

予測手法

(1)：時系列予測法

上田 徹

はじめに

オイルショック等の社会の急激な変化や消費者・顧客の嗜好のばらつきが増大、技術進歩、それらに伴う多数の新サービスの出現のために、需要予測は非常に難しくなっている。

個々の予測手法は、基本的には過去の何らかの傾向が将来も続くと考えて外挿するものがほとんどであり、外挿の結果がその企業存続を許さないようなものであれば、その企業は何らかの政策変更をするであろう。政策変更自身は予測の責任ではなく、むしろ成果と考えるべきであり、その結果をふまえて新たな予測に臨まなくてはならない。

政策変更を予測担当者が知らなければずっと誤った予測が繰り返されることになる。予測値は予測のみを担当する予測専門家が予測対象の特性も知らずにいろいろな予測法を試行錯誤して得られるものではない。予測担当者は予測値には誤差が含まれ、長期予測ほど誤差は大きくなるとの認識のもとに予測対象の特徴、既存類似サービスの普及過程などを考慮しながら予測を行なわねばならない。予測法には、短期予測に向くものと長期予測に向くものがある。ひとつの方法で短期も長期もと欲張るとどちらも精度の低い予測になってしまう危険性がある。また短期予測法による予測値の外挿値と長期予測値とがいちじるしく異なっている場合にはその原因を追求し、相互修正する必要もでてくるであろう。

また、ロケットの実時間軌道推定（[1]など）を行なう場合と売上高の予測を行なう場合とでは予測モデルのあいまいさや予測期間などに大きな差がある。ここでは売上高や需要量などの経済活動に関連する予測をとりあげる。

経済活動に関連する予測では、予測を予測担当者の専売特許にしておくわけにはいかない。むしろ、新サービス開発者や投資計画策定者なども既存の予測法の特徴を把握している必要がある。そこで

(i) てっとり早い知識習得の手助けとなるよう、

(ii) 経済活動に関連する予測に役立つよう

本講座では

(1) 時系列データのみで行なえる予測法、

(2) 競争状態を考慮できる生態学モデル、

(3) 変数間の相関を利用する統計的方法、

(4) 個人または企業ごとなど個体の選択行動を表わす非集計モデル

の概要、特徴、問題点などを示す。これらの他にもいろいろな予測手法があり、実際の予測では社会変化や経済変動に関する洞察をまじえつつ、いろいろな予測値を比較しつつ、予測値が決定されるはずであり、特定の手法をよく知っているからといってそればかりで予測しては駄目なことは言うまでもない。

1. 時系列予測法

時系列データを分析する方法としては

(1) 指数平滑法[2]、

(2) ARIMAモデル[3]、

(3) カルマンフィルタ[4]、

(4) CENSUS局法 (X-11) [5]、

(5) EPA法 (経済企画庁) [6]、

(6) BAYSEA (統計数理研究所) [7]

などがある。これらの方法では

(I) トレンド、

(II) 循環的変動要素、

(III) 季節成分

の扱い方に差があるが、モデルの柔軟性の観点からカルマンフィルタに焦点を合わせることにする。他の方

うえだ とおる NTT通信網総合研究所

〒180 武蔵野市緑町3-9-11

法はモデルの構造が柔軟でないぶんだけシステム化しやすく、有効な市販システムもある ([8],[9]など)。カルマンフィルタのもつ柔軟性は逆にモデル設定の段階でのプログラム作りを困難にしており、普及の妨げにもなっていると思われるのでモデル設定の段階を中心に紹介する。

そしてBox-Jenkins 流アプローチ[3]で苦勞して適切なモデルとパラメタの選択を行ってきた人々には、その成果をそのままカルマンフィルタに生かせるようARIMAモデルの状態空間表現も示す。

