

# 広告・販促効果に関する外部入力付自己回帰モデル解析

## Auto-Regressive eXogenous input model analysis of advertisement and sales promotion

深田 健太\*<sup>1</sup>      鷲尾 隆\*<sup>1</sup>      矢田 勝俊\*<sup>2</sup>      元田 浩\*<sup>1</sup>  
 Kenta Fukata      Takashi Washio      Katutosi Yada      Hiroshi Motoda

\*<sup>1</sup>大阪大学産業科学研究所知能システム科学研究部門高次推論方式研究分野

Department of Advanced Reasoning, Division of Intelligent Systems Science,  
 The Institute of Scientific and Industrial Research, Osaka University

\*<sup>2</sup>関西大学商学部

Faculty of Commerce, Kansai University

In marketing study, static response of sales to some exogenous factors such as advertisement and sales promotion have been analyzed. However, if we can model dynamic response of sales to exogenous factor, changes more precise strategies of the sales can be designed in marketing. To achieve this modeling, an Auto-Regressive eXogenous input (ARX) model which has been widely used in engineering fields. A difficulty in this model application is the determination of an appropriate model order for each data. In this paper, we develop a new and practical approach to determine the appropriate order. Moreover, we apply the developed technique to a real marketing data, and analyse dynamic response character of sales revenue to advertisement and sales promotion.

### 1. はじめに

スーパーマーケットで物を販売する場合において、どの程度チラシを配り、CMを打ち、どのような条件下で販売を行った場合に、売上がどのように変化したかを記録した時系列のデータが存在する。そのような時系列の実データにおいて、それぞれの行動がどの程度、売上高に貢献したかを知ることができれば、例えば、広告への投資額を売上高の増加によりどの程度回収、増幅可能か否かのリスク予測を行うことができる。その結果に基づいて、あと10%多く広告を打てば最大の利益を得られることや何日後に販売促進活動を店内で行えば最高に売り上げを伸ばすことができるなどの情報を得ることができ、マーケティング分野全般において非常に役立つ。しかし、以上のような広告や販売促進のタイミングまでを含めた効果や影響を評価するには、対象データの動的変化に関するモデルが必要となる。

そこで、本研究では時間変化を陽に考慮して外部要因変化に対する動的変化のモデルを得ようとする。特に、外的条件の変化に対する対象の動的変化に関する代表的モデルである外部入力付自己回帰モデル [2] は、多くの実時系列データのマイニングにおいて有力な解析手段になり得る。しかし、その外部入力付自己回帰モデルを作成する際に重要なモデル次数決定において、現手法にはいくつかの問題が存在しており、うまくモデル化できないことが多い。そこで本稿では外部入力付自己回帰モデル作成に伴う現状の次数決定手法である AIC をより発展させた  $\Delta AIC^*$  による新しい次数決定手法の開発を行う。また、それを用いて実データへの適用検証を行うことで、マーケティング分野における広告や販促に対する動的売り上げ変化特性の解析方法の提案とその可能性評価を行う。

### 2. 現状課題

#### 2.1 外部入力付き自己回帰モデル

以下の外部入力付自己回帰モデルは予測対象とする変数を、その自身の過去の時系列と他の外部入力変数の過去の時系列か

連絡先: 大阪大学産業科学研究所

〒567-0047 大阪府茨木市美穂ヶ丘 8-1

E-mail: kenta-f@ar.sanken.osaka-u.ac.jp

ら、漸化的に線形予測するモデルである。

$$x(s) = \sum_{i=1}^p a_i x(s-i) + \sum_{j_1=1}^{q_1} b_{j_1} y_1(s-j_1-k_1) + \dots + \sum_{j_h=1}^{q_h} b_{j_h} y_h(s-j_h-k_h) + e(s) \quad (1)$$

