

時変パラメーター回帰モデルに 対するカルマン・フィルターの 応用について

原 田 桂 一 郎

目 次

1. はじめに
2. 時変パラメーター回帰モデルの定式化と推定
3. 時変パラメーターの計算と予測
4. カルマン・フィルター応用上の問題
5. むすび
- 付一1 線型動的システム・モデル
- 付一2 Bayes 推定法

1. はじめに

計量モデル（回帰モデル）において、パラメーター（回帰係数）が時間経過と共に変動するものと仮定した時変パラメーター回帰モデルを定式化することができる。この場合、パラメーターは確率変数と考えて、その変動は非定常確率過程によって定式化が行われる。またどのような確率過程をあてはめるかは分析者の判断によるものである。

パラメーターが、時々、急激な変化を伴いながら非定常に推移するものと仮定した場合、その変動はランダム・ウォーク過程で定式化することが適切であろう。すなわち、パラメーター・ベクトル $\beta_t = (\beta_{1t} \beta_{2t} \cdots \beta_{kt})'$

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

の変動は,

$$\beta_{t+1} = \beta_t + \varepsilon_t \quad (1-1)$$

と表わされ, また回帰式は

$$y_t = x_t' \beta_t + u_t \quad (1-2)$$

である。ここで, $\{\varepsilon_t\}$ は $\varepsilon_t = (\varepsilon_{1t} \varepsilon_{2t} \cdots \varepsilon_{kt})'$, $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma$ である相互独立なランダム・ベクトル・プロセス, $\{u_t\}$ は $E u_t = 0$, $E u_t^2 = \sigma_u^2 > 0$ の相互独立なスカラー・プロセス, $\{x_t\}$ は $x_t' = (x_{1t} x_{2t} \cdots x_{kt})$ である既知の非確率変数の系列, そして $\{y_t\}$ は観測スカラー系列である。

ところで, 工学の分野では Kalman[6] によって, Linear Filtering と称される多変数時系列モデルとその推定法が開発されている。この Kalman Filter モデルは次の二つの線型定差方程式によって表現される。すなわち, m -ベクトル y_t の時系列 y_1, y_2, \dots, y_T は出力方程式 (観測方程式)

$$y_t = X_t \beta_t + u_t \quad t = 1, 2, \dots, T \quad (1-3)$$

と状態推移方程式

$$\beta_t = \Phi \beta_{t-1} + \varepsilon_t \quad t = 2, \dots, T \quad (1-4)$$

によって構成されることになる。^①ここで $X_t (m \times k)$ は既知である非確率変数行列, ベクトル $\beta_t (k \times 1)$ を t 期の状態と称し, $\Phi (k \times k)$ は状態推移行列, また $u_t (m \times 1)$ と $\varepsilon_t (k \times 1)$ はそれぞれランダム・ベクトルである。

(1-3) の各方程式は変数ベクトル $x_t (1 \times k)$ と t 期の状態変数と呼ばれているベクトル $\beta_t (k \times 1)$ の積とランダム項の和であり, 回帰式としての形式を保持している。また状態推移方程式 (1-4) は状態変数ベクトルの動的推移を表現したものである。このように Kalman Filter モデルは,

(1-1) と (1-2) で示した時変パラメーター回帰モデルと類似の形式を有している。

本稿は, この点に着目して, 回帰係数がランダム・ウォーク過程にした

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）
 がって変動するとした時変パラメーター回帰モデルの定式化とその推定
 に、Kalman Filter 法を応用することを試み、応用に際して直面する問題
 の検討を行うものである。

注

① 付一1 参照。

2. 時変パラメーター回帰モデルの定式化と推定

回帰係数（パラメーター）ベクトルが時変であり、その変動がランダム・ウォーク過程にしたがうと仮定した回帰式を、Kalman Filter モデルに則して次のように定式化する。

$$\beta_{t+1} = \beta_t + \varepsilon_t \quad (2-1) \quad \text{①}$$

$$y_t = x'_t \beta_t + u_t \quad t=1, 2, \dots, T \quad (2-2)$$

ここで、 β_t はパラメーター・ベクトルであり $\beta_t = (\beta_{1t} \cdots \beta_{kt})'$, $x_t' = (x_{1t} x_{2t} \cdots x_{kt})$ は独立変数（非確率変数）ベクトル、 y_t は t 期における従属変数の観測値であり、 $\{\varepsilon_t\}$ は $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$, $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma$, $\varepsilon_t \sim N(0, \Sigma)$ であるランダム・ベクトル・プロセス、そして $\{u_t\}$ は $E u_t = 0$, $E(u_t^2) = \sigma_u^2 > 0$, $u_t \sim N(0, \sigma_u^2)$ であるランダム・スカラー・プロセスを表わす。また $\{\varepsilon_t\}$ と $\{u_t\}$ は相互独立である。