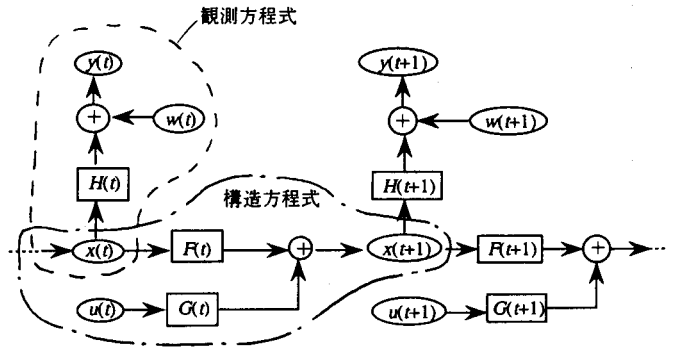


図 1.1 状態空間モデル

1. 1 カルマンフィルタ

(1) 状態空間表現

時点 t でのトレンドや季節性などに関する状態を表わす変数 $x(t)$ を推定、予測したいものとする。時点 t から時点 $(t+1)$ への状態変化は、雑音 $u(t)$ を考慮して構造(システム)方程式

$$x(t+1) = F(t)x(t) + G(t)u(t)$$

で与えられるものとする。時点 t での観測値 $y(t)$ は状態 $x(t)$ と関連づけられるが、雑音 $v(t)$ を含んでいるため、観測方程式

$$y(t) = H(t)x(t) + v(t)$$

で表わされるものとする。このように構造方程式と観測方程式で表現されるモデルを状態空間表現という

(図1.1参照)。このとき、カルマンフィルタは状態 $x(t)$ を逐次的に推定するアルゴリズムである[10](連続時間線形システムに対する推定アルゴリズムは[11]を参照)。

ここで $x(t)$ は直接観測できないものであってもかまわず、観測可能な $y(t)$ を通じて推定される量である。

状態空間表現を行なうときに、まず戸惑うのは状態をどう定義し、 F 、 G 、 H をどのように作るかということである。ここでは経済時系列に典型的な状態空間表現の例を示し、カルマンフィルタを使いたい読者の参考としたい。

時点 n での売上高 $y(n)$ は

$$y(n) = \text{トレンド成分 } T(n) + \text{季節成分 } S(n) + \text{価格効果 } L(n, N_0) + \text{不規則成分 } W(n) \quad (1.1)$$

と表現できるものとする。[14],[15]

(I) トレンド成分

トレンド成分の変化はなめらかであることが予測を

行なう場合、望ましい。この制約を表現するためには、トレンド成分の階差 $|\nabla T(n) = T(n) - T(n-1)|$ あるいは $|\nabla^2 T(n) = T(n) - 2T(n-1) + T(n-2)|$ が小さな値になることが要請される。後者はトレンドを局所的に1次式で近似することであり、

$$\begin{aligned} \nabla^2 T(n) &= \{T(n) - T(n-1)\} - \{T(n-1) - T(n-2)\} \\ &= u(n) \end{aligned} \quad (1.2)$$

と表わす。ただし $u(n)$ [$n=1, 2, \dots, N$]は互いに独立で、平均0、分散 τ^2 の正規性ホワイトノイズと仮定する。分散 τ^2 は未知とする。

トレンド成分 $T(t)$

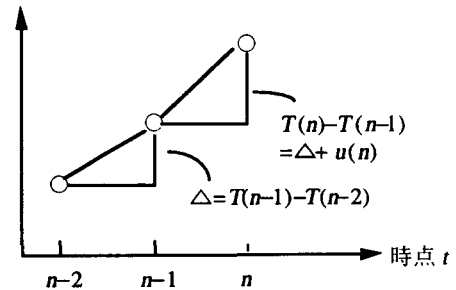


図 1.2 トレンド・モデル

(II) 季節成分

季節成分 $S(n)$ は、月単位のデータであれば周期 $\{s=12\}$ 、四半期データであれば周期 $\{s=4\}$ をもつと考えられるので、 B を後退作用素とすると $S(n)$ は、次式で表わすことができる。

$$(1 - B^s)S(n) = v(n); BS(n) = S(n-1) \quad (1.3)$$

ただし $v(n)$ [$n=1, 2, \dots, N$]は、互いに独立で平均0、分散 σ^2 の正規性ホワイトノイズと仮定する。分散 σ^2 は未知とする。季節成分 $S(n)$ は、式(1.3)のかわりに

