

ベイズ最適な店舗の組み合わせの探索と 予測式の混合によるレストランチェーンの来客数予測モデルの構築

1X07C053-3 小宮洋一
指導教員 後藤正幸

1 研究背景と目的

首都圏を中心に店舗を展開するレストランチェーン S 社では、経済情勢の悪化に伴う環境変化の煽りを受け、効率的な経営を行う必要に迫られている。効率的な経営とは、最適な人員数による業務運営や発注量最適化などが挙げられるが、これらを実現するためには、正確な来客数予測を行うことが必要不可欠である。その際、各店舗において来客数に影響を与える基本要因の効果を定量化するなど、現場レベルで理解可能な扱いやすい予測モデルである必要がある。本研究では S 社を事例とし、店舗ごとの基本要因の差異分析と来客数の予測精度向上を目的とした予測モデルを提案する。

来客数予測に対しては、一般的に POS データを用いた多変量解析などの方法が考えられる。しかし来客数予測に使用できる POS データは、需要構造の変化やメニュー改変などにより分析に利用できる期間が限定的である場合が多く、店舗単位のモデル化では十分なデータ数が確保できないことが多い。そこで、店舗効果を 1 つの要因と考えて、予測モデルに取り込み、モデル選択基準によって最適なモデルを選択する方法が考えられる。

一方で、ベイズ統計の観点によれば全ての予測モデルを混合することで予測精度が向上することが知られているが、考え得るモデル全ての混合を取ることは、店舗の特性が見えにくい複雑な予測モデルの構築に繋がるうえに、時間計算量的な困難も生じる。

そこで本研究では、モデル選択基準を用いて基本要因の影響が類似した店舗の組み合わせを 1 つ選択し、その組み合わせの店舗群ごとに、考え得る全ての予測モデルを混合するというモデルの構築法を提案する。さらにモデルの混合に [1] の方法を適用し、回帰係数パラメータの混合値を得る方法により、店舗群ごとの基本要因の効果を定量的に把握でき、かつ来客数予測の精度も向上することを示す。

2 準備

2.1 研究対象

世界中に店舗を展開する S 社は、日本において首都圏に 9 店舗（東京都：7 店舗，神奈川県：2 店舗）を出店している。従来 S 社各店舗は、地域に根差した運営を目指してきたこともあり、来客数予測を各店長が経験と勘で行ってきた。このため現場では、実務的な来客数予測モデルを構築することが求められている。また、店長や現場スタッフの暗黙知によれば、来客数を変動させる基本要因は曜日、月、天気である。さらにそれら要因の効き方は店舗によって異なるが、傾向が類似する店舗もあることが分かっている。このため基本要因の影響が店舗によってどの程度異なっているかを定量的に把握し、傾向の類似する店舗をまとめて管理したいという要求がある。

そこで類似した店舗を組み合わせ、POS データを統合して予測モデルを構築することを考える。例えば 3 つの要素の組み合わせを考えると、 $\{(1),(2),(3)\}$ 、 $\{(1),(2,3)\}$ 、 $\{(2),(1,3)\}$ 、

$\{(3),(1,2)\}$ 、 $\{(1,2,3)\}$ の 5 通りが組み合わせの総数である。この組み合わせの総数は、ベル数という数で求めることができる。S 社では 9 店舗を出店しているため、店舗の組み合わせの総数は 21,147 にもなる。

2.2 予測モデル

来客数予測モデルとして、線形回帰モデルを採用する。来客数を変動させる基本要因は、質的変数であるので、基本要因が取る値は質的なカテゴリである。例えば日曜日は曜日効果のカテゴリである。基本要因数を R 、基本要因 j のカテゴリ数を c_j 、 α_{jk} を基本要因 j におけるカテゴリ k の効果の大きさ、データ数を n 、定数項を α_0 、ダミー変数を $\delta_{i(jk)}$ (データ i が基本要因 j のカテゴリ k ならば 1、その他は 0) とすると、来客数予測モデルは (1) 式で表される。

$$\hat{y}_i = \hat{\alpha}_0 + \sum_{j=1}^R \sum_{k=1}^{c_j} \hat{\alpha}_{jk} \delta_{i(jk)} \quad (i = 1, 2, \dots, n). \quad (1)$$