ここで  $s$  は時間ステップを表し、 $x$  は予測対象変数、 $y_g$  ( $g = 1, \dots, h$ ) は外部入力変数を表す。 $a_i$  は過去の対象変数から現在の対象変数への影響係数、 $b_{j_g}$  は各外部入力変数の対象変数への影響係数、 $k_g$  は各外部入力変数の対象変数への影響の時間遅れステップ数、 $p, q_g$  ( $g = 1, \dots, h$ ) は対象変数が自身や外部入力変数から受ける影響の過去最大時間ステップ数で次数と呼ばれる。また、 $\hat{x}(s)$  を  $x(s)$  の予測値とすると、 $e(s) = x(s) - \hat{x}(s)$  は  $x(s)$  の予測誤差である。モデル係数である  $a_i$  ( $i=1, \dots, p$ ) と  $b_{j_g}$  ( $j_g=1, \dots, q_g$ ) ( $g=1, \dots, h$ ) は最小二乗法を用いて、この予測誤差を最小にするように決定される。モデルパラメータ数は従来は次節に説明する AIC を用いて決定してきた。

#### 2.2 従来の手法の問題点

外部入力付自己回帰モデルを求める際、従来は AIC (Akaike information criterion) を用いてモデル化がなされてきた。AIC とは実際の確率分布  $x(s)$  とその予測値の確率分布  $\hat{x}(s)$  の誤差を最小にするために用いる情報量規準である。カルバックライブラーの情報量を用いて、この二つの確率分布の誤差を測ると AIC は次式のように定義できる。

$$AIC = N \log(2\pi \hat{\sigma}_M^2) + N + 2(M+1) \quad (2)$$

AIC が小さい程、実際の分布と予測値の分布が近くなり、正確なモデル化ができると考えられる。

しかし、従来の AIC を用いた外部入力付自己回帰モデル作成における問題点として、AIC はモデリング対象データの測定誤差や対象の非線形性によってもたらされるモデリング誤差を考慮していないことが挙げられる。時系列データに誤差や非線形性が含まれない場合、ある一定のモデル複雑さまでは AIC が下がり、あるモデル複雑さを超えると上昇する。そして、最小 AIC のモデル複雑さを選択すればよいことが知られている。しかし、測定誤差や非線形性を含む時系列データでは、測定や非線形性誤差までを意味のある差異と見なし、実際以上の複雑さ

のモデルを選択することが多い。そこで実際以上に複雑なモデル選択を回避することが可能な新しい基準を考案する必要があると考えられる。

### 3. モデル次数決定手法の開発

#### 3.1 $\Delta AIC^*$ の提案

2. 2節で述べたように、モデルが実際以上に複雑にならない、モデル選択を行う指標が必要である。そこで、誤差よりも十分大きく、急激に AIC が減少するモデル複雑さを考える。急激に AIC が減少する複雑さのモデルでは、モデルを複雑にすることで、測定データの特性をうまくモデル化できるようになると考えられる。また、これ以上に大きく AIC が減少する複雑さのモデルがなければ、それより複雑なモデルを選択することは意味がないと考えられる。つまり、モデルを単純化すると急激に AIC の値が増加してしまう複雑さのモデルが最も妥当である。そこで、モデルの単純化により AIC の値が急激に増加してしまう複雑さである次数パラメータの組み合わせを探索するために、次に挙げるような  $\Delta AIC^*$  という新しい指標を提案する。 $\Delta AIC^*$  はある次数パラメータの組み合わせの AIC とその 1 パラメータ数単純な次数パラメータ組み合わせの AIC の差を取り、その中で最も値が大きいものと定義する。そして、すべての次数パラメータの組み合わせについて  $\Delta AIC^*$  を計算し、 $\Delta AIC^*$  が最も小さい次数パラメータの組み合わせを最も適切な次数パラメータの組み合わせであると考えられる。定量的には  $\Delta AIC^*$  は次式のように定義される。

$$\Delta AIC^* = \max(\Delta AIC_p, \max_{g=1, \dots, h} (\Delta AIC_{q_g})) \quad (3)$$

$$\begin{aligned} \Delta AIC_p &= AIC(p, q_1, q_2, \dots, q_h) - AIC(p-1, q_1, \dots, q_h) \\ \Delta AIC_{q_g} &= AIC(p, q_1, \dots, q_g, \dots, q_h) - AIC(p, q_1, \dots, q_g-1, \dots, q_h) \end{aligned}$$