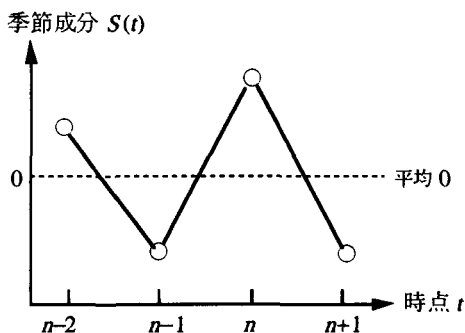


図 1.3 季節成分モデル

次式の移動平均を用いて定義することもできる。

$$\sum_{i=0}^{s-1} S(n-i) = v(n) \quad (1.4)$$

式(1.4)は、式(1.3)の左辺を

$$(1-B^s)S(n) = (1-B)(1+B+\dots+B^{s-1})S(n) \quad (1.5)$$

と変形し、 $(1-B)$ 項を落として季節成分を定義していることになる。式(1.4)は、季節成分を平均0のまわりで周期的な変動をする量としてとらえるためのモデル化の1つである。

(III) 価格改定の影響分

価格改定の時期があらかじめわかっている場合、すなわち時点 N_0 に起きた価格改定の影響 $L(n, N_0)$ は、時点 N_0 に $\{L(n) = L_0\}$ のステップ入力がかわったと考えて定式化する。(図1.4参照)

$$L(n, N_0) = 0 \quad : n < N_0$$

$$L_0 \quad : n \geq N_0$$

この成分は他の政策変更にも用いることができるが、価格の値など市場変数を用いたいときには回帰分析の導入が必要である。[17]

各成分を定めたので、これらを用いて次のように状態空間モデルを構成することができる。

$$x(n+1) = Fx(n) + GU(n) \quad (1.6)$$

$$y(n) = H(n)^t x(n) + W(n) \quad (1.7)$$

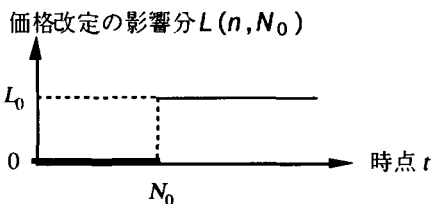


図 1.4 価格成分モデル

$x(n)$ は状態、 $y(n)$ は観測値、 $W(n)$ は観測ノイズ、 $U(n)$ はシステムノイズと呼ばれる量である。システムノイズ $U(n) [n=1, 2, \dots, N]$ は互いに独立で、平均0、分散 Q の正規性ホワイトノイズと仮定される。また初期状態 $x(0)$ は各ノイズと独立な正規分布に従っていると仮定する。

状態としてトレンド成分と式(1.4)による季節成分(ただし $s=4$)および価格改定の影響を用いることとし、次式で定義する。

$$x(n) = (T(n), T(n-1), S(n), S(n-1), S(n-2), L(n))^t$$

このとき、 F, G, H, U は

$$F = \begin{bmatrix} 2-1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$H(n) = (1, 0, 1, 0, \dots, 0, d(n, N_0))^t$$

$$U(n) = (u(n), v(n))^t$$

で与えられる。ただし $u(n), v(n) [n=1, 2, \dots, N]$ は互いに独立で、 $U(n)$ は

$$\text{平均} \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \text{分散} \begin{bmatrix} \tau^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

