

Sparse Modeling for Marketing

YU TOMITA

1. イントロダクション
2. 数学的定式化
3. スパースモデリングの応用
4. 論文の紹介
5. 研究のモチベーション

Outline

スパースモデリングとは何か

○スパース性の定義:

「物事やデータの本質的な特徴を決定づける要素はわずかである」という性質

○重要性:

- ①データが不足していても解析可能
- ②大量のデータを使わずにすみ、データ解析にかかる時間を短縮できる
- ③複雑な構造を分かりやすく表現する

スパースモデリングとは何か

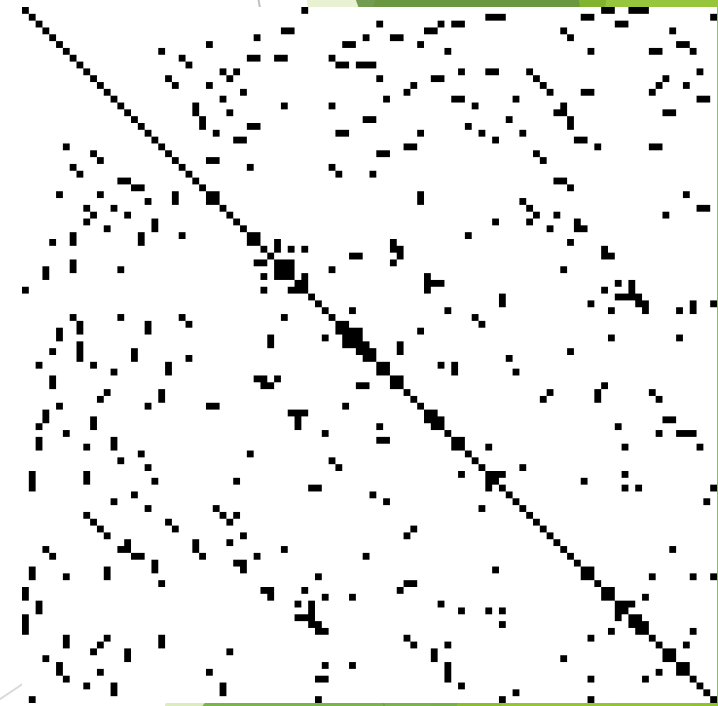
○スパースなデータの例

$$\begin{pmatrix} 1.0 & 0 & 5.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.0 & 0 & 0 & 0 & 0 & 11.0 & 0 \\ 0 & 0 & 0 & 0 & 9.0 & 0 & 0 & 0 \\ 0 & 0 & 6.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7.0 & 0 & 0 & 0 & 0 \\ 2.0 & 0 & 0 & 0 & 0 & 10.0 & 0 & 0 \\ 0 & 0 & 0 & 8.0 & 0 & 0 & 0 & 0 \\ 0 & 4.0 & 0 & 0 & 0 & 0 & 0 & 12.0 \end{pmatrix}$$

スパース推定

○データの本質的な部分を抽出する

- ① データに合わせて近似を行う
- ② いくつかの要素を排除(重要でないパラメータを0と推定する)



Source: Sparse matrix

数学的定式化

○定式化: ノルム正則化

$$\min_w L(W) + \lambda \sum |w|$$

W: モデルのパラメータ

λ : 正則化パラメータ ($\lambda > 0$)

第一項: 解の候補を選ぶ

第二項: 解の候補を絞る

数学的定式化

○定式化: ノルム正則化

$$\min_w L(W) + \lambda \sum |w|$$

最適化の方法は . . .

1. Coordinate Descent
2. Iteratively Reweighted Shrinkage
3. proximal gradient method etc...

$L(W)$ に合わせて選ぶ必要がある

スパースモデリングの応用例

1. 不足するデータを補完

モデルのパラメータ推定に十分にデータを得られない場合でも、精度の高い推定値を得られる

例：①反射法地震探査 ②カンニング検出

2. ビッグデータを省く

手元にあるビッグデータのすべてを使うことなく、モデルのパラメータを推定できる

例: MRI

3. モデルを単純化

データの一部を省いた方が精度が高くなる時もある

例：津波がどこまで到達したかを地層の堆積物で分析

論文の紹介①

題名:飲食サービスを利用したおもてなし消費構造の分析

著者: 宮井、西尾(2014)

要約:

1. 「おもてなし」という概念に注目
2. 「価値の共創」「顧客間相互作用」といった概念を手がかりに消費構造をモデリング
3. ロジスティック回帰モデルを使ってアンケートデータを解析
4. 「顧客間相互作用」と「店舗側のサービス提供」の間に交互作用効果を確認した

