

時系列予測について: カルマン・モデルの応用

メタデータ	言語: jpn 出版者: 公開日: 2017-10-03 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	http://hdl.handle.net/2297/37177

時系列予測について

— カルマン・モデルの応用 —

平 館 道 子

1. ま え が き

時系列分析における R. E. カルマンのモデルは工学分野の予測理論の発展と応用に大きな貢献をしたが、経済時系列への応用はこれまであまり見られなかった。それは、カルマン・モデルにおけるインプットが工学分野の問題では制御可能な確率変量（白色雑音）であるのに対して、経済分野の場合には実際上その性格を客観的には明確にすることが難しいからであろうと思われる。しかし、インプットが性格の規定された白色雑音であるからこそ、カルマン・フィルターの単純さという特徴が生かされるのであり、もしその特性を統計的に追求しなければならぬとしたら、利点の多くが失われてしまうことになる。そういった意味から伝統的な方法によるアプローチには難点が伴うと思われる。Harrison-Stevens はベイジアン観点からアプローチを試み、応用への道を拓くのに成功している〔2〕。本稿では彼等のアプローチをもとに、具体的な応用例によってカルマン・モデルの検討を試みた。

2. カルマン時系列モデル

通信における信号と雑音の統計的処理の問題、いわゆるフィルタリング問題について、これを最初に定式化したのは N. ウィーナーであり、彼の統計的予測の理論が時系列分析に大きな貢献をしたことはよく知られている〔5〕。しかし彼の定式化から最適フィルターを求める一般の解法が発見されず、実際への応用があまり見られなかったが、1960年に R. E. カルマンが線形ダイナミックシステムにもとづく定式化を行うことによって、ウィーナーを乗り越える程の拡張が可能となり、実際的な応用への道が拓かれた。これらの詳細については、例えば〔1〕に述べられている。カルマンは連続時間の場合と離散時間の場合の両者を定式化しているが、本稿で応用するのは後者の場合であるから、離散時間のカルマン・モデルについて、結果だけを述べておきたい。詳細に関しては R. E. カルマン〔4〕を参照されたい。

一つの確率過程からその値が生成される変数 X_t があるとしよう。もしこれ

らの値が観測可能ならば、ある期間の観測値を得ることによって、その将来の値がどのようなものであるかについて確率的に述べたい。これは X の観測値を得て、 X の将来の値の条件つき確率分布を求めることである。しかし、もし X_t が観測不可能である場合には、それと関連した確率過程から生じる観測可能な他の変数 Y_t を観測することによって X_t を推定したい。これは Y の値が与えられた時の X_t の条件つき確率分布を求めることによって行われる。例えば X は信号であり、 Y は信号と雑音とがまじったものである。この場合には Y を観測して、これから雑音を取り除き、信号 X を取り出すことが問題なのである。これをフィルタリング問題という。フィルタリング問題を一般的に解くことはできないから、何らかの制約をおくことが必要であるが、カルマンは確率過程がマルコフ性をもつことを仮定し、 X_t の線形推定量を求めるという制約において、次の様な線形システムモデルを定式化した。

$$X_t = G_t X_{t-1} + H_t w_t \dots \dots (1)$$

$$Y_t = F_t X_t + v_t \dots \dots \dots (2)$$

ここに、 t は時間を表わす整数であり、 X は n 次元ベクトルで、システムの状態とよばれる。 w は m 次元ベクトルで、システムのインプットであり、制御可能な場合も不可能な場合もある。工学分野では前者の場合制御関数とよぶが、経済時系列への応用を考える場合には一般に制御することは不可能であるから w は確率変量である。 Y は p 次元ベクトルでシステムのアウトプットであり、観測可能である。 v は p 次元ベクトルで、確率変量である。 G, H, F はそれぞれ $n \times n, n \times m, p \times n$ 行列であり、これらがすべて定数行列である場合（時間 t と共に変化しない場合）、システムはコンスタントであるという。

式 (1) は時点 t におけるシステムの状態が 1 期前の状態 X_{t-1} と w_t とから生成される過程をあらわしており、式 (2) は Y_t が観測される過程を示している。したがって v は観測上の雑音あるいは誤差と考えることができる。

次に w と v に関する仮定を述べよう。 w_t は正規分布に従う m 次元確率変量で、平均を \bar{w}_t 、自己共分散を $E[(w_t - \bar{w}_t)(w_t - \bar{w}_t)'] = W_t$ とする。必ずしも必要ではないが以下では $\bar{w}_t = 0$ であるとする。また異なる時点での w は互いに独立である。同様に v_t は平均が 0 で、自己共分散 V_t の正規変量で、異なる時点での v はやはり互いに独立であると仮定しよう。これは正規白色雑音とよばれるものである。また v_t と w_t とは同一時点で共分散をもつものが一般的であるが、ここでは、同一時点においても、異なる時点においても独立

であると仮定する。最後に両者は X_{t-1} と独立であると仮定しよう。

初期の状態 X_0 が w とは独立な正規変量であると仮定すれば、(1) によって定義される X_t はマルコフ性をもつことを示すことができる。すなわち、 X_0 が与えられた時の X_t の条件つき確率分布は、 X_0 とそれ以前の X が与えられた時のそれと等しい。