この  $\Delta AIC^*$  を用いることで、モデルの単純化による最も大きな AIC の変化ではなく、最も小さな変化が最大になる次数パラメータ組み合わせを選択する。これは、それ以上の単純化によって少なくとも誤差オーダーを超えて大きく精度が失われる可能性の高い最も単純なモデル次数パラメータの組み合わせを発見するためである。ここで、ある次数パラメータが 0 のときはそれより単純なモデルが存在しないため、 $\Delta AIC^*$  を計算できない。しかし、ある次数パラメータが 0 であるときはそれに対応する変数と目的変数にモデル上の関係がないので、その変数をモデルから除いて  $\Delta AIC^*$  を計算する。上記 AIC の定義式から、 $\Delta AIC^*$  は下式のように書き換えることができる。

$$\Delta AIC^*_{p, q_1, q_2, \dots, q_h} = N \log \frac{\sigma_{p, q_1, q_2, \dots, q_h}^2}{\min_{g=1, \dots, h} (\sigma_{p-1, q_g}^2, \sigma_{p, q_1, \dots, q_g-1, \dots, q_h}^2)} + 2 \quad (4)$$

これより、すべての次数パラメータの組み合わせにおいて、二乗誤差比を計算すれば、 $\Delta AIC^*$  の大小は評価でき、かつ二乗誤差比が最も小さくなる次数パラメータの組み合わせを探索することで、最も妥当にモデル化できる次数パラメータの組み合わせを探索できることが分かる。

### 4. モデル次数探索手法の開発

次数パラメータの探索において、 $k_1, k_2$  を変えても、大きく AIC が減少することはなく、 $\Delta AIC^*$  では  $k_1, k_2$  に関する次数パラメータの組み合わせが決定できないため、 $k_1, k_2$  の AIC 値が最小のモデルが最良と考えて、 $k_1, k_2$  を決定することにした。 $k_1, k_2$  について AIC の値が大きく変化しない理由は、これらがモデル構造に、多大な影響を与えるパラメータではないためである。

#### 4.1 全探索法の問題点

すべての次数パラメータの組み合わせにおいて、 $\Delta AIC^*$  を求め、その中で  $\Delta AIC^*$  が最も小さくなる次数パラメータの組み合わせを求めるための手法として、まず、for loop による全探索法が挙げられる。for loop による全探索法は  $p, q_g$  のそれぞれのパラメータの探索範囲を  $1, 2, 3, 4, \dots, L$  ( $L$  は探索範囲の上限) に設定し、各パラメータ探索範囲の整数値のすべての組み合わせの中で、最小の  $\Delta AIC^*$  を探索することで、最適な次数パラメータの組み合わせを決定する手法である。つまり、for loop による全探索法の計算量は  $O(L^{h+1})$  となる。