論文の紹介②

○モデルのコンセプト

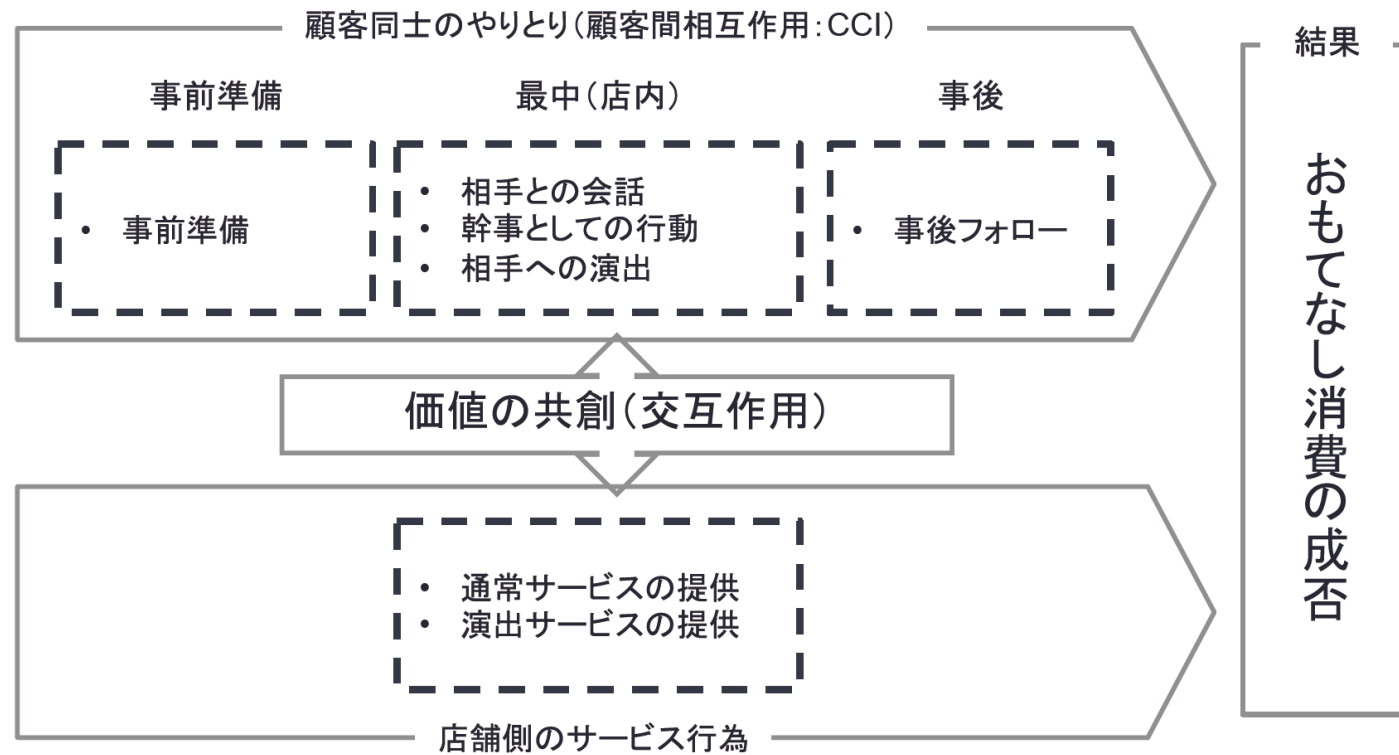


図 1 おもてなし消費のステップ

論文の紹介③

○実際のアンケート

おもてなしプロセスを構成する要素

一般的なおもてなしのすすめ方をいくつかの要素に分けて、提示いたします。それぞれの段階ごとのあなたの評価について、お答えください。

評価は、以下を基準に100点満点で数値をご記入ください。

あなたがお答えになったおもてなし経験で行われなかった要素については「999」をご記入ください。

- 100点

（予想よりもはるかによかった・うまくできた）
- 60点

（当初の期待通り・まあ合格点）
- 0点

（予想よりもはるかに悪かった・うまくいかなかった）
- 999

（この要素については特に評価は無い わからない）

| | |
|--|----------|
| 1.おもてなし当日を迎えるまでの準備 （メニューや席の予約，お店との相談など） | 【 】点 |
| 2.おもてなし当日の，おもてなし相手とあなたの会話 | 【 】点 |
| 3.おもてなし当日の，貴方の幹事としての行動 （メニューのオーダー，店員とのやりとりなど） | 【 】点 |
| 4.おもてなし当日の，成功のために貴方が行ったこと （プレゼント，サプライズ，写真撮影など。貴方がお願いして，お店の人にやってもらったことも含みます） | 【 】点 |
| 5.おもてなし当日の通常のお店のサービス （料理メニュー等の説明や配膳など） | 【 】点 |
| 6.おもてなしの成功のためにお店が自主的に行ったこと（プレゼント，サプライズ，写真撮影など。貴方がお店の人をお願いして，やってもらったことは除きます） | 【 】点 |
| 7.店を出た直後から後日までのアフターフォロー （電話・メールのやりとりなど） | 【 】点 |