X_t の最良線形推定量を求める問題は、線形フィルタリング問題とよばれ、この最適線形推定量を \hat{X}_t とかくと、これは観測値の線形関数で平均2乗誤差 $E\{|X_t - \hat{X}_t|^2\}$ を最小にするものである。最良線形推定量を求めるためには、その平均値と共分散のみが必要であるが、正規分布は平均と共分散とによってユニークに定められるから、正規性によるフィルタリング問題の解は一般の線形フィルタリング問題の解と同じである。そして正規性の場合、平均2乗誤差を最小にするのは観測値が与えられた時の条件つき分布の平均であり、これは最も新しい期の観測値の線形関数である。

一般に τ 期 ($\tau \leq t$) までの観測値が与えられた時 X_t を推定することが問題となるが、以下では $\tau = t$ として Y の t 期までの観測値が得られている場合の推定を考えることにしよう。

いま τ 期までの観測値の系列を $Y^\tau = (Y_1, \dots, Y_\tau)$ とあらわすと、最適フィルタ問題には、 $t-1$ 期までの観測値 Y^{t-1} が与えられた時の X_t の条件つき分布と、 Y^t が与えられた時の X_t の条件つき分布とが関連し、上述の仮定の下ではいずれも正規分布であることがわかる。従って平均と共分散行列が求まればよいから、前者のそれらを \tilde{X}_t, R_t 、後者のそれらを \hat{X}_t, Σ_t とおくことにしよう。この時 \tilde{X}_t, R_t は次の通りである。

$$\tilde{X}_t = G_t \hat{X}_{t-1} \dots \dots \dots (3)$$

$$R_t = G_t \Sigma_{t-1} G_t' + H_t W_t H_t' \dots \dots \dots (4)$$

さらに Y^{t-1} が与えられた時の Y_t の条件つき分布が必要であり、これを求めると正規分布になり、平均 \tilde{Y}_t と共分散行列 V_e とはそれぞれ次の様になる。

$$\tilde{Y}_t = F_t G_t \hat{X}_{t-1}, \dots \dots \dots (5)$$

$$V_e = F_t R_t F_t' + V_t \dots \dots \dots (6)$$

\hat{X}_t, Σ_t は $N(\tilde{X}_t, R_t)$ を事前分布としてベイズの定理から上の諸量を用いて次の様に求められる。

$$\hat{X}_t = \tilde{X}_t + A_t (Y_t - \tilde{Y}_t) \dots \dots \dots (7)$$

$$\Sigma_t = R_t - A_t V_e A_t' \dots \dots \dots (8)$$

ここに

$$A_t = R_t F_t' (V_t)^{-1} \dots \dots \dots (9)$$

である。 \hat{X}_t は X_t の最良線形推定量であり、カルマン・フィルタとよばれている。なお \hat{X}_t は Y_t に依存しているが、 Σ_t は依存していないことに注意されたい。この様にカルマン・モデルによれば、時点 t の観測値と 1 期前の \hat{X}_{t-1} 、 Σ_{t-1} によって次々に X_t の推定を行うことができる。

以上カルマンの結果を概略だけ述べたが、経済時系列についても同様な線形システムがモデルとして妥当と考えられる場合が多い。例えばよく用いられる線形回帰モデルでは回帰係数が時間的に変化しないが、これは式 (2) において F_t を説明変数、 X_t を回帰係数とすると、 $X_t = X_{t-1}$ の場合と考えることができる。分析の期間が長くて X_t が変化すると考えられる場合はダイナミックな回帰モデルということができるだろう。

ところで通信問題の場合、信号と雑音とが交ったものを観測し、これから雑音をろ過して信号をとり出すことが課題であった。したがって専ら X_t の推定が興味を中心となっているが、経済時系列の場合は、回帰モデルの様に X_t の推定と共に Y_t の将来の値の予測という点に興味を持たれるだろう。 Y_t の予測に際して問題になるのは、 Y^{t-k} が与えられた時の条件つき分布であるが、 $k=1$ の場合、これは既に上で求められており、式(5),(6)をそれぞれ平均、共分散とする正規分布であることがわかっている。この様にシステムのアウトプット Y_t そのものに焦点をあてる場合、モデル(1),(2)は Y_t をパラメーター X_t によって表現しているものと解釈することができる。(1)はパラメーターが時間の経過と共に変化する過程を示している。以下では通常の記法に合せて X_t を θ_t と書くことにしよう。 v_t については、観測上の誤差である場合もあるが、また Y_t の変動のうち θ_t に依らない変動を表現していると解釈することもできる。次節で述べるように本稿では後者の解釈に従って、 v_t が通常の程度の変動を示す場合と、異常に大きな変動を示す場合とを区別した。次に w_t はパラメーター θ_t の変動要因であるが、 θ_t は必ずしも変動するとは限らないので、 w_t の共分散 W_t によって、パラメーターが変動する場合としない場合とを表現することができる。この様に W_t と v_t の共分散 V_t とによって、いくつかの意味のある場合を表現することが可能となる。

ここで改めて次の様にモデルを書き直しておこう。

$$\theta_t = G_t \theta_{t-1} + H_t w_t \dots \dots \dots (1)'$$

$$Y_t = F_t \theta_t + v_t \dots \dots \dots (2)'$$

w_t は $N(O, W_t)$ に、 v_t は $N(O, V_t)$ に従い、 W_t と V_t は時点 t にお

いて既知であるとする。次元および他の仮定は上述と同じである。 θ はゼロベクトルである。また観測値の系列が与えられた時の条件つき分布とは、事後分布にはかならない。以下では、 Y の予測は 1 期後について行ない、 θ の推定はその期まで観測値が得られる場合を考えよう。