このモデルでは、パラメーター・ベクトルが固定された未知数ではなく、確率変数と考えているので、その推定には最大尤度法や最小二乗法をそのまま適用することはできない。Kalman Filter モデルの推定には Ho and Lee〔4〕による Bayes 推定法、Duncan and Horn〔2〕などによるリカーシブ最小二乗法、そして Sant〔10〕による一般化最小二乗法の適用が、これまでに試みられている。

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

本稿では、時変パラメーター回帰モデル（2—1）（2—2）の係数ベクトルの推定に Bayes 推定法^②を採用する。そのために、先の事項に加えて次のことを仮定する。

未知のパラメーター・ベクトルは正規分布にしたがう確率変数として取り扱われ、その初期値 β_0 は $\beta_0 \sim N(\hat{\beta}_0, P_0)$ また $\hat{\beta}_0, P_0$ （初期値の共分散行列）は既知とする。先に示した Σ, σ_u^2 も既知であるとの仮定を置く。さらに、 $\beta_0, \{\varepsilon_t\}$ と $\{u_t\}$ は相互独立、すなわち、

$$E(u_t \varepsilon_s) = 0, E(\beta_0 \varepsilon_s') = 0, E(u_t \beta_0) = 0 \\ t=1, 2, \dots; s=1, 2, \dots \quad (2-3)$$

である。

いま y_t のサンプル Y_t が得られたとき、パラメーター・ベクトルの事後密度関数 $p(\beta_t/Y_t)$ は次のように与えられる（Bayes の定理）。

$$p(\beta_t/Y_t) = \frac{p(y_t/\beta_t) p(\beta_t/Y_{t-1})}{p(y_t/Y_{t-1})} \quad (2-4)$$

この事後密度関数は三つの密度関数によって構成される。また、損失関数^③を推定誤差 $(\hat{\beta} - \beta)$ の二次形式で与え、その期待値（この場合は分散）を最小とするパラメーター・ベクトルの推定量はパラメーター・ベクトルの事後密度関数（2—4）の平均^④で与えられる。したがって、事後密度関数を定めることが必要となり、それは以下のようにして求められる。

まず、 $p(\beta_t/Y_{t-1})$ は正規分布にしたがい、（2—1）を考慮して、

$$E(\beta_t/Y_{t-1}) = \hat{\beta}_{t-1} \\ Cov(\beta_t/Y_{t-1}) = E[\{\beta_t - E(\beta_t/Y_{t-1})\} \{\beta_t - E(\beta_t/Y_{t-1})\}' / Y_{t-1}] \\ = E[\{(\beta_{t-1} - \hat{\beta}_{t-1}) + \varepsilon_{t-1}\} \{(\beta_{t-1} - \hat{\beta}_{t-1}) + \varepsilon_{t-1}\}' / Y_{t-1}] \\ = p_{t-1} + \Sigma = P_t = M_t$$

同様にして、 $p(y_t/Y_{t-1})$ も正規分布にしたがい、（2—2）を考えて

時変パラメーター回帰モデルに対するカルマン・フィルターの応用 (原田)

$$E(y_t/Y_{t-1}) = x_t' \hat{\beta}_t$$

$$\begin{aligned} Cov(y_t/Y_{t-1}) &= E[\{y_t - E(y_t/Y_{t-1})\} \{y_t - E(y_t/Y_{t-1})\}' / Y_{t-1}] \\ &= x_t' P_t x_t + \sigma_u^2 \end{aligned}$$

最後に $p(y_t/\beta_t)$ もまた正規分布にしたがい、(2-2) から

$$E(y_t/\beta_t) = x_t' \beta_t$$

$$Cov(y_t/\beta_t) = \sigma_u^2$$

となる。以上の三つの密度関数を結び合せるとパラメーター・ベクトルの事後密度関数は、

$$\begin{aligned} P(\beta_t/Y_t) &= \frac{|x_t' M_t x_t + \sigma_u^2|^{1/2}}{(2\pi)^{k/2} |\sigma_u^2|^{1/2} |M_t|^{1/2}} \\ &\quad \cdot \exp[-1/2\{(\beta_t - \hat{\beta}_{t-1}) M_t^{-1} (\beta_t - \hat{\beta}_{t-1})' \\ &\quad + (y_t - x_t' \beta_t) (\sigma_u^2)^{-1} (y_t - x_t' \beta_t)' \\ &\quad - (y_t - x_t' \hat{\beta}_t) (x_t' M_t x_t + \sigma_u^2)^{-1} (y_t - x_t' \hat{\beta}_t)'\}] \end{aligned} \quad (2-5)$$

のようになる。(2-5) の { } 内を整理すると

$$\begin{aligned} P(\beta_t/Y_t) &= \frac{|x_t' M_t x_t + \sigma_u^2|^{1/2}}{(2\pi)^{k/2} |\sigma_u^2|^{1/2} |M_t|^{1/2}} \\ &\quad \cdot \exp\{-1/2(\beta_t - \hat{\beta}_t) P_t^{-1} (\beta_t - \hat{\beta}_t)'\} \end{aligned} \quad (2-6)$$