モデルに取り込むべき基本要因の数やカテゴリ数の取り方で、複数の予測モデルが想定される。説明変数の数は $t = \sum_{j=1}^R c_j$ で与えられ、考え得るモデルは説明変数を 1 つも取り込まない場合も含め 2^t 通りになる。これらのモデルは、例えば基本要因 j のカテゴリ k の効果を取り込まない場合には、 $\hat{\alpha}_{jk} = 0$ と制約を課すことで得られる。

2.3 従来研究

店舗売上などの予測モデルに関する研究では、山口らの研究 [2] などがある。これらの研究では最適なモデルを 1 つ選択する方法を採用している。

統計的モデル選択問題の際に用いられるモデル選択基準の 1 つに Bayesian Information Criterion (BIC) がある。BIC は真のモデルの探索を目的としたモデル選択基準で、モデルの事後確率の漸近式で与えられる。

一方ベイズ統計の観点で考えた場合、考え得る全てのモデルを事後確率の重み付けで混合することが、ベイズ最適な予測となることも知られている。しかしモデルを混合するための事後確率計算には、複雑な積分計算が伴うため、鈴木らは漸近近似的に計算量を抑えた予測法を提案している [1]。この漸近近似式は本質的には BIC と等価である。

3 提案手法

店舗効果を取り込んだ学習データに対しては、店舗の組み合わせの総数が膨大であるため、鈴木らの手法を取り入れても時間計算量的な困難が生じる。この問題は類似した店舗の組み合わせを 1 つ選択し、その組み合わせの下で考え得る全てのモデルの混合を取ることで解決できると考えられる。

本研究では、店舗の要因分析と予測精度を加味した 2 段階の手順を踏む手法を提案する。第 1 段階では BIC を用いて最適な店舗の組み合わせを 1 つ探索し、第 2 段階ではその組み合わせの店舗群で考え得る来客数予測モデル全てを混合する。本提案では混合の重み付けの際に BIC が必要なため、モデル選択基準として BIC を用いた。

3.1 最適な店舗の組み合わせ

店舗の組み合わせにおいて、同じ要素としてまとめられた店舗をグループと呼ぶことにする。この店舗の組み合わせ 1 つ 1 つに対し BIC を計算するが、その際組み合わせを構成するグループごとに、BIC の値が最大になるように変数選択を行う。変数選択は BIC による変数増減法を用いる。ここで店舗の組み合わせ C_r ($r = 1, \dots, 21147$) における、グループ数を G_r 、グループのデータ数を $n_{r,g}$ 、グループに対して BIC で変数選択をし得られた残差平方和を $S_{r,g}$ 、グループのパラメータ数を $k_{r,g}$ とすると、店舗の組み合わせ C_r の BIC は (2) 式で算出できる。

$$BIC(C_r) = -\frac{\sum_{g=1}^{G_r} n_{r,g}}{2} \left(\log 2\pi - \frac{\sum_{g=1}^{G_r} S_{r,g}}{\sum_{g=1}^{G_r} n_{r,g} - 1} - 1 \right) - \frac{\sum_{g=1}^{G_r} k_{r,g}}{2} \log \sum_{g=1}^{G_r} n_{r,g}. \quad (2)$$

全ての c_r に対し BIC を計算し、 $\hat{C}_r = \arg \max_{C_r} BIC(C_r)$ を満たす \hat{C}_r を最適な店舗組み合わせとする。

3.2 予測式の混合方法

3.1 で選択された組み合わせにおいて、考え得る全ての予測モデルに対し、事後確率による重み付けを用いて混合を行う。この全てに対し BIC を算出し、予測モデル m_l の重み w_l を (3) 式で与える。

$$w_l = \frac{e^{\text{BIC}(m_l)}}{\sum_{l=0}^{2^t-1} e^{\text{BIC}(m_l)}} \quad (0 \leq l \leq 2^t - 1). \quad (3)$$

このとき混合された定数項 $\hat{\beta}_0$ と混合された基本要因 j におけるカテゴリ k の効果の大きさ $\hat{\beta}_{jk}$ を

$$\hat{\beta}_0 = \sum_{l=0}^{2^t-1} \omega_{m_l} \hat{\alpha}_0^{m_l}, \quad \hat{\beta}_{jk} = \sum_{l=0}^{2^t-1} \omega_{m_l} \hat{\alpha}_{jk}^{m_l}, \quad (4)$$

とすると、混合予測モデルは、

$$\hat{y}_i = \hat{\beta}_0 + \sum_{j=1}^R \sum_{k=1}^{c_j} \hat{\beta}_{jk} \delta_{i(jk)} \quad (i = 1, 2, \dots, n), \quad (5)$$