最初の観測に先立って θ_i の初期値 θ_0 をカルマンは 0 とすべきだとしているが、ここでは θ_0 の事前分布として $N(m_0, C_0)$ が付与されるものとしよう。この時 θ_1 の事前分布は (1)' より、

$$N(G_1 m_0, G_1 C_0 G_1' + H_1 W_1 H_1')$$

となる。

$$G_1 C_0 G_1' + H_1 W_1 H_1' = R_1$$

とおくと、上で述べたカルマンの結果から、 θ_1 の Y_1 が観測された後の事後分布は、 $N(m_1, C_1)$ となり、 m_1, C_1 は次の様になる。

$$m_1 = G_1 m_0 + A(Y_1 - \tilde{Y}_1)$$

$$C_1 = R_1 - A V_e A'$$

ここに

$$\tilde{Y}_1 = F_1 G_1 m_0$$

$$V_e = F_1 R_1 F_1' + V_1$$

$$A = R_1 F_1' (V_e)^{-1}$$

である。

Y_1 の予測に関しては、事前分布が付与された時点で問題になるのであるが、 $N(m_0, C_0)$ が与えられた時、 Y_1 の分布は $N(\tilde{Y}_1, V_e)$ となるから、最小平均 2 乗誤差を与える予測値として平均 \tilde{Y}_1 をとると、 V_e はその予測誤差を示すことになる。

同様に $t-1$ 期までの観測値 Y^{t-1} が得られ、 (m_{t-1}, C_{t-1}) が求められた時、 θ_t の事前分布は $N(G_t m_{t-1}, R_t)$ であり、 Y_t の分布は $N(F_t G_t m_{t-1}, F_t R_t F_t' + V_t)$ である。ここに $R_t = G_t C_{t-1} G_t' + H_t W_t H_t'$ である。従って Y_t の予測は後者の分布にもとづいて行い、予測値は

$$\tilde{Y}_t = F_t G_t m_{t-1},$$

その予測誤差は、

$$V_e = F_t R_t F_t' + V_t$$

である。

次に Y_t が観測された後の θ_t の事後分布は $N(m_t, C_t)$ であり、

$$m_t = G_t m_{t-1} + A(Y_t - \tilde{Y}_t),$$

$$C_t = R_t - A V_e A'$$

$$A = R_i F_i' (V_i)^{-1}$$

である。 θ_i の推定値は m_i であり、 C_i は推定の誤差を示している。

(m_i, C_i) を M_i とかくことにすると、この様に (M_{i-1}, V_i, W_i) と Y_i とから M_i が得られ、また新たな観測値を得る毎に次の期 (M_{i+1}) へと進むことができる。 θ_i の事前分布は正確には上に述べたように $N(Gm_{i-1}, R_i)$ であるが、この様に M_i を求める際のインプットは (M_{i-1}, V_i, W_i) であるから、以下では $N(M_{i-1})$ を θ_i の事前分布とよぶことにする。

3. いくつかの状態があるモデル

経済時系列にカルマンの結果を応用する場合、 W と V とがそれぞれの時点で既知であるか、あるいは制御できるという様な状況はほとんどないであろう。また常に一定であるということもありそうもない。これらの問題への一つの現実的な対応として、 W と V がそれぞれの時点で幾通りかある定められた値をとる可能性があるものと想定しよう。これらを $(W_i, V_i), i = 1, \dots, h$ として、モデル (1)', (2)' と共に (W_i, V_i) が成立する時、モデルは状態 i にあるということにしよう。これは例えば $V = 0$ であれば、 Y_i は全くパラメーター θ にのみ依存してそれ以外の要因による変動がない状態をあらわし、 V が大きな値をとればパラメーター以外の変動要因が大きく働く状態をあらわす。 W についても同様に意味のある値を考えることが可能であろう。モデルは各時点において状態 1 から h までのいずれかにあるが、どれであるか不明で、やはり観測データにもとづいて確率的に判断されるものとし、 t 期に状態 j にあるこの確率を P_t^j としよう。前節の結果からわかる様に、 t 期のカルマン・フィルターは $t-1$ 期の M_{t-1} と t 期の (W, V) とから成っているから、 $t-1$ 期の状態と t 期の状態とに依存する。前者を i 、後者を j としよう。まず $t=0$ においてそれぞれの状態 i に対して $M_0^i = (m_0^i, C_0^i)$, P_0^i , $i = 1, \dots, h$ が付与されたとすると、事前確率分布は

$$\sum_{i=1}^h P_0^i N(M_0^i)$$

となる。これは、 t 期の状態を S_t とあらわすと

$$P(\theta_0) = \sum_{i=1}^h P(S_0=i) P(\theta_0 | S_0=i)$$

であり、右辺の後の項が $N(M_0^i)$ であることから求められる。次に $t=0$ において状態 i にあり、 $t=1$ 期に状態 j になるとすれば、前に述べた結果を

適用して Y_1 が観測された後には, (M_0^i, V_j, W_j) から $M_1^{ij} = (m_1^{ij}, C_1^{ij})$ $i=1, \dots, h, j=1, \dots, h$ が求められる。

θ_1 の事後確率は,

$$P(\theta_1 | Y^1) = \sum_{i,j} P(\theta_1, S_0=i, S_1=j | Y^1) \\ = \sum_{i,j} P(\theta_1 | Y^1, S_0=i, S_1=j) P(S_0=i, S_1=j | Y^1)$$