ここで

$$\hat{\beta}_t = \hat{\beta}_{t-1} + M_t x_t (x_t' M_t x_t + \sigma_u^2)^{-1} (y_t - x_t' \hat{\beta}_{t-1}) \quad (2-7) \quad \textcircled{5}$$

$$P_t^{-1} = M_t^{-1} + (\sigma_u^2)^{-1} x_t x_t' \quad (2-8) \quad \textcircled{5}$$

あるいは

$$P_t = M_t - M_t x_t (x_t' M_t x_t + \sigma_u^2)^{-1} x_t' M_t \quad (2-9) \quad \textcircled{5}$$

また

$$M_t = P_{t-1} + \Sigma \quad (2-10)$$

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）である。さらに

$$K_t = (P_{t-1} + \Sigma)x_t[x_t'(P_{t-1} + \Sigma)x_t + \sigma_u^2]^{-1} \quad (2-11)$$

と置くと、事後密度の平均と共分散行列は、それぞれ

$$\hat{\beta}_t = \hat{\beta}_{t-1} + K_t(y_t - x_t'\hat{\beta}_{t-1}) \quad (2-12)$$

$$P_t = (I - K_t x_t')(P_{t-1} + \Sigma) \quad (2-13)$$

のように与えられる。

なお、(2-12)、(2-13) は (2-1) と (2-2) で表わされる動的システム（モデル）の Kalman Filter, また (2-11) は Kalman Gain と呼ばれる。

パラメーター・ベクトルの推定量 (2-12) は、それが線型形式で与えられ、また推定誤差の二次式で損失関数を定めた Bayes 推定量であるから、最良線型不偏推定量となる。

注

- ① Kalman Filter モデルでは、(2-1) は、 $\beta_{t+1} = \phi\beta_t + \varepsilon_t$ と表わされ、推移行列 $\phi(k \times k)$ が単位行列 I に等しくなるとはかぎらない。 β_t の変動をランダム・ウォーク過程にしたがうと仮定したので (2-1) において $\phi=1$ とした。また、経済データから ϕ の良好な推定量を得ることは実際には不可能であるから、その推定を行う代りに便宜上、これを単位行列に等しいと置くことは現実の問題を分析する上で都合がよい。
- ② Bayes 推定法については付-2 を参照。
- ③ 推定値 $\hat{\theta}$ を得たことによる、または、この推定値が真の値から乖離していることによる損失を、パラメーターの真の値 θ と推定値 $\hat{\theta}$ の関数 $l(\hat{\theta}; \theta)$ で表わしたものを損失または損失関数という。この損失の期待値、 $E[l(\hat{\theta}; \theta)]$ をリスクまたはリスク関数という。
- ④ 付-2 を参照。
- ⑤ (2-7)、(2-8) および (2-9) の導出についての詳細は Zazwinski [5] chap. 7 を参照。

3. 時変パラメーターの計算と予測

観測期間内 ($0 < t < T$) の時変パラメーターは, Kalman Filter と Kalman Gain を用いて, パラメーター (ベクトル) の初期値 $\hat{\beta}_0$, その共分散行列 P_0 , そして, σ_u^2 , Σ が既知であるとして次のように計算される。

まず, $t=1$ の場合,

$\hat{\beta}_0$ と P_0 が与えられているから, (2-11) から, $t=1$ の Gain 行列

$$(1) \quad K_{t=1} = (P_0 + \Sigma)x_{t=1}[x'_{t=1}(P_0 + \Sigma)x_{t=1} + \sigma_u^2]^{-1}$$

を求める。次にパラメーター $\hat{\beta}_{t=1}$ を (2-12) から

$$(2) \quad \hat{\beta}_{t=1} = \hat{\beta}_0 + K_{t=1}(Y_{t=1} - x'_{t=1}\hat{\beta}_0)$$

のように計算する。さらに, この時点での $\beta_{t=1}$ の分布の共分散行列を (2-13) から

$$(3) \quad P_{t=1} = (I - K_{t=1}x'_{t=1})(P_0 + \Sigma)$$

と算定する。

$t=2$ の場合, ステップ(1)に戻ってこの時点での Gain 行列と Filter を計算する。こうして, 時点が次々とすすむごとに以上を逐次繰り返すことにより, 観測期間内でのパラメーター・ベクトル, $\hat{\beta}_1 \hat{\beta}_2 \cdots \hat{\beta}_T$ が算出される。

次に $T < t$ の範囲でのパラメーター・ベクトルの値を求めること, すなわちパラメーター・ベクトルの予測は, 初期条件 ($\hat{\beta}_T, P_T$) をもって次の関係を用いて行われる。

$$\hat{\beta}_{t+1} = E\{\beta_{t+1}/Y_t\} = \hat{\beta}_t$$

$$P_{t+1} = E\{(\beta_{t+1} - \hat{\beta}_{t+1})(\beta_{t+1} - \hat{\beta}_{t+1})'/Y_t\}$$