#### 4.2 A\*探索法の提案

$\Delta AIC^*$  は式 (4) により表されるので、 $\Delta AIC^*$  の値が最小となるのは二乗誤差比が最小のときである。ここで二乗誤差は、与えられたデータに対する当てはめモデルの次数パラメータ  $p, q_g (g=1, 2, \dots, h)$  が増加すると必ず減少する。そこで、この単調な性質を利用して二乗誤差比が最小となる値を見逃すことなく、効率的な探索を行うことを考える。この単調性より探索のある時点での次数パラメータの組み合わせに対する二乗誤差と、その時点で更に探索を行わねばならない最大次数パラメータの組み合わせに対する二乗誤差との比は、必ず  $\Delta AIC^*$  を決める二乗誤差比の下界になる。この性質を利用して最適解を効率的に探索可能な A\*探索法 [1] を考える。A\*探索法は下界を評価関数として探索を行うことで、完全探索を行えることが数学的に示されている。図 1 に A\*探索法のアルゴリズムを示す。ここで、 $R=[r_0=p, r_1=q_1, \dots, r_h=q_h]$  であり、 $R_{max}$  は各探索時点の最大次数パラメータ組み合わせ、 $\Delta AIC^*_R$  は  $R$  の  $\Delta AIC^*$ 、 $\Delta AIC^*_{R_{max}}$  は  $\Delta AIC^*_R$  の下界、 $R_{opt}$  は最適次数パラメータ組み合わせ、 $\Delta AIC^*_{R_{opt}}$  は最適次数パラメータ組み合わせの  $\Delta AIC^*$  である。A\*探索はこのアルゴリズムを用いて、最適次数パラメータの組み合わせ  $(p, q_g) (g=1, 2, \dots, h)$  を探索する。

Main

- (1)  $R=[r_0=1, \dots, r_h=1], R_{max}=[r_{max_0}=L, \dots, r_{max_h}=L]$
- (2)  $\Delta AIC^*_R, \Delta AIC^*_{R_{max}}$  を計算
- (3)  $R_{opt} = R, \Delta AIC^*_{R_{opt}} = \Delta AIC^*_R$
- (4) If  $\Delta AIC^*_R \geq \Delta AIC^*_{R_{max}}$  then  
 $[R_{opt}, \Delta AIC^*_{R_{opt}}] = A^*(R, R_{max}, R_{opt}, \Delta AIC^*_{R_{opt}}, 1)$

Function

- (1)  $[R_{opt}, \Delta AIC^*_{R_{opt}}] = A^*(R, R_{max}, R_{opt}, \Delta AIC^*_{R_{opt}}, g_s)$
- (2) for  $g=gs$  to  $h$
- (3)  $r_g = r_g + 1$
- (4)  $\Delta AIC^*_R, \Delta AIC^*_{R_{max}}$  を計算
- (5) If  $\Delta AIC^*_R \leq \Delta AIC^*_{R_{opt}}$  then  
 $R_{opt} = R, \Delta AIC^*_{R_{opt}} = \Delta AIC^*_R$
- (6) If  $(r_g < L) \text{ and } (\Delta AIC^*_R \geq \Delta AIC^*_{R_{max}})$  then  
 $[R_{opt}, \Delta AIC^*_{R_{opt}}] = A^*(R, R_{max}, R_{opt}, \Delta AIC^*_{R_{opt}}, g)$
- (7)  $r_{max_g} = r_g$
- (8) end

図 1: A\*探索のアルゴリズム

#### 4.3 ヒューリスティック探索法の提案

前節で提案した最大次数との二乗誤差比の値は必ず下界であり、完全な探索を行うことができるが、その下界が実際の最小値と比べて小さすぎて、あまり多くの枝狩りを行うことができないことがある。その場合には、あるパラメータの組み合わせと最大次数パラメータの組み合わせとの二乗誤差比の  $n$  乗根を

取って、それを評価関数に置き換えて先ほどと同様に次数パラメータの組み合わせを探索する方法を考える。これにより二乗誤差比が  $[0,1]$  区間の値を取るので  $n$  乗根は必ず二乗誤差比よりも  $[0,1]$  区間で 1 に近づき大きくなることで、実際の最小値に近づく保証はないが多くの枝狩りができ、ある程度小さい  $\Delta AIC^*$  の値を有すると考えられる次数パラメータの組み合わせを短時間で発見できる可能性がある。