であり, 最後の式の前項は $N(M_1^{ij})$ であるから, 後の項を P_1^{ij} とすれば,

$$\sum_{i,j} P_1^{ij} N(M_1^{ij}) \dots \dots \dots (10)$$

となる。 P_1^{ij} は $Y^1 = (Y_1)$ であることに注意すれば,

$$P_1^{ij} = P(S_0=i, S_1=j | Y^1) \\ = P(Y_1 | S_0=i, S_1=j) P(S_1=j | S_0=i) P(S_0=i) / P(Y_1)$$

である。ここで $P(Y_1 | S_0=i, S_1=j)$ は前と同様にして $N(\tilde{Y}_1^{ij}, V_e^{ij})$ であることが求められる。ここに

$$\tilde{Y}_1^{ij} = F_1 G_1 m_0^{(i)} \\ V_e^{ij} = F_1 R_1^{ij} F_1' + V_j \\ R_1^{ij} = G_1 C_0^i G_1' + H_1 W_j H_1'$$

である。 \tilde{Y}_1^{ij} は i にのみ依存するから \tilde{Y}_1^i とすることにしよう。 $P(S_0=i) = P_0^i$ であるから, $P(S_1=j | S_0=i) = \pi_{ij}$ とすれば

$$P_1^{ij} \propto |V_e^{ij}|^{-1/2} \exp\left\{-\frac{1}{2} (Y_1 - \tilde{Y}_1^i)' (V_e^{ij})^{-1} (Y_1 - \tilde{Y}_1^i)\right\} \pi_{ij} P_0^i$$

となる。なお π_{ij} は時間的に変動せず, 一定とした。

次に $t=2$ に進む時, (i, j) の組は h^2 通りあり, 時間が進むにつれて (i, j, k, \dots) の様に状態の組の数が増大してしまい, 計算が困難になるから, 事後分布 (10) を何らかの形で縮約しなくては現実的でない。ここでは Harrison-Stevens [2] に従って $t=1$ の状態が同じところでまとめて, 近似的な事後分布を求めることにする。すなわち,

$$P_1^j = \sum_{i=1}^h P_1^{ij}, \\ m_1^j = \sum_i P_1^{ij} m_1^{ij} / P_1^j \\ C_1^j = \sum_i P_1^{ij} \{C_1^{ij} + (m_1^{ij} - m_1^j) (m_1^{ij} - m_1^j)'\} / P_1^j$$

とすると, 近似的な事後分布は

$$\sum_{j=1}^h P_1^j N (M_1^j)$$

となって、また同じ形になる。この様に近似すれば、各期のインプットは同じ形に戻り、 $t=2$ においても同じステップを経ればよいことになる。

予測と推定の手順は一般には次の様になる。

i) Y_t の予測: (M_{t-1}^i, W_j, V_j) , P_{t-1}^i から次の様に求める。

$$\tilde{Y}_t^i = F_t G_i m_{t-1}^i$$

$$V_e^{ij} = F_t R_t^{ij} F_t' + V_j$$

$$R_t^{ij} = G_i C_{t-1}^i G_i' + H_i W_j H_i'$$

$$\tilde{Y}_t = \sum_{i=1}^h P_{t-1}^i \tilde{Y}_t^i$$

$$V_e = \sum_{i,j=1}^h V_e^{ij} P_{t-1}^i \pi_{ij}.$$

ii) M_t^{ij} を求める。

$$m_t^{ij} = G_i m_{t-1}^i + A^{ij} (Y_t - \tilde{Y}_t^i)$$

$$C_t^{ij} = R_t^{ij} - A^{ij} (V_e^{ij})^{-1} A^{ij'}$$

$$A^{ij} = R_t^{ij} F_t' (V_e^{ij})^{-1}.$$

iii) P_t^{ij} , $i, j = 1, \dots, h$ を求める。

$$P_t^{ij} \propto |V_e^{ij}|^{-1/2} \exp\left\{-\frac{1}{2} (Y_t - \tilde{Y}_t^i)' (V_e^{ij})^{-1} (Y_t - \tilde{Y}_t^i)\right\} \pi_{ij} P_{t-1}^i.$$

iv) 事後分布を縮約する。

$$P_t^j = \sum_{i=1}^h P_t^{ij} \quad j = 1, \dots, h$$

$$m_t^j = \sum_{i=1}^h P_t^{ij} m_t^{ij} / P_t^j$$

$$C_t^j = \sum_{i=1}^h P_t^{ij} \{C_t^{ij} + (m_t^{ij} - m_t^j) (m_t^{ij} - m_t^j)'\} / P_t^j.$$

v) θ_t の推定: $m_t = \sum_{j=1}^h m_t^j P_t^j$

$$C_t \doteq \sum_{j=1}^h C_t^j P_t^j.$$

なお C_t は近似的な値である。

4. 応用モデルによる予測

前節までの結果を対ドル為替レートについて応用し、若干の検討を行ってみた。^注

為替価格に影響を与える主要な要因は、関係する国々の金融政策と物価上昇率およびそれらの将来の予想であろうと思われる。また先物価格には当然金利が関係するであろう。これらの要因は当然関連しているものであるが、大雑把に云って、為替価格の基本的な動向は両国の通貨の需要供給政策の差異によって影響をうけ、その基本的動向からの調整が物価上昇率、あるいは金利の差異にもとづいて行われ、更に先物は将来の予想にもとづいてディスカウントされたり、プレミアムがつけられたりするものと思われる。これらの事はモデルの論拠というよりは、単に想定にすぎないかもしれないが、これらの考え方から暫定的に、モデルを次の様に構成した。ここで扱う先物は1ヶ月先物である。