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

$$\begin{aligned} &= E[\{(\beta_t - \hat{\beta}_t) + \varepsilon_t\} \{(\beta_t - \hat{\beta}_t) + \varepsilon_t\}' / Y_t] \\ &= P_t + \Sigma \end{aligned}$$

であるから、結局

$$\hat{\beta}_{t+1} = \hat{\beta}_t \quad (3-1)$$

$$P_{t+1} = P_t + \Sigma \quad (3-2)$$

によって予測値が求められることになる。しかし、(3-1) から明らかのように、予測値は初期条件の値のままである。これはパラメーター・ベクトルの推移方程式 (2-1) において、 β_t の推移行列を単位行列に等しいと置いたことによるものである。(2-1) のようにパラメーター・ベクトルの推移を仮定すれば、そのモデルは観測期間内のパラメーター・ベクトルの推定に供することはできるが、予測を目的とするモデルとしては適切ではない。予測のためには、パラメーター・ベクトル推移方程式の推移行列 Φ を単位行列とせず、

$$\beta_{t+1} = \Phi \beta_t + \varepsilon_t \quad (3-3)$$

$$y_t = x_t' \beta_t + u_t \quad (3-4)$$

というモデルが適している。このモデルの Kalman Filter は (2-12)、(2-13) のように与えられるが、Kalman Gain は

$$K_t = (\Phi P_{t-1} \Phi' + \Sigma) x_t \{x_t' (\Phi P_{t-1} \Phi' + \Sigma) x_t + \sigma_u^2\}^{-1} \quad (3-5)$$

となる。また、(3-3) により

$$\begin{aligned} \hat{\beta}_{t+1} &= E\{\beta_{t+1} / Y_t\} = \Phi \hat{\beta}_t \\ P_{t+1} &= E\{(\beta_{t+1} - \hat{\beta}_{t+1})(\beta_{t+1} - \hat{\beta}_{t+1})' / Y_t\} \\ &= E[\{\Phi(\beta_t - \hat{\beta}_t) + \varepsilon_t\} \{\Phi(\beta_t - \hat{\beta}_t) + \varepsilon_t\}' / Y_t] \\ &= \Phi P_t \Phi' + \Sigma \end{aligned}$$

であるから、パラメーター・ベクトルの予測は $(\hat{\beta}_T, P_T)$ を初期値とし

$$\hat{\beta}_{t+1} = \Phi \hat{\beta}_t \quad (2-18)$$

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

$$P_{t+1} = \Phi P_t \Phi' + \Sigma \quad (2-19)$$

によって行なわれる。この場合の問題は、推移行列 Φ の良好な特性をもつ推定値を、経済分析に供されるような系列数の少ないデータによって求めることが不可能なことである。

また、パラメーター・ベクトル β_t の予測に際しては、変数 y_t および x_t' の値が $t > T$ においても入手可能でなければならない。Kalman Filter による予測は、従属変数 y_t の将来値を予測するための方法ではなく、それらの将来値を与えて、パラメーターの将来値の予測を行うものである。Kalman Filtering はあくまでもパラメーター・ベクトル β_t の変動を分析するのに有効な方法である。

4. カルマン・フィルター応用上の問題

Kalman Filter 法を計量モデルの推定に応用する上での問題は、 σ_u^2 、 Σ 、 β_0 および P_0 が既知であることを要すところにある。時変パラメーター回帰モデルではこれらの量が未知であるから、何らかの方法で推定しておかなければならない。

Sarris [11] は最大尤度法により、パラメーター・ベクトル β_t の初期値 β_0 を求めることを試みている。我々のモデルではそれは以下のように行われる。

まず (2-1) における β_t は繰り返し代入を行うことにより次のように表わすことができる。

$$\beta_t = \beta_0 + \sum_{j=1}^t \varepsilon_j \quad (4-1)$$

これを (2-2) に代入すると

時変パラメーター回帰モデルに対するカルトン・フィルターの応用（原田）

$$y_t = x_t' \beta_t + u_t = x_t' \beta_0 + x_t' \sum_{j=1}^t \varepsilon_j + u_t \quad (4-2)$$

を得る。さらに,

$$v_t \equiv x_t' \sum_{j=1}^t \varepsilon_j + u_t \quad t=1, 2, \dots, N \quad (4-3)$$

