一部は，専修大学情報科学研究所の共同研究助成を受けたものである．

参考文献

[1] Baker, J., Parasuraman, A., Grewal, D., & Voss, G. B. (2002). The influence of multiple store environment cues on perceived merchandise value and patronage intentions. *Journal of Marketing, 66*, 120–141.

[2] Bertsimas, D., & King, A. (2015). An algorithmic approach to linear regression. *Operations Research, 64*, 2–16.

[3] Bertsimas, D., King, A., & Mazumder, R. (2016). Best subset selection via a modern optimization lens. *Annals of Statistics, 44*, 813–852.

[4] Bien, J., Taylor, J., & Tibshirani, R. (2013). A lasso for hierarchical interactions. *Annals of Statistics, 41*, 1111–1141.

[5] Bloemer, J., & de Ruyter, K. (1998). On the relationship between store image, store satisfaction and store loyalty. *European Journal of Marketing, 32*, 499–513.

[6] Blum, A. L., & Langley, P. (1997). Selection of relevant features and examples in machine learning. *Artificial Intelligence, 97*, 245–271.

[7] Briesch, R. A., Chintagunta, P. K., & Fox, E. J. (2009). How does assortment affect grocery store choice?. *Journal of Marketing Research, 46*, 176–189.

[8] Efroymson, M. A. (1960). Multiple regression analysis. *Mathematical Methods for Digital Computers, 1*, 191–203.

[9] Furnival, G. M., & Wilson, R. W. (2000). Regressions by leaps and bounds. *Technometrics, 42*, 69–79.

[10] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research, 3*, 1157–1182.

[11] Huang, J., Zhang, T., & Metaxas, D. (2011). Learning with structured sparsity. *Journal of Machine Learning Research, 12*, 3371–3412.

[12] Jacob, L., Obozinski, G., & Vert, J. P. (2009). Group lasso with overlap and graph lasso. In *Proceedings of the 26th International Conference on Machine Learning* (pp. 433–440).

[13] Jenatton, R., Audibert, J. Y., & Bach, F. (2011). Structured variable selection with sparsity-inducing norms. *Journal of Machine Learning Research, 12*, 2777–2824.

[14] 河原吉伸 (2013). 構造的な事前情報を用いた機械学習：構造正則化と劣モジュラ性. 情報処理, *52*, 734–740.

[15] Kim, S., & Xing, E. P. (2010). Tree-guided group lasso for multi-task regression with structured sparsity. In *Proceedings of the 27th International Conference on Machine Learning* (pp. 543–550).

[16] Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial Intelligence, 97*, 273–324.

[17] Konno, H., & Yamamoto, R. (2009). Choosing the best set of variables in regression analysis using integer programming. *Journal of Global Optimization, 44*, 273–282.

[18] Leszczyc, P. T. P., & Timmermans, H. (2002). Experimental choice analysis of shopping strategies. *Journal of Retailing, 77*, 493–509.

[19] Liu, H., & Motoda, H. (Eds.). (2007). *Computational methods of feature selection*. CRC Press.

[20] Miyashiro, R., & Takano, Y. (2015). Subset selection by Mallows' C_p: A mixed integer programming approach. *Expert Systems with Applications, 42*, 325–331.

[21] Miyashiro, R., & Takano, Y. (2015). Mixed integer second-order cone programming formulations for variable selection in linear regression. *European Journal of Operational Research, 247*, 721–731.

[22] Pan, Y., & Zinkhan, G. M. (2006). Determinants of retail patronage: A meta-analytical perspective. *Journal of Retailing, 82*, 229–243.

[23] Reutterer, T., & Teller, C. (2009). Store format choice and shopping trip types. *International Journal of Retail & Distribution Management, 37*, 695–710.

[24] Sato, T., Takano, Y., & Miyashiro, R. (2015). Piecewise-linear approximation for feature subset selection in a sequential logit model. *arXiv preprint, arXiv:1510.05417*.

[25] Sato, T., Takano, Y., Miyashiro, R., & Yoshise, A. (2016). Feature subset selection for logistic regression via mixed integer optimization. *Computational Optimization and Applications, 64*, 865–880.

[26] Sato, T., Takano, Y., & Nakahara, T. (2016). Using mixed integer optimisation to select variables for a store choice model. *International Journal of Knowledge Engineering and Soft Data Paradigms, 5*, 123–134.

[27] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, B58*, 267–288.

[28] Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., & Knight, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society, B67*, 91–108.

[29] Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society, B68*, 49–67.

[30] Yusta, S. C. (2009). Different metaheuristic strategies to solve the feature selection problem. *Pattern Recognition Letters, 30*, 525–534.

[31] Zhao, P., Rocha, G., & Yu, B. (2009). The composite absolute penalties family for grouped and hierarchical variable selection. *Annals of Statistics, 37*, 3468–3497.