$Y_t = (Y_t^0, Y_t^1)'$, $\theta_t = (\mu_t, \beta_t, \gamma_t)'$ とする。ここに Y_t^0 は直物価格, Y_t^1 は先物価格, μ_t はトレンド, β_t は調整, γ_t はディスカウントあるいはプレミアムをあらわす。 β_t, γ_t は単独であるいは共に価格を上向かせたり下向かせたりするから、勾配要因と云うことができよう。このとき、

$$Y_t^0 = \mu_t + \beta_t + v_1$$

$$Y_t^1 = \mu_t + \beta_t + \gamma_t + v_2$$

$$\mu_t = \mu_{t-1} + w_1$$

$$\beta_t = \beta_{t-1} + w_2$$

$$\gamma_t = \gamma_{t-1} + w_3.$$

$$W = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix}, \quad V = \begin{pmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{pmatrix}$$

である。したがって F, G, H_t は定数行列で、 G, H_t は3次元単位行列である。なお上述の考え方からみると、対数線形モデルの方が適当であろうと思われるが、まだ思わしい結果を得ていない。

パラメーターの推定は興味あるものであり、その実質的な意味との関連で推定結果を検討しなければならないのであるが、筆者の知識からみて、それには

注 筆者は為替相場について経験も知識も極めて不足しており、またその理論を十分検討したとも云えないので、以下で述べるモデルは全く暫定的なものであり、目的は為替価格そのものの分析ではなく、カルマン・フィルターの有用性の検討であることをお断りしておきたい。しかしそれでも実際に計算してみた結果、いくつかの示唆が得られたと思う。

まだ無理があるので、ここでは、専ら予測を通じて検討を行い、どの程度よい予測が得られるか、インプットおよび初期値が予測にどのような影響を与えるか、また各期における状態の確率 P_i は観測値の変動とくらべて妥当なものであるか、等についてみてみたい。

このモデルのインプットは π, V_i, W_i , 初期値 $m_0^i, C_0^i, P_0^i, i=1, \dots, h$ である。アウトプットは Y_i の事後分布の平均と共分散行列 $\tilde{Y}_i, V_e, \theta_i$ の事後分布の平均と共分散行列 m_i, C_i , および状態の確率 $P_i^i, i=1, \dots, h$ である。

状態は次の4つとした。

状態1：無変化

状態2：トレンド変化

状態3：勾配変化

状態4：一時的変動

状態1は前期に比べて変化がなく通常の状態が続く場合で、パラメーターは変動せず W は 0 行列で、 V は通常値をとる。これらを W_N, V_N としよう。状態2は価格水準が段階的に変化する場合で、 W のうち μ_t の分散 σ_{11} が正の値をとり他はすべて0とした。これを W_T としよう。また V は V_N とした。状態3は勾配要因の変化がある場合であり、 W のうち β と r の分散 σ_{22}, σ_{33} が正で他は0とした。これを W_S としよう。 V は V_N である。状態4は、価格は大きく変動するが、これはパラメーター以外の要因によっている場合で、 V の値が通常値よりかなり大きいことを意味する。これを V_A としよう。 W は W_N である。これらをまとめると次表のようになる。

表 1

		W	V
状態	1	W_N	V_N
	2	W_T	V_N
	3	W_S	V_N
	4	W_N	V_A

具体的な値には根拠はないが、予測値への影響をみるために、幾通りかの値を与えてみた。例えば、 V_A については、 V_N の9倍と25倍の値について試みた。

π については、これは $P(S_t=j | S_{t-1}=i)$ であるから、 i, j と t とに依存するが、ここでは時間的に変化せず、しかも j のみに依存するものと想定

した。すなわち、ある期の状態とその次の期の状態とは確率的に独立であると仮定した。現実にはこの仮定は妥当とは云えないであろうが、主として確率付与のための便宜上からの仮定である。

インプットと初期値の値を変えて幾通りかの計算を試みたが、予測の一例を上げると次頁の2つの図の様になる。詳細は後の表2に示してある。尚、データは、1974年3月上旬から60週にわたる週別データである(日銀データによる)。図からわかる様に、この期間為替レートは直物先物共に24期近辺から46期近辺にかけてかなりはつきりとした高水準の安定期があり、その後で水準の変化があったという点が特徴である。

予測値に対するインプットの影響

予測値に対して与えるインプットの影響はかなり錯綜しており、計算式から直ちにわかるという訳ではないので、インプットの値を種々変えて影響をみることにした。予測を評価する尺度は二乗誤差であるが、それと共にここではとくに変動期における予測値への影響、および状態の確率が妥当なものであるかどうかをみることにしよう。

初期値は m_0 , C_0 , P_0 であるが、これらははじめの数期ないし10期くらいまで影響を与えるだけで、その後はほとんど同一の予測値が得られた。これは勿論モデルとデータに依存しているだろうと思われる。しかし観測に先立って行わなければならない事前分布の付与はかなり困難な作業であるから、初期値の影響がどの程度の期間継続するかは関心もたれる。かなり大きく値を変えて計算を試みたが、大体上で述べた期間を過ぎると影響はみられなくなるから、ある程度の長い期間にわたって観測を行うことができる場合ならば、初期値の付与についてそれほど心配しなくてもよいと云えるだろう。

かなり重大な影響を及ぼすのは他のインプット、 π , V , W であった。これは予想されたことではあるが、とりわけ V と W の影響が大きくみられる。

π については、既に述べたように、時間的に変化せず、しかも現期の状態にのみ関係すると仮定した。Harrison-Stevens の前掲論文では4つの状態に対して (0.9, 0.003, 0.003, 0.094) という値 (これを π_1 としよう) を勧めているが、更にこれを少し修正したものも試みたところ、予測値にはあまり差がないが、2乗誤差が小さくなるものも得られた。これは状態2, 3, 4の確率を3倍程度にしたもの (これを π_2 としよう) であるが、予測に差がでるのは、為替価格に大きな変化があるところであり、データはかなり典型的なトレンド