### 5. 性能評価実験

この節では、まず 5.1 節で述べる手法検証用人工データを用いて外部入力付自己回帰モデルを作成し、その次数決定手法である AIC と  $\Delta AIC^*$  の性能評価を行う。その後、外部入力付自己回帰モデルと提案する次数決定手法を用いて、解析対象実データの解析を行い、その結果と考察について述べる。なお、本実験においては、データ生成の背景メカニズムが複雑なデータは扱っていないため、 $AIC^*$  を用いて次数パラメータの組み合わせを探索する際の次数パラメータの探索範囲は  $L=9$  とする。

#### 5.1 手法検証用人工データの概要

手法検証用人工データを作成する際、まず、二つの外部入力変数  $y_1(s), y_2(s)$  の時系列データを作成し、さらに次式 (5) を用いて、対象変数  $x(s)$  を計算する。このとき、 $s=1,2,\dots,10000$  とした。

$$x(s) = \sum_{i=1}^p a_i x(s-i) + \sum_{i_2=1}^p a_{i_2} x^2(s-i) + \sum_{j_1=1}^{q_1} b_{j_1} y_1(s-j_1-k_1) + \dots + \sum_{j_h=1}^{q_h} b_{j_h} y_h(s-j_h-k_h) + e(s) \quad (5)$$

ここで、 $y_1(s), y_2(s)$  はステップ応答又は分散 1 の正規乱数とした。次数パラメータの組み合わせを  $p=3, q_1=3, q_2=4$  又は  $p=6, q_1=7, q_2=6$  とし、遅れ時間ステップ数  $k_1=0, k_2=0$  とした。また、モデルの係数  $a_i, b_{j_1}, b_{j_2}$  は *Infinite-duration Impulse Response (IIR)* フィルタの設計手法により計算機を用いて、 $x(s)$  の値が発散することがないようにした。 $a_{i_2}$  は  $a_i$  に比べて十分小さい値とし、 $x(s)$  の値が発散することがないようにした。そこで、与えた次数パラメータの組み合わせが  $p=3, q_1=3, q_2=4$  のときは  $a_{i_2}=a_i/20$  とし、与えた次数パラメータの組み合わせが  $p=6, q_1=7, q_2=6$  のときは  $a_{i_2}=a_i/500$  と決定した。また、非線形性が存在しない人工データを作成するときは、非線形項のモデル係数である  $a_{i_2}$  はすべて 0 とした。

このような条件下で求めた対象変数  $x(s)$  に実際の測定環境に存在するノイズを模擬するため、様々な標準偏差を持つ正規雑音を加えた。ここで加える正規雑音の標準偏差は対象変数  $x(s)$  の  $t\%$  の標準偏差であり、 $t$  には様々な値を用いた。

#### 5.2 手法検証用人工データにおける $\Delta AIC^*$ による次数パラメータ組み合わせ選択性能の評価

手法検証用人工データで外部入力付自己回帰モデルを作成する際、AIC と  $\Delta AIC^*$  を用いて、次数パラメータの組み合わせ  $(p, q_1, q_2)$  を決定した。与えた真の次数パラメータの組み合わせが  $p=3, q_1=3, q_2=4$  のときの結果を表 1 に示し、 $p=6, q_1=7, q_2=6$  のときの結果を表 2 に示す。ここで、対象変数作成方法がステップ、正規雑音と表記したのは、それぞれ  $y_1, y_2$  にステップ応答、正規雑音を用いたことを表す。また、 $a_{i_2}=0$  は対象変数が非線形性を持たないことを表す。表 1、表 2 から、雑音が多い又は非線形項が存在すると AIC により決定した次数パラメータの値が著しく大きくなるのがわかる。ここで、与えた真の次数パラメータの組み合わせを考慮する