の正規性ホワイトノイズと仮定する。また

$$d(n, N_0) = 0 \quad : n < N_0$$

$$1 \quad : n \geq N_0$$

である。

これらの F, G, H, U および $x(n)$ を式(1.6)、(1.7)に代入すると、式(1.1)、(1.2)、(1.4)が組み込まれていることは容易にわかる。たとえば式(1.6)の第1要素は

$$T(n+1) = 2T(n) - T(n-1) + u(n)$$

すなわち式(1.2)を表現しており、第4要素は単に

$$S(n) = S(n)$$

と左右両辺に同じ要素を配置したに過ぎない。

(2) 未知パラメータの推定

状態空間モデルの雑音の分散 $\theta = \{\omega^2, \tau^2, \sigma^2\}$ および初期状態 $x(0)$ を推定するために、状態空間モデルの尤度を計算する。 θ および $\{x(0) = x(0|0)\}$ を与えたときのデータ $\{y(1), y(2), \dots, y(N)\}$ の尤度は、

$$L(\theta, x(0)) = \prod_{n=1}^N \{2\pi R(n|n-1)\}^{-1/2} \exp\{-\alpha(n)^2 / 2R(n|n-1)\} \quad (1.8)$$

となる[14]。ここで

$$e(n) = y(n) - y(n|n-1) \quad (1.9)$$

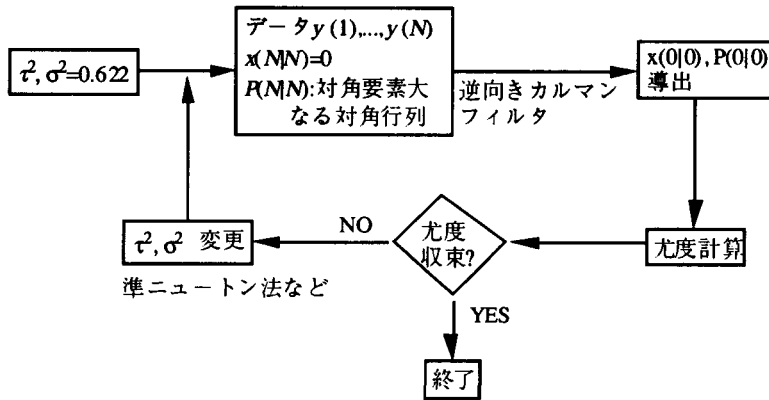


図 1.5 パラメータ推定手順

であり、 $y(n|n-1)$ と $R(n|n-1)$ は、それぞれ $\{y(1), y(2), \dots, y(n-1)\}, x(0)$ および θ を与えたときの $y(n)$ の条件つき期待値とその誤差分散である。

$y(n|n-1)$ と $R(n|n-1)$ は

$$y(n|n-1) = H(n)^t x(n|n-1) \quad (1.10)$$

$$R(n|n-1) = H(n)^t P(n|n-1) H(n) + \omega^2 \quad (1.11)$$

で与えられる。 $x(n|n-1)$ 、 $P(n|n-1)$ は $x(0)$ と θ が与えられれば、以下に示すカルマンフィルタのアルゴリズムから求めることができる。

(時間更新)

$$x(n+1|n) = Fx(n|n) \quad (1.12)$$

$$P(n+1|n) = FP(n|n)F^t + GQG^t \quad (1.13)$$

(観測による修正)

$$x(n|n) = x(n|n-1) + K_n \{y(n) - H(n)^t x(n|n-1)\} \quad (1.14)$$

$$P(n|n) = P(n|n-1) - K_n H(n)^t P(n|n-1) \quad (1.15)$$

$$K_n = \frac{P(n|n-1)H(n)}{H(n)^t P(n|n-1)H(n) + \omega^2} \quad (1.16)$$

データ、 $x(0)$ および θ が与えられれば尤度が計算できるので、 $x(0)$ と θ は尤度関数(1.8)を最大にすることによって決定される。

カルマンフィルタの定常ゲインはシステムノイズの分散と観測ノイズの分散の比によって決まるので、未知分散 θ のうち観測ノイズの分散 ω^2 を1に規格化し、システムノイズの分散 Q を尤度の数値的最大化によって推定する。