と定める。また

$$Y \equiv [y_1 \ y_2 \ \dots \ y_N]'$$

$$X \equiv [x_1' \ x_2' \ \dots \ x_N']'$$

$$v \equiv [v_1 \ v_2 \ \dots \ v_N]'$$

とおけば, (4-3) は

$$Y = X\beta_0 + v \quad (4-4)$$

と書き表わすことができる。

ベクトル v は多変量正規分布にしたがい, その平均は零 ($E(v)=0$), ε_t の共分散行列を

$$\Sigma = \sigma_\varepsilon^2 R \delta_{ts}, \quad R(k \times k)$$

と書きかえると, v の共分散行列は次のように与えられる。

$$\begin{aligned} E v_t^2 &= E \left[x_t' \sum_{j=1}^t \varepsilon_j + u_t \right] \left[u_t + \left(\sum_{j=1}^t \varepsilon_j' \right) x_t \right] \\ &= \sigma_u^2 + \sigma_\varepsilon^2 x_t' R x_t \end{aligned} \quad (4-5)$$

$$\begin{aligned} E v_t v_i &= E \left[x_t' \sum_{j=1}^t \varepsilon_j + u_t \right] \left[u_i + \left(\sum_{j=1}^i \varepsilon_j' \right) x_i \right] \\ &= x_t' R x_i \end{aligned} \quad (4-6)$$

であるから v の共分散行列は

$$E(vv') = \sigma_u^2 I + \sigma_\varepsilon^2 Q \quad (4-7)$$

となる。ここでは $I(N \times N)$ 単位行列, $Q(N \times N)$ 行列の (i, j) 要素は

$$Q_{ij} = \min(N-i+1, N-j+1) x_i' R x_j \quad (4-8)$$

である。

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

Sarris はここで、 $Q(N \times N)$ 行列については全要素にわたるまで既知であるという仮定を置いている（したがって R についても既知と仮定される）。さらに、

$$\theta \equiv \sigma_e^2 / \sigma_u^2 \quad (4-9)$$

と定めると、 v の共分散行列は

$$E(vv') = \sigma_u^2 [I + \theta Q] = \sigma_u^2 P(\theta) \quad (4-10)$$

と表わすことができる。

したがって、 Z, Y の対数尤度関数は

$$\begin{aligned} L(Y; X, \beta_0, \sigma_u^2, \theta) = & -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma_u^2 - \frac{1}{2} \ln |P(\theta)| \\ & - \frac{1}{2\sigma_u^2} (Y - X\beta_0)' P(\theta)^{-1} (Y - X\beta_0) \end{aligned} \quad (4-11)$$

と書くことができる。 β_0 と σ_u^2 はこの尤度を最大にするものとして

$$\hat{\beta}_0(\theta) = [X'P^{-1}(\theta)X]^{-1}X'P(\theta)^{-1}Y \quad (4-12)$$

$$\begin{aligned} \sigma_u^2(\theta) = & \frac{[Y - X\hat{\beta}_0(\theta)]' P(\theta)^{-1} [Y - X\hat{\beta}_0(\theta)]}{N} \\ = & \frac{Y' P(\theta)^{-1} [I - X[X'P(\theta)^{-1}X]^{-1}X'P(\theta)^{-1}] Y}{N} \end{aligned} \quad (4-13)$$

のように与えられる。

いま (4-12)、(4-13) を (4-11) の β_0 と σ_u^2 に代入することにより θ についての “concentrated” 尤度は次のように与えられる。

$$\begin{aligned} L(Y; X, \theta) = & -\frac{N}{2} \ln(2\pi) \\ & - \frac{N}{2} \ln \frac{1}{N} \{Y' P(\theta)^{-1} [I - X(X'P(\theta)^{-1}X)^{-1}X'P(\theta)^{-1}] Y\} \\ & - \frac{1}{2} \ln |P(\theta)| - \frac{N}{2} \end{aligned} \quad (4-14)$$

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

θ の値を変化させ、(4—14) が最大となる $\hat{\theta}_*$ を選定し、この $\hat{\theta}_*$ を (4—12), (4—3) に代入して $\hat{\beta}_0, \hat{\sigma}_u$ を計算することになる。

以上の手続きにおいての問題点は、(2—1) の ε_t の共分散行列 Σ をその全要素にわたって既知と仮定していることである。測定不可能な ε_t の共分散行列が既知であるというようなことは現実にはありえない。したがって、モデル (2—1), (2—2) における時変回帰パラメーター・ベクトル β_t の初期値 β_0 の値を、最尤法で推定することは、實際上、不可能であるといえよう。

ところで、観測期間以前のある一定期間にわたって独立変数 x_t' と従属変数 y_t の値（データー）が入手可能な場合、(2—2) のパラメーター・ベクトル β_t と誤差項 u_t の分散 σ_u^2 を、最尤法あるいは最小二乗法により、それぞれの、推定値 $\hat{\beta}$ および $\hat{\sigma}_u^2$ （最小二乗法では不偏分散 s^2 の推定）を求めてみる。そして、これらの値をパラメーター・ベクトルの初期値 $\hat{\beta} = \beta_0$ 、また誤差項 u_t の分散 $\hat{\sigma}_u^2 = \sigma_u^2$ と置くことが考えられる。

しかし、係数ベクトルの初期値の共分散行列 P_0 と ε_t の共分散行列 Σ については、それらを推定することができない。したがって、(2—1), (2—2) のモデルに Kalman Filter 法を適用するためには、 P_0 と Σ を恣意的に定めておくことが必要となる。