論文の紹介④

○提案モデル

▶
$$P_h = Pr(y_h = 1) = \frac{\exp(v_h)}{1 + \exp(v_h)}$$

▶
$$v_h = \beta_0 + \sum_{i=1}^7 \beta_i x_{ih} + \sum_{i=1}^6 \sum_{j=i+1}^7 \beta_{ij} x_{ij,h}$$

以上の式のそれぞれのパラメータ β
を推定する

表 1 本研究において仮定する変数

| 記号 | 変数 | 変数の詳細 |
|------------|------------|--|
| y_h | おもてなし消費の成否 | 成功=1, 失敗=0 |
| $x_{1,h}$ | 事前準備 | おもてなし当日店内に入るまでの準備(メニューや席の予約、お店との相談など) |
| $x_{2,h}$ | 顧客同士の会話 | おもてなし当日の自分と相手との会話 |
| $x_{3,h}$ | 幹事としての行動 | おもてなし当日の自分の幹事としての行動(メニューのオーダー、店員とのやりとりなど) |
| $x_{4,h}$ | 相手への演出 | おもてなし当日の自分が行なった演出(プレゼント、サプライズ、写真撮影など。お店の人にお願いしてやってもらったことも含む) |
| $x_{5,h}$ | 店舗側通常サービス | おもてなし当日のお店のサービス(料理メニュー等の説明や配膳など) |
| $x_{6,h}$ | 店舗側演出サービス | おもてなしの成功のためにお店が自主的に行なったこと(プレゼント、サプライズ、写真撮影など。お店の方があくまで自主的に行なったことに限る) |
| $x_{7,h}$ | 事後フォロー | 店を出た直後から後日までのアフターフォロー |
| $x_{ij,h}$ | 各変数の交互作用項 | $x_{1,h}$ から $x_{7,h}$ までの変数各々の交互作用項 (ただし, $j > i$) |

論文の紹介⑤

- ▶ $\beta_{MLE} = \operatorname{argmin}\{-l(B)\} = \operatorname{argmin}[-\sum_{i=1}^n \{y_h \log P_h + (1 - y_h) \log(1 - P_h)\}]$
- ▶ $\beta_{lasso} = \operatorname{argmin}\{-l(B) + \lambda(\sum_{i=1}^7 |\beta_i| + \sum_{i=1}^6 \sum_{j=i+1}^7 |\beta_{ij}|)\}$
- ▶ 最尤法で推定すると多重共線性が発生する
- ▶ この論文では、Lasso推定を用いている

論文の紹介⑥

～推定結果～

表 3 モデルのパラメータ推定結果

| 変数 | 推定値 |
|--|--------|
| β_0 切片項 | -2.130 |
| β_1 事前準備 ($x_{1,h}$) | -1.483 |
| β_6 店舗側演出サービス ($x_{6,h}$) | -7.860 |
| β_{24} 顧客同士の会話 ($x_{2,h}$)と相手への演出 ($x_{4,h}$)の交互作用項 | 0.225 |
| β_{25} 顧客同士の会話 ($x_{2,h}$)と店舗側通常サービス ($x_{5,h}$)の交互作用項 | 0.279 |
| β_{27} 顧客同士の会話 ($x_{2,h}$)と事後フォロー ($x_{7,h}$)の交互作用項 | 0.048 |
| β_{45} 相手への演出 ($x_{4,h}$)と店舗側通常サービス ($x_{5,h}$)の交互作用項 | 0.052 |
| λ_1 正則化パラメータ | 0.010 |

研究のモチベーション

1. 多くの場合、マーケティングデータはスパースである
2. 十分なデータ量を得られないケースに用いることが可能
3. 欠損値を推定するのに有効
4. 複雑な構造を人間が理解できるように表現できる
5. 他の手法と組み合わせるのが比較的容易

Reference

1. スパース性に基づく機械学習(富岡,2015,講談社)
2. スパースモデリングってなんだ(日高他,2017,カットシステム)
3. 飲食サービスを利用したおもてなし消費構造の分析(宮井弘之,西尾チヅル,2014,マーケティング・サイエンスVol. 22 pp. 31 – 46)
4. The Elements of Statistical Learning(Hastie et al.,2008)