初期値

$$x(0) = (T(0), T(-1), S(0), S(-1), \dots, S(-s+2), L(0))^t \quad (1.17)$$

はスムージング値を用いる方法やすべての要素を零とする方法などがあるが、ここでは次のような中間的な方法を紹介する。データ $y(1), y(2), \dots, y(N)$ が利用できるとき、逆向きにカルマンフィルタを適用する。すなわち式(1.12)~(1.16)を次のように逆向きに使う。

(i) $x(N|N)$ はゼロベクトル、 $P(N|N)$ は十分大きな対角成分をもつ対角行列に設定する。

$$(ii) x(n-1|n) = Fx(n|n) \quad (1.18)$$

$$P(n-1|n) = FP(n|n)F^t + GQG^t \quad (1.19)$$

$$x(n-1|n-1) = x(n-1|n) + K_{n-1} \{y(n-1) - H(n-1)^t x(n-1|n)\} \quad (1.20)$$

$$P(n-1|n-1) = P(n-1|n) - K_{n-1} H(n-1)^t P(n-1|n) \quad (1.21)$$

$$K_{n-1} = P(n-1|n)H(n-1) / \{H(n-1)^t P(n-1|n)H(n-1) + \omega^2\} \quad (\text{ただし } \omega^2=1); n=N, N-1, \dots, 1 \quad (1.22)$$

(iii) $n=1$ では $x(0|1)$ の要素

$$(T(0), T(1), S(0), S(1), S(2), L(0))^t$$

および $P(0|0)$ が得られる。

(iv) $x(-m|1) = F^m x(0|1)$; $m=1, 2$ ($=s-2$)

すなわち、その第1、3要素から

$$T(-m) = 2T(-m+1) - T(-m+2)$$

$$S(-m) = -S(-m+1) - S(-m+2) - \dots - S(-m+s-2)$$

として、 $T(-1), S(-1), S(-2), \dots, S(-s+2)$ を求めることにより式(1.17)の全要素を特定化できる。

これらのプロセスをまとめると図1.5のようになる。

(3) 長期予測

カルマンフィルタによる予測は、式(1.12)~(1.16)か

ら明らかのように、時点 n までのデータを用いて1期先の状態を予測するアルゴリズムである。 m ($m > 1$ なる整数)期先の予測を行なうことを長期予測と呼び、時点 n までのデータを与えたときの状態 $x(n+m)$ の条件付期待値を $x(n+m|n)$ 、その誤差分散を $P(n+m|n)$ で表わす。

$$x(n+m|n), P(n+m|n) \text{は次式で与えられる.}$$

$$x(n+m|n) = F^m x(n|n) \quad (1.23)$$

$$P(n+m|n) = F^m P(n|n) (F^t)^m + \sum_{i=0}^{m-1} F^i G Q G^t (F^t)^i \quad (1.24)$$

m 期先の最適な予測を行なうのに必要なパラメータ推定法については[14],[16]などを参照してもらいたい。

1. 2 ARIMAモデル

(1) 基本モデル

等間隔で観測された過去の時系列データ y_1, y_2, \dots, y_n を用いて将来の値 y_{n+i} ($i \geq 1$)を予測する問題に対して、BoxとJenkinsは自己回帰和分移動平均ARIMA (Auto Regressive Integrated Moving Average) モデル、すなわち

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) (1 - B)^d y_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) a_t \quad (1.25)$$

あるいは

$$\Phi(B) (1 - B)^d y_t = \Theta(B) a_t \quad (1.26)$$

で表わされるモデルについて論じた[3]。ここで B は、式(1.3)と同じであり、 $\Phi(B)$ 、 $\Theta(B)$ は B に関する多項式である。式(1.25)は、時間とともにたどるなだらかな経路(トレンド成分)を除去するために、差分

$$(1 - B)^d y_t = z_t$$

を採っている。また a_t は平均0、一定分散をもつ確率変数列で、 a_t と a_s ($t \neq s$)は無相関である。 a_t には通常、正規性が仮定される。式(1.25)で表わされるモデルはARIMA(p, d, q)モデルと表示され、 $(1 - B)^d$ を含まないモデルはARMA(p, q)モデル、 $\{\Phi(B) y_t = a_t\}$ は自己回帰AR(p)モデル、 $\{y_t = \Theta(B) a_t\}$ は移動平均MA(q)モデルと呼ばれる。