いずれにしろ、何らかの方法で $\sigma_u^2, \Sigma, \beta_0$ および P_0 の値を与えなければ、Kalman Filter 法により、パラメーター・ベクトルの推定をすることは不可能である。

5. む す び

本稿では、時変パラメーター回帰モデルのパラメーター（回帰係数）の

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

推定に、Kalman Filter 法を適用することを試みた。Kalman Filtering は線型動的システム（時変パラメーター回帰モデル）に、ランダム・シリーズが入力として加えられると、それがシステムの状態変数（パラメーター）を変動させ、また出力（従属変数）に影響する様を分析するものである。この方法によれば、従属変数と独立変数の値が与えられると、時変パラメーター（回帰係数）の変動の経過（値）が逐次推定されることになり、データ観測期間内でのパラメーターの変動を分析することができる。

また、Kalman Filtering による予測は、従属変数と独立変数の将来値をそれぞれ与えて、このときパラメーターの変動の模様を求めていくことになる。通常、計量モデルではパラメーターは不変なものとされる。したがって、計量モデルに基づく予測は、推定したパラメーター（構造）が予測期間中は不変として、これをベースに従属変数の将来値を定めていくことになる。しかし、時変パラメーター回帰モデルをベースに Kalman Filtering にしたがう予測は、時変パラメーターの将来値をフォローするものとなる。

Kalman Filtering による時変パラメーターの推定に際して、パラメーターの初期値の事前分布の平均値と共分散行列、パラメーター・プロセスの誤差の共分散行列、および回帰式の誤差の分散が既知であることを必要としている。

一般に経済データは受動的にしか得られず、実験が不可能であるから、これ等の値は何らかの方法で推定しなければならない。しかし、先述のようにこれ等の推定が不可能であることは、Kalman Filtering を計量分析に応用することを難しいものにしている。

付一1 線型動的システム・モデル

1. 連続時間モデル

連続的な時間経過にしたがって状態が推移していく動的システムを考える。この連続時間を t で表わすと、そのシステムは時間の関数である r 個の入力 $u(t)$ と m 個の出力 $Z(t)$ をもつものとする。そのシステムの状態を表わす n 個の状態変数 $x(t)$ は、入力 $u(t)$ によって影響を受け、その一部が出力 $Z(t)$ に反映される。

入力 $u(t)$ 、出力 $Z(t)$ および状態変数 $x(t)$ の間に、次のような連立の線型微分方程式

$$\frac{d[x(t)]}{dt} = \mathbf{F}(t)x(t) + \mathbf{B}(t)u(t) \quad x(t_0) = x_0 \quad (\text{A-1})$$

$$Z(t) = \mathbf{H}(t)x(t) \quad (\text{A-2})$$

が成立するとき、これらの式で表わされる体系を線型動的システムという。ここで、 $u(t)$ は r 次元の入力ベクトル、 $Z(t)$ は m 次元の出力ベクトルそして $x(t)$ は n 次元の状態変数ベクトルである。また、 $\mathbf{F}(t)$ は同次のシステム・ダイナミクスを表現する時間と共に変化する行列 ($n \times n$)、 $\mathbf{B}(t)$ は入力行列 ($n \times r$)、そして $\mathbf{H}(t)$ は出力行列 ($m \times n$) であり、システムの構造（時間経過と共に変化する）を定める。

なお、(A-2) は入力変数が直接出力に影響する場合を考慮して、

$$Z(t) = \mathbf{H}(t)x(t) + \mathbf{D}(t)u(t) \quad (\text{A-3})$$

と表わされることもある。

(A-1) を状態微分方程式、(A-2) あるいは (A-3) を出力方程式と呼び、この二つによって動的システム・モデルを構成する。

2. 状態微分方程式の解

まず次の微分方程式と初期条件を満足する行列 ($n \times m$) を“状態推移行列”と定義する。

$$d[\Phi(t, t_0)]/dt = F(t)\Phi(t, t_0) \quad (A-4a)$$

$$\Phi(t_0, t_0) = I \quad (A-4b)$$

さらに、状態推移行列は次の性質をもつ。

(1) $\Phi(t, t_0)$ は区間 $[0, \infty)$ におけるすべての t および t_0 について唯一定義される。

(2) 時間的経過 $t_1 < t_2 < t_3$ について、 t_1 から t_3 へ推移させる状態推移行列は、 t_1 から t_2 、 t_2 から t_3 へ推移させるそれぞれの状態推移行列の積に等しい。

$$\Phi(t_3, t_1) = \Phi(t_3, t_2) \Phi(t_2, t_1) \quad (A-5)$$

(3) $\Phi(t, t_0)$ は非特異であり、また

$$\Phi(t, t_0) \Phi(t_0, t) = \Phi(t, t) = I \quad (A-6)$$

したがって、

$$\Phi^{-1}(t, t_0) = \Phi(t_0, t)$$

である。

次に、状態微分方程式 (A-1) の解は以下の手順にしたがって求められる。

(A-1) の両辺に左から $\Phi(t, t_0)$ を乗じて

$$\Phi(t, t_0) \frac{d[x(t)]}{dt} - F(t) \Phi(t, t_0) x(t) = \Phi(t, t_0) B(t) u(t)$$