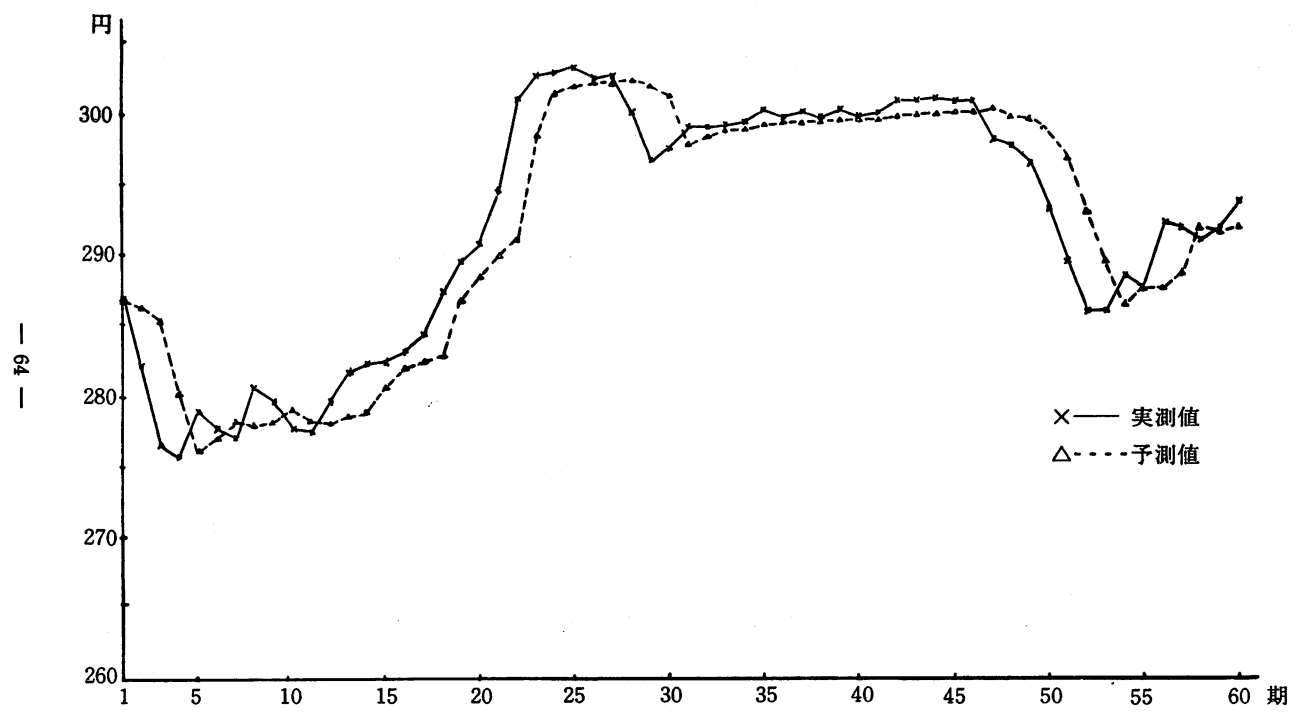


図1. 直物価格と予測値

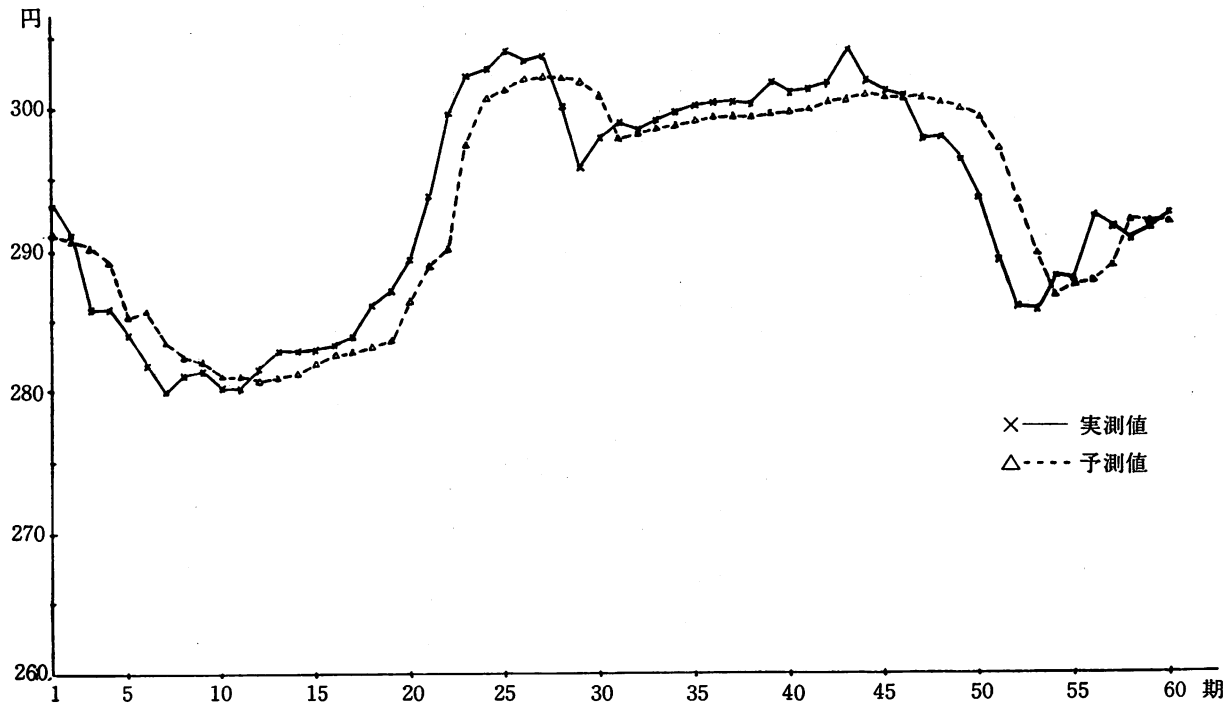


図2. 先物価格と予測値

変化を示しているから当然なことであるかもしれない。次に問題になるのは、上述した π に関する便宜的な仮定である。これに対しては、時間的に一定という仮定はそのままにして i と j に依存する様に改めるか、あるいは毎毎に計算されるようなメカニズムを考案するか、等の方法が考えられる。ここではまだこの問題には手をつけるに至っていないが、後者の方法をとれば、 m_0 、 C_0 等と同様の扱いができる。これらのことが予測あるいは推定の改善にどれだけ貢献するかはわからないが、モデルとしての整合性を強めることにはなるだろう。

今度行った試算でもっとも大きな影響を示したのは V であった。 V で問題になるのは状態 4 の場合の V_4 であるから、これを平常の値 V_N の 9 倍にしたもの（これを V_1 としよう。）と 25 倍にしたもの（これを V_2 としよう）とについて検討しよう。まず状態の確率に対する影響をみると、 V_1 の場合、観測値の変動に対してより敏感で状態 4 の確率が大きくなる傾向があるのに対して、 V_2 の場合は鈍感で状態 1 の確率が比較的大きくなる。またかなり大幅の上昇あるいは下降の場合、前者がトレンド変化（状態 2）の確率を大きくするのに対し、後者では一時的変動（状態 4）の確率が大きくなる。例えば、20～22 期の急激な上昇の場合、前者は状態 2 の確率を大きくしてトレンド変化を示しているのに対し、後者は一時的変動状態の確率を大きくしている。51～53 期あたりの下降についても同様な現象がみられる。結局、 V_4 が小さい場合、ステップ変化がかなりはっきりと指示されるのに対し、 V_4 が大きい場合は曖昧になるが、一方前者は少しの変化にも敏感すぎる様に思われる。次に予測に関しては、 V_1 の場合の方が 2 乗誤差が小さくなっているが、これは、安定期である 32～45 の期間には両者は殆ど同じ予測値を与えるのに対して、急速な下降および上昇の期間では V_1 の方が大体においてよい予測値を与えていることによる。また予測誤差の分散共分散行列 V_0 は V_1 の場合の方が V_2 の場合より小さくなっているが、これは計算式からもわかることである。