表 1: 与えた真の次数パラメータが  $(p, q_1, q_2)=(3,3,4)$  のときに探索した次数パラメータ組み合わせ

対象変数作成方法	$a_{i_2}$	指標	雑音なし	雑音 5 %	雑音 20 %	雑音 50 %
ステップ	0	AIC	(3,3,4)	(9,3,4)	(9,3,3)	(9,3,2)
ステップ	0	AIC*	(1,1,1)	(1,1,1)	(1,1,1)	(1,1,1)
正規雑音	0	AIC	(3,4,5)	(9,8,9)	(9,9,9)	(9,9,9)
正規雑音	0	AIC*	(1,2,3)	(1,2,3)	(1,2,3)	(2,2,3)
正規雑音	$a_i/20$	AIC	(6,8,8)	(3,5,1)	(9,9,9)	(9,9,9)
正規雑音	$a_i/20$	AIC*	(1,2,3)	(1,2,3)	(1,2,3)	(3,2,3)

表 2: 与えた真の次数パラメータが  $(p, q_1, q_2)=(6,7,6)$  のときに探索した次数パラメータ組み合わせ

対象変数作成方法	$a_{i_2}$	指標	雑音なし	雑音 5 %	雑音 20 %	雑音 50 %
正規雑音	0	AIC	(7,8,8)	(9,9,9)	(9,8,9)	(9,9,9)
正規雑音	0	AIC*	(5,7,7)	(5,3,8)	(3,4,4)	(5,5,5)
正規雑音	$a_i/500$	AIC	(8,9,9)	(9,9,9)	(8,8,9)	(9,9,9)
正規雑音	$a_i/500$	AIC*	(6,3,8)	(2,3,9)	(3,4,4)	(5,5,5)

と、AIC を用いて探索した次数パラメータの組み合わせは非線形項がなく、雑音が少ない、つまり誤差がほとんどないときは真に近い値を求められるが、誤差がある状況では複雑になりすぎるのがわかる。それに対して、 $\Delta AIC^*$  を用いて探索した次数パラメータの組み合わせは与えた真の次数パラメータの組み合わせより少し簡単になるものの、次数パラメータの組み合わせが複雑化しすぎるのを防ぎ、誤差がある状況でもある程度、真に近い値を求めることができる。このことから、測定誤差を多く含むデータに対しては  $\Delta AIC^*$  が AIC よりよい指標であることがわかる。

#### 5.3 解析対象実データの概要

解析対象実データの採取方法を説明する。菓子製造会社が 4 月 13 日～5 月 3 日までの 3 週間に亘り、ある特定菓子のテレビ CM を流した。テレビ CM 期間中とその前後 1 ヶ月 (3 月～6 月) において、あるスーパーマーケットチェーンで店舗状況を変え、それぞれの店舗の当該特定菓子の売上金額を毎日算出し、記録した。店舗状況は、販促対象商品 (特定菓子) を特売品設置用棚に置いたときを 1、置かないときを 0 としている。また、CM 量は GRP [4] と呼ばれる広告露出量の単位で測定される。これはテレビ CM が流される時の視聴率を足し上げたものである。解析対象実データにおいて、GRP は  $4/15 \sim 4/26$  まで一日につき、100 程度であった。本実験の解析上で着目した特定菓子が販売されたスーパーマーケットの各店舗は限られた特定地域に存在するため、店舗のどの地域にも同様の GRP 量のテレビ CM が流されたと考える。解析対象実データとして、2 店舗における毎日の売上金額と店舗状況、広告量 (GRP) からなる時系列データが 4 ヶ月分ある。店舗には次の 2 種類がある。一時店舗状況変動型店舗では広告期間中の 1 週間のみ特売品設置用棚に特定菓子を置いて、販売を行った。また、常時店舗状況変動型店舗では広告期間中、常に特売品設置用棚に特定菓子を置いて、販売を行った。

#### 5.4 様々な探索アルゴリズムの計算速度、精度の評価

実データにおける次数パラメータに関する  $AIC^*$  の変化は人工データにおける変化に比較して複雑である。また探索過程は  $AIC^*$  の値に強く依存する。そのため計算時間に関しては実データを用いた評価を行った。そこで、 $\Delta AIC^*$  と考案した様々な探索アルゴリズムを一時店舗状況変動型店舗の解析対象実データに適用して、外部入力付自己回帰モデルを作成する。ここで、解析対象実データ中の日付により、標本時間間隔が一日毎の時系列データとし、売上金額を対象変数とみなし、GRP と店舗状況を外部入力変数とみなす。外部入力付自己回帰モデルを作成する際に、様々な探索アルゴリズムを用いて、次数パラメータの組み合わせを探索した結果を表 3 に示す。この表