で与えられる。従って、混合されたパラメータ (4) 式を用いて、混合予測モデルは、(5) 式の 1 つの回帰式の形で表現できる。そのため (4) 式による値を、各要因の効果とみなすことが可能になる。

4 評価実験

4.1 実験条件

S 社の 9 店舗における 2008 年 10 月 22 日から 1 年分を学習データ、2009 年 10 月 22 日から 1 年分をテストデータとして実験を行う。データには欠損もあるため、店舗ごとにデータ数は異なっている。なお、ゴールデンウィーク・クリスマス・三が日など来客数の挙動が異なる日時を除去した上で、予測モデルを構築する。

4.2 分析結果

店舗の組み合わせ全てにおいて、グループごとに変数選択を行い、グループごとに来客数予測モデルを構成し、その後 BIC を算出した。その結果を表 1 に示す。表中の店舗組み合わせの項目番号は、9 店舗に番号を付与したものである。

表 1. 店舗の組み合わせによる BIC の値

順位	店舗組み合わせ	BIC
1	(9)(3,5,6,7,8)(4)(1,2)	-16984.978
⋮	⋮	⋮
3086	(9)(8)(7)(6)(5)(4)(3)(2)(1)	-17157.660
⋮	⋮	⋮
21147	(2,4,9)(1,8)(7)(6)(5)(3)	-17651.260

表 1 より、BIC が最大である最適な店舗の組み合わせは、 $\{(9), (3, 5, 6, 7, 8), (4), (1, 2)\}$ である。この組み合わせの各グループにおける混合予測モデルの予測精度の結果を表 2 に示す。評価基準は平均二乗誤差と平均絶対誤差を用いた。

表 2. 変数選択と予測式の混合による予測精度の評価

評価	店舗 1, 2		店舗 3, 5, 6, 7, 8	
	選択	混合	選択	混合
平均二乗誤差	73.390	73.693	66.651	64.353
平均絶対誤差	58.087	58.469	50.819	47.960
店舗 4	店舗 9		店舗 9	
	選択	混合	選択	混合
平均二乗誤差	119.323	118.386	53.256	53.011
平均絶対誤差	87.783	86.640	42.136	41.987

5 考察

店舗の組み合わせ全ての BIC を算出した結果、表 1 より店舗ごとに来客数予測を行うよりも店舗を組み合わせた方が、BIC の値が向上したことが分かる。その理由としては、店舗別の限られたデータで予測モデルを構築するという問題を解決できたためと考えられる。最適な店舗の組み合わせの各グループは、実務的に説明のつく組み合わせとなった。

また表 2 より、多くの場合、選択された単一のモデルによる予測よりも、混合モデルで予測を行う方が精度が向上したことが分かる。

次に最適な店舗の組み合わせの下で、考え得る全ての予測モデルを混合した際の、特徴的な偏回帰係数を表 3 に示す。

表 3. 各店舗の混合による偏回帰係数の値

偏回帰係数	9	3, 5, 6, 7, 8	4	1, 2
定数項	416.745	486.227	436.786	481.976
日	-0.160	1.276	-5.623	5.009
金	-236.573	-278.681	-263.635	-134.292
土	-59.550	-66.094	-34.398	-0.896
3月	49.386	24.398	0.483	0.445
7月	0.491	9.598	199.874	-0.476
8月	-0.187	27.900	442.591	-17.908
10月	-41.212	-33.954	-2.598	-92.644
晴(AM)	-0.126	-0.057	58.140	1.441
店舗1	-	-	-	61.904
店舗3	-	4.325	-	-

表 3 より、例えば店舗 4 における 7 月と 8 月の偏回帰係数の値は、他の店舗と比べ特異であると捉えることができる。

6 まとめと課題

本研究では、最適な店舗の組み合わせを導出し、その下で考え得る予測モデル全てを混合することで、店舗ごとの基本要因の差異分析と予測精度の向上を可能とする方法を提案した。今後の課題として、基本要因ごとに混合を行い、予測精度の向上を図ることなどが考えられる。

参考文献

- [1] 鈴木友彦, 後藤正幸, 俵信彦, “線形回帰モデルのベイズ最適な予測法に関する研究,” 日本経営工学会論文誌 51(1), pp.59-69, 2000
- [2] 山口類, 土屋映子, 樋口知之, “状態空間モデルを用いた飲食店売上の要因分解,” オペレーションズ・リサーチ: 経営の科学 49(5), pp.316-324, 2004