を得る。さらに、(A-4a) の関係を用いると、

$$\Phi(t, t_0) \frac{d[x(t)]}{dt} - \frac{d[\Phi(t, t_0)]}{dt} x(t) = \Phi(t, t_0) B(t) u(t)$$

であるから

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

$$\frac{d[\Phi(t, t_0) x(t)]}{dt} = \Phi(t, t_0) \mathbf{B}(t) u(t)$$

が成立する。この両辺を積分すると

$$\Phi(t, t) x(t) - \Phi(t, t_0) x(t_0) = \int_{t_0}^t \Phi(t, \tau) \mathbf{B}(\tau) u(\tau) d\tau$$

のようになるから、(A-6)を用いることにより、(A-1)の解は

$$x(t) = \Phi(t, t_0) x_0 + \int_{t_0}^t \Phi(t, \tau) \mathbf{B}(\tau) u(\tau) d\tau \quad (\text{A-7})$$

のように表わされることになる。

3. 不連続時間モデル

経済データーのように、測定が連続して実行できず、一定の時間を置いて行われる場合には、不連続時間についての線型動的システム・モデルが適切である。

$t_i (i=1, 2, \dots)$ を時点を表わすものとすれば、出力方程式は

$$Z(t_i) = \mathbf{H}(t_i) \cdot x(t_i) \quad (\text{A-8a})$$

あるいは

$$Z(t_i) = \mathbf{H}(t_i) x(t_i) + \mathbf{D}(t_i) u(t_i) \quad (\text{A-8b})$$

となる。

また、不連続時間モデルの状態変数の値、 $x(t_1), x(t_2), \dots$ が連続時間モデルの特定時点 t_1, t_2, \dots における値と一致するならば、状態微分方程式の解 (A-7) から

$$x(t_{i+1}) = \Phi(t_{i+1}, t_i) x(t_i) + \int_{t_i}^{t_{i+1}} \Phi(t_{i+1}, \tau) \mathbf{B}(\tau) u(\tau) d\tau \quad (\text{A-9})$$

を得る。また

$$u(t) = u(t_i) \quad t \in [t_i, t_{i+1}) \quad (\text{A-10})$$

と仮定すると (A-9) は

時変パラメーター回帰モデルに対するカルマン・フィルターの応用 (原田)

$$x(t_{i+1}) = \Phi(t_{i+1}, t_i) x(t_i) + \left[\int_{t_i}^{t_{i+1}} \Phi(t_{i+1}, \tau) B(\tau) d\tau \right] u(t_i) \quad (\text{A-11})$$

のように表わせる。したがって、(A-11) と (A-8) から、不連続時間の定差方程式システム・モデルは

$$x(i+1) = \Phi(i+1, i) x(i) + B(i) u(i) \quad (\text{A-12a})$$

$$Z(i) = H(i) x(i) \quad (\text{A-12b})$$

と表わされる。なお i は時点を示すものである。

付—2 Bayes 推定法

Bayes 推定法はパラメーターを固定された未知の量と考えることはせず、分析者はパラメーターの値に関してある先験的な判断をもち、またその判断は確率によって表現できるという仮定の下ですすめられる。特に、パラメーターは確率変数であると考えられ、また分析者はその先験的な判断を数量的に表現しているパラメーターについての先験的確率分布を仮定できるとしている。

いま、 $X = (X_1, \dots, X_n)$ を未知のパラメーター・ベクトル θ によって規定されている密度 $f(x/\theta)$ からのランダム・サンプルとする。そして、その未知のパラメーター・ベクトル θ を推定することを考える。ここで、 $P(\theta)$ を θ の周辺密度そして $l(\hat{\theta}; \theta)$ を損失関数とする。たとえば、 θ をランダム変数であると仮定しても、 θ の特定の値を推定しようとするのである。すなわちその θ の値は、ランダム・サンプルが抽出された密度 $f(x/\theta)$ を決定するものである。換言すれば、 θ が変化すれば異なった密度が定められることになり、この密度を決定する θ の値を推定するのである。

損失関数の期待値であるリスク関数を $E[l(\theta; \theta)] = R(d, \theta)$ と表現す

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

る。 θ がランダム変数であるから，“期待リスク”を最小とするような関数 d を見い出すことが関心事である。その期待リスクは次のように書くことができる。

$$B(d) = E[R(d, \theta)] = \int_{-\infty}^{\infty} R(d, \theta) P(\theta) d\theta \quad (B-1)$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} l[d(x_1, \dots, x_n); \theta] g(x_1, \dots, x_n/\theta) dx_1 \cdots dx_n \right\} p(\theta) d\theta \quad (B-2)$$