3はfor loopを用いた全探索,再起呼び出しを用いた全探索,評価関数に下界を用いて完全探索を行うA\*探索,その評価関数の2~5乗根を評価関数に置き換えて探索を行うヒューリスティック探索のそれぞれについて,探索ノード数,計算時間,探索した回数パラメータ組み合わせを示している.ここで,

表 3: 一時店舗状況変動型店舗での計算時間や探索ノード数

評価関数	探索ノード数	時間	p,q1,q2,k1,k2の値
for loop 処理による全探索	1000	20 秒	1,4,1,7,0 (最適解)
再帰呼び出し処理による全探索	729	99 秒	1,4,1,7,0 (最適解)
A*探索	458	61 秒	1,4,1,7,0 (最適解)
2乗根での探索	231	39 秒	1,4,1,7,0 (最適解)
3乗根での探索	107	19 秒	1,4,1,7,0 (最適解)
4乗根での探索	54	10 秒	1,4,1,7,0 (最適解)
5乗根での探索	42	8 秒	2,6,1,9,0 (非最適解)

単純にfor loopによる全探索法とA\*探索法の計算時間だけを比較すると,プログラム処理系の性質からA\*探索法が時間をより要するという結果が得られる.しかし,探索範囲の上限Lを大きくして探索を行ったところ,L=15でfor loopによる全探索法とA\*探索法の計算時間がどちらも630秒となり,等しくなる.このことから,より回数パラメータ組み合わせが増大する際に探索を行う場合はA\*探索法がfor loopによる全探索法より計算時間が早くなると考えられる.また,A\*探索法はfor loopによる全探索法と比べて,探索ノード数を大幅に減らすことができ,優れていると考えられる.また,ヒューリスティック探索法はA\*探索法より格段に計算時間を短くすることができ,本対象データに対しては下界の2乗根や3乗根,4乗根を評価関数として用いても,全探索を用いて求めた回数パラメータの組み合わせと同じ最適解を導ける.評価関数として,下界の5乗根を用いたときは最適解を導けないが,最適解とヒューリスティック探索法によって決定される回数パラメータ組み合わせは近く,評価関数に下界のベキ乗根を用いるヒューリスティック探索法は優れていると考えられる.

### 5.5 解析対象実データ解析結果の考察

常時店舗状況変動型店舗に,A\*探索法を適用して得た解析結果を考察する.店舗の売り上げ特性を詳しく知るために,得られた外部入力付自己回帰モデルのインパルス応答を求めた.インパルス応答とはある外部入力変数が特定の1日のみ1になるような入力に対して,売上金額(対象変数)が示す応答波形である.例えば,G R Pの変化に対する売上金額の変化を知りたいときは,実際の計測値に変えて,G R Pの値を0,0,...,0,1,0,...,0とし,店舗状況の値を0,0,...,0として,データから得た外部入力付自己回帰モデルに入力することでインパルス応答が得られる.

提案手法を用いて外部入力付自己回帰モデルを作成し,そのインパルス応答を求めた結果を図2に示す.図2の上図は,店舗状況の60日目に単位インパルスを入れたとき,売上金額が60日目以降にどのように変化するかを示す.また,下図は,G R Pの60日目に単位インパルスを入れたとき,売上金額が60日目以降にどのように変化するかを示す.ここで,下図のG R Pに単位インパルスを入れたときの売上金額の変化から,この店舗では,単位G R P当たりの売上金額に対する効果は±10~20円程度であり,その効果が出るまでの遅れ時間は9日程度とわかる.また,店舗状況とG R Pが売上金額に与える効果について考える.店舗状況は当該特定菓子を特売品設置用棚で売る際の外部入力値を1としている.そして,上図の店舗状況に対する売上金額の変動は+1000程度である.また,下図のG R Pに単位インパルスを入れたときの売上金額の変化は±10程度である.このこととG R Pの投入量が一日に100~150程度であることを考慮すると,1日のG R Pによる売上金額の変化はの±1000~1500程度であると評価できる.