W の影響をみるためには、 W_T と W_S （表 1）の値を変えてみればよいから、両者が小さな場合（これを W_1 としよう）と大きな場合（これを W_2 としよう）について検討した。 W_2 は分散要素が W_1 のその 4 倍程度の大きさである。まず状態の確率については、 W_2 の場合トレンド変化をかなりはっきりと識別するのにに対し、 W_1 の場合はこれを一時的変動と識別する傾向がある。小さな変動については、両者にはほとんど差がみられなかった。予測値については、2 乗誤差は W_2 の場合の方が小さいが、これは大きく変動する期によ

い予測値を与えることに依っている。しかし V_0 にはあまり差が出なかった。結局 W は大きい方が状態の確率の推移も自然で、よい予測値が得られる様であるが、これは W の特殊な形のために、予測値を構成する前期のパラメーターの推定値において予測誤差 ($Y_{t-1} - \hat{Y}_{t-1}$) により大きなウェイトがおかれることから、観測値の変動により敏感であるためと思われる。

計 算 例

最後に最も良い予測値が得られた計算例について少し述べておこう。予測結果は表 2 に示してある。この計算は、 V は V_1 を、 W は W_2 を、 π は π_2 をとった場合である。

まず状態の確率についてみると、観測値があまり変動しない 32~45 期あたりでは、例えば (0.9650, 0.0026, 0.0036, 0.0288) の様な値をとるが、変動のあるところでは、これとはかなり異なる様相を示す。初期の頃の下降においては、2 期で (0.0000, 0.0014, 0.0170, 0.9816) となっており、まだこれを状態 4 と識別しているが、更に落ち込む 3 期には (0.0000, 0.2319, 0.0013, 0.7668) となってトレンド変化の確率がかなりの大きさになっている。しかしこの程度の下降ではあまり強くトレンド変化を指示していない。次に多少の変動を示す 4~11 期では、状態 1 と 4 の区別が微妙で、大体等しい程度の確率になったり、交互に大きくなったりするが、状態 2 と 3 の確率は小さい。12~23 期にかけてはじめは緩慢に、後に急激に上昇しているが、ここでは、はじめは状態 1 の確率が大きく、17 期あたりから状態 4 の確率が大きくなり、23 期の急激な上昇において [0.0000, 0.6632, 0.0001, 0.3367] となって、トレンド変化をかなり明確に識別している。その後の長い安定期については既に述べた様に状態 1 の確率が大きい。後期の下降は 50 期から 52 期にかけて識別されている。この様に、この計算例の場合、状態の確率は比較的自然的な推移を示している。すでに述べたように、 V_4 が小さな値をとる場合、少しの変動にも敏感に反応するから、状態 1 と 4 の確率が急に入れかわったりすることがあり、細部ではおかしなことも起っているが、大体において妥当である。他の計算例の場合、推移が自然でないものもあった。