パラメータ ϕ_i, θ_j の推定法については最尤推定量や最小自乗推定法などがあげられるが、それらを用いてパラメータおよび p, d, q は求められているものとする。ARIMAモデルも状態空間表現することができ、カルマンフィルタのアルゴリズムを適用することができる。

ここでは、まずARMA(p, q)モデルを状態空間表

現した後、それを用いてARIMA(p, d, q)モデルの状態空間表現を行なう。

(2) ARMA(p, q)モデルの状態空間表現[17]

1変量時系列 z_t がARMA(p, q)過程に従うとすると

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (1.27)$$

これは、次のように書き直すことができる。

$$z_t = \phi_1 z_{t-1} + \dots + \phi_m z_{t-m} + a_t + \theta_1 a_{t-1} + \dots + \theta_{m-1} a_{t-m+1} \quad (1.28)$$

ただし、

$$m = \max(p, q+1)$$

$$j > p \text{ ならば } \phi_j = 0, \quad j > q \text{ ならば } \theta_j = 0.$$

上式の状態空間表現を作るには、状態を

$$x_t \triangleq (z_t, x_{1t}, \dots, x_{m-1,t})^t$$

と定義し、

$$x_t = F x_{t-1} + G a_t \quad (1.29)$$

$$z_t = H x_t \quad (1.30)$$

とすればよい。ただし、

$$F \triangleq \begin{bmatrix} \phi_1 & 1 & & 0 \\ \phi_2 & & 1 & \\ \vdots & & & \ddots \\ \phi_m & 0 & \dots & 0 \end{bmatrix}, \quad G \triangleq \begin{bmatrix} 1 \\ \theta_1 \\ \vdots \\ \theta_{m-1} \end{bmatrix}$$

$$H \triangleq (1, 0, \dots, 0)$$

である。式(1.29)を書き下すと、

$$z_t = \phi_1 z_{t-1} + x_{1,t-1} + a_t$$

$$x_{1,t} = \phi_2 z_{t-1} + x_{2,t-1} + \theta_1 a_t$$

$$\vdots$$

$$\vdots$$

$$x_{m-2,t} = \phi_{m-1} z_{t-1} + x_{m-1,t-1} + \theta_{m-2} a_t$$

$$x_{m-1,t} = \phi_m z_{t-1} + \theta_{m-1} a_t$$

となるから、下から上への代入を繰り返すことにより、もとのARMAモデルを作ることができる。

(3) ARIMA(p, d, q)の状態空間表現[18]

ARIMA(p, d, q)モデルでは、観測値 y_t の d 階差分がARMA(p, q)に従うことから、式(1.29),(1.30)において、 $z_t = (1 - B)^d y_t$ として考えればよい。すなわち、

$$x_t \triangleq ((1 - B)^d y_t, x_{1t}, \dots, x_{m-1,t})^t$$

と定義すれば

$$x_t = F x_{t-1} + G a_t \quad (1.31)$$

$$(1 - B)^d y_t = H x_t \quad (1.32)$$

である。しかし、予測は $(1-B)^d y_t$ ではなく、 y_t について行なうので状態空間表現を修正する必要がある。

$$y_t = (1-B)^d y_t + \sum_{j=1}^d \binom{d}{j} (-1)^{j+1} y_{t-j} \quad (1.33)$$

ただし $\binom{d}{j}$ は二項係数

だから、

$$y_t \triangleq (y_t, y_{t-1}, \dots, y_{t-d+1})^t$$

と定義し、

$$A \triangleq (1, 0, \dots, 0)^t$$

$$D \triangleq \begin{bmatrix} \binom{d}{1} & -\binom{d}{2} & \dots & (-1)^{d+1} \binom{d}{d} \\ 1 & & & 0 \\ & 1 & & 0 \\ & & \ddots & \vdots \\ 0 & & & 1 & 0 \end{bmatrix}$$