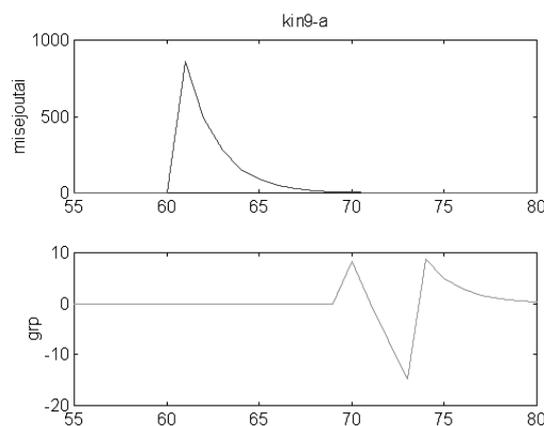


図 2:  $\Delta AIC^*$ による常時店舗状況変動型店舗のインパルス応答

これより,G R Pと店舗状況の変化が売上金額に与える効果は同程度だと考えられる.ここで,G R Pが売上金額に与える効果がマイナスになる部分が見られるが,一般には理解し難い.これは外部入力付自己回帰モデルがそれぞれの外部入力変数の影響を完全に分離できないことがあり,店舗状況の変化による売上金額の変化とG R Pによる売上金額の変化とが影響しあってインパルス応答が変形されたためだと考えられる.ただしこれらより,店舗状況やG R Pが売上金額に与えるおよその定量的効果やそのタイミングがわかる.また,解析対象実データにAICを用いた解析も実施したが,非常に高い回数のパラメータ組み合わせを有する複雑なモデルが得られ,そのインパルス応答も誤差を反映した複雑なものとなり,解釈困難であった.

この結果を利用することで,例えば今回のデータではG R Pが売上金額に与える効果やその遅れ時間が9日後程度であることがわかる.これを考慮して,菓子製造会社や店舗側はG R Pの量やエンド展開のタイミングを計ることで,より当該特定菓子の売上金額を伸ばすことができ,効果的な販売を行うことが可能である.また,様々なマーケティング分野において,このような時系列データは多数存在しているため,外部入力付自己回帰モデルは多大な貢献をもたらすことができると考えられる.

## 6. むすび

本研究において,外部入力付自己回帰モデルの回数パラメータ組み合わせを決定する際, $\Delta AIC^*$ と様々な探索手法を提案した. $\Delta AIC^*$ は従来手法と比べ,誤差のある実際の時系列データにおいて,より妥当な外部入力付自己回帰モデル作成を可能とした.また,様々な探索手法を用いることで効率よく回数パラメータの組み合わせ探索が行えることを示した.新手法により外部入力付自己回帰モデルを作成し,解析を行うことで,外部入力付自己回帰モデルを用いたデータマイニング手法による分析のマーケティング分野への応用可能性を考察した.

## 参考文献

- [1] Stuart Russel, Peter Norvig, 古川 康一, エージェントアプローチ 人工知能, pp, 51-120 (1995).
- [2] 赤池 弘次, 尾崎 統, 北川 源四朗, 尾形 良彦, 石黒 真木夫, 時系列データの解析方法, pp, 1-116 (1998).
- [3] Lennart Ljung, 足立 修一, System Identification TOOLBOX For Use with MATLAB,(1996).
- [4] <http://www.systrat.co.jp/terminology/advertising.html>, 真面目なシストラット流マーケティング用語集