$Y_t = (Y_t^0, Y_t^1)$ の Y^{t-1} が与えられた時の事後分布は、平均 \hat{Y}_t 、分散共分散行列 V_0 の 2 次元正規分布であり、 \hat{Y}_t を予測値とすることは、2, 3 節で述べた通りである。 V_0 は近似的なものであるが、この分布から信頼領域 (ベイジアン) を構成することができる。厳密に云えば、これは楕円領域になるが、近似的に \hat{Y}_t のまわりに $\pm\sigma$, $\pm 2\sigma$, $\pm 3\sigma$ の長方形領域をとって、観測値が

これらの領域に入るかどうかをみると、60の観測値のうち $\pm 3\sigma$ の中に入ったものは87%、 $\pm 2\sigma$ の場合は78%、 $\pm \sigma$ の場合は55%であった。

表 2 予 測 結 果

期	直 物		先物(1ヶ月)		期	直 物		先物(1ヶ月)	
	観測値	予測値	観測値	予測値		観測値	予測値	観測値	予測値
1	286.80	286.60	293.00	290.75	31	299.10	297.86	298.95	297.71
2	282.20	286.25	290.80	290.69	32	299.00	298.33	298.40	298.17
3	276.55	285.76	285.70	290.14	33	299.15	298.71	299.05	298.49
4	275.85	280.90	285.70	289.11	34	299.95	298.83	299.67	298.63
5	279.00	276.11	284.00	285.12	35	300.41	299.10	300.50	298.89
6	277.85	277.10	281.85	285.70	36	299.77	299.33	300.10	299.15
7	277.10	278.31	279.90	283.46	37	300.03	299.37	300.13	299.23
8	280.60	278.04	281.10	282.29	38	299.90	299.46	300.10	299.41
9	279.75	278.15	281.40	281.95	39	302.25	299.50	301.85	299.58
10	277.60	279.12	280.10	281.23	40	299.84	299.53	301.10	299.34
11	277.55	278.30	280.10	281.07	41	300.00	299.57	301.25	299.85
12	279.60	278.07	281.45	280.84	42	300.90	299.67	301.90	300.30
13	281.90	278.45	282.65	280.90	43	300.98	299.82	304.10	300.51
14	282.40	278.78	282.60	281.04	44	300.99	299.90	301.84	300.85
15	282.70	278.63	282.90	281.91	45	300.85	299.99	301.15	300.79
16	283.15	282.00	283.05	282.53	46	300.93	300.06	300.90	300.73
17	284.27	282.36	283.82	282.68	47	298.00	300.16	297.60	300.67
18	287.35	282.82	285.96	282.96	48	297.98	299.92	297.98	300.32
19	289.50	283.46	287.00	283.45	49	296.39	299.66	296.37	300.04
20	290.75	286.83	289.40	286.24	50	293.30	299.04	293.45	299.39
21	294.50	289.96	293.90	288.75	51	289.50	296.72	289.07	297.04
22	301.00	291.10	299.60	289.93	52	285.90	293.12	285.95	293.39
23	302.65	298.47	302.30	297.30	53	285.97	289.53	285.60	289.76
24	302.89	301.55	302.70	300.56	54	288.50	286.57	288.01	286.73
25	303.17	301.95	304.00	301.14	55	287.55	287.64	287.85	287.74
26	302.55	302.15	303.00	301.63	56	292.30	287.56	292.32	287.68
27	302.64	302.23	303.30	301.91	57	291.90	288.71	291.40	288.82
28	300.00	302.31	299.90	302.20	58	291.00	292.02	290.65	292.08
29	296.65	301.96	295.60	301.87	59	291.90	291.84	291.58	291.85
30	297.40	301.09	297.80	300.95	60	293.70	291.91	292.50	291.89

5. おしまいに

カルマン・モデルの一つの応用例についてこれまで述べて来たことは、きわめて大雑把なことであり、またデータに強く依存していることも考えられるので、更に緻密な検討と広範な試算が必要なことは勿論である。しかしインプット、とりわけ初期値以外のインプットの影響の大きさの程度が示唆されたことは予想される面があったとは云え、貴重な経験であった。また既にふれたことではあるが、予測の手続を更に整合的にし、またモデルを豊富にするために、 π の取扱いについて考えなければならないであろう。その点では、 w の平均をここでは 0 としたが、状態の確率を比較的妥当なものに計算できることがわかったので、状態を表現するために 0 ではない平均をもつけ加えれば、予測値は観測値の変動にもっと敏感になるのではないかと思われる。これらの点について更に検討したいと思う。最後に、データを提供して下さった日本銀行調査室の折谷吉治氏に感謝したい。

文 献

- 1 有本卓 (1977). カルマン・フィルター. 産業図書
- 2 Harrison, P. J. and Stevens, C. F. (1971). A Bayesian approach to short-term forecasting. *Oper. Res. Quart.* **22**, 341-362
- 3 Harrison, P. J. and Stevens, C. F. (1976). Bayesian Forecasting. *J. R. Statist. Soc. B*, **38**, 205-247
- 4 Kalman, R. E. (1963). *New methods in Wiener filtering theory*. Wiley
- 5 Wiener, N. (1949). *Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications*, Wiley.