とおくと、

$$y_t = A z_t + D y_{t-1} = A H x_t + D y_{t-1}$$

となる。そこで、新たに状態

$$T_t \triangleq (x_t, y_{t-1})^t$$

を定義すると、

$$T_{t+1} = \begin{bmatrix} F & 0 \\ AH & D \end{bmatrix} T_t + \begin{bmatrix} G \\ 0 \end{bmatrix} a_t \quad (1.34)$$

という新しいシステム方程式を作ることができる。観測方程式は式(1.33)より、

$$y_t = H x_t + \sum_{j=1}^d \binom{d}{j} (-1)^{j+1} y_{t-j} \\ = \left[H, \binom{d}{1}, -\binom{d}{2}, \dots, (-1)^{j+1} \binom{d}{j}, \dots, (-1)^{d+1} \binom{d}{d} \right] T_t \quad (1.34)$$

となる。

式(1.34),(1.35)は、 y_t がARIMA(p,d,q)に従う場合の状態空間表現になっており、これにカルマンフィルタのアルゴリズムを適用して予測を行なうことができる。

本資料をまとめるにあたってはNTT研究所 矢田健氏の助言を得た。ここに謝意を表わす。

参考文献

[1] 野村, 石谷, 馬場, 前田: 「科学観測衛星打上げにおける電波誘導方式」, 計測と制御, Vol.14, No.11, pp.836-847, 1975.

[2] L.A. Johnson and D.C. Montgomery: "Forecasting with Exponential Smoothing and Related Methods", TIMS Studies in the Management Science Vol.12, Forecasting, pp.31-44, 1979.

[3] G.E.P.Box and G.M. Jenkins: "Time Series Analysis: Forecasting and Control", Revised ed., Holden-Day, 1976.

[4] 有本卓: 「カルマン・フィルター」, 産業図書, 1977.

[5] Bureau of the Census: "The X-11 Variant of the Census Method II Seasonal Adjustment Program", Technical Paper No. 15, 1965.

[6] 阿部喜三他: 「季節変動調整法」, 経済企画庁 経済研究所 研究シリーズ第22号, 1971.

[7] 石黒真木夫: 「ベイズ型季節調整法」, 数理科学, No.213, pp.57-61, 1981.

[8] 社会情報サービス: 「マルチ時系列」(EPA法)

[9] メタテクノ: 「時」(ARIMAモデル)

[10] R.E. Kalman: "A New Approach to Linear Filtering and Prediction Problems", Trans, ASME, Series D, Journal of Basic Eng., Vol.82, pp.35-45, 1960.

[11] R.E. Kalman and R.S. Bucy: "New Results in Linear Filtering and Prediction Theory", op. cit., Vol.83, pp.95-107, 1961.

[12] H.J. Kushner: "Dynamical Equations for Optimal Nonlinear Filtering", Journal of Differential Equations, Vol.3, pp.179-190, 1967.

[13] 樫木義一, 砂原善文: 「制御における推定理論」, 計測と制御, 第8巻, pp.537-549, 1969.

[14] 阿部威郎, 上田徹: 「状態空間モデルを用いた電話収入予測」, 信学論, J68-A, 5, pp.437-443, 1985.

[15] H. Akaike: "Seasonal Adjustment by a Bayesian Modeling", Journal of Time Series Analysis, Vol.1, No.1, pp.1-13, 1980.

[16] W. Gersch and G. Kitagawa: "The prediction of time series with trends and seasonalities", Journal of Business & Economic Statistics, Vol.1, No.3, pp.253-264, 1983.

[17] A.C. Harvey: "Time Series Models", Philip Allan Publishers Limited, Oxford, 1981. 国友, 山本訳, 「時系列モデル入門」, 東京大学出版会, 1985.

[18] P.J. Brockwell and R. A. Davis: "Time Series: Theory and Methods", second ed., Springer-Verlag, 1991.