良好な特性をもつ推定量は $B(d)$ を最小とする x_i に関する関数 d である。この関数を“Bayes 推定量”という。

(B-2) の x_i と θ についての積分の順序を取り換えると、

$$B(d) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} l[d(x_1, \dots, x_n); \theta] g(x_1, \dots, x_n/\theta) \cdot p(\theta) d\theta \right\} dx_1 \cdots dx_n \quad (B-3)$$

のようになる。 x のあらゆる集合について、(B-3) の $\{ \}$ 内の量を最小とする x_i に関する関数 d が見い出されると、 $B(d)$ が最小となる。すなわち、次の式で表わされる量

$$\int_{-\infty}^{\infty} l[d(x_1, \dots, x_n); \theta] g(x_1, \dots, x_n/\theta) P(\theta) d\theta \quad (B-4)$$

を最小とする $d(x_1, \dots, x_n)$ を見い出すことを要す。(B-4) における $g(x_1, \dots, x_n/\theta) \cdot P(\theta)$ は x_1, \dots, x_n, θ に関する結合分布である。これを $q(x_1, \dots, x_n, \theta)$ で表わすと、 x に関する周辺分布は

$$\begin{aligned} h(x_1, \dots, x_n) &= \int_{-\infty}^{\infty} q(x_1, \dots, x_n, \theta) d\theta \\ &= \int_{-\infty}^{\infty} g(x_1, \dots, x_n/\theta) \cdot P(\theta) d\theta \end{aligned} \quad (B-5)$$

と与えられる。また X_1, X_2, \dots, X_n が与えられたときの θ 条件付分布は

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）

$$h(\theta/x_1, \dots, x_n) = \frac{q(x_1, \dots, x_n, \theta)}{k(x_1, \dots, x_n)}$$

であるから結局

$$h(\theta/x_1, \dots, x_n) = \frac{g(x_1, \dots, x_n/\theta) P(\theta)}{k(x_1, \dots, x_n)} \quad (\text{B-6})$$

となり、これを“事後密度”と称す。(B-6) の関係を用いると (B-4) は

$$k(x_1, \dots, x_n) \int_{-\infty}^{\infty} l[d(x_1, \dots, x_n; \theta)] h(\theta/x_1, \dots, x_n) d\theta \quad (\text{B-7})$$

と書き換えることができる。

かくして、Bayes 推定量は、それぞれの可能なサンプルについて、次の量

$$v(\hat{\theta}; x_1, \dots, x_n) = \int_{-\infty}^{\infty} l(\hat{\theta}; \theta) h(\theta/x_1, \dots, x_n) d\theta \quad (\text{B-8})$$

を最小にする値 $\hat{\theta}$ である。また関数 v は、 $X_1=x_1, \dots, X_n=x_n$ と与えられ、 θ を推定するための“事後損失”を表わしている。

Bayes 推定法の考え方では、事後密度（事後分布）はサンプルおよび分析者の先験的判断によって与えられるすべての情報を含んでいることになる。またこの推定法では、パラメーター θ はランダム変数として取り扱う。 θ の Bayes 推定量を $\hat{\theta} = \hat{\theta}(X)$ と書くと、推定は事後密度（事後分布） $h(\theta/X)$ を基礎にし、かつサンプル X を条件として行われるので、 $\hat{\theta}$ が固定されることになる。これは最尤法と異なる点である。次に Bayes 推定法では損失関数 $l(\theta; \theta)$ を定式化し、これを事後密度（事後分布）についての期待値である事後リスクを最小にする $\hat{\theta}$ を求めることになる。

いま、損失関数を推定誤差 $(\hat{\theta} - \theta)$ について、行列 Q ($Q \geq 0$) を有す二次形式、すなわち $(\hat{\theta} - \theta)' Q (\hat{\theta} - \theta)$ 、の場合を考える。また、 ξ をもって

時変パラメーター回帰モデルに対するカルマン・フィルターの応用（原田）
事後密度（事後分布）についての期待値をとる演算子とし，次の損失関数の期待値（リスク）を考える。

$$\begin{aligned} & \xi[(\hat{\theta}-\theta)' Q(\hat{\theta}-\theta)] \\ &= \xi[\{\hat{\theta}-\xi\theta-(\theta-\xi\theta)\}' Q\{\hat{\theta}-\xi\theta-(\theta-\xi\theta)\}] \\ &= (\hat{\theta}-\xi\theta)' Q(\hat{\theta}-\xi\theta) + \xi[(\theta-\xi\theta)' Q(\theta-\xi\theta)] \quad (\text{B-9}) \\ & \therefore \xi[(\theta-\xi\theta)' Q(\hat{\theta}-\xi\theta)] = \alpha Q(\hat{\theta}-\xi\theta) = 0 \end{aligned}$$

(B-4) の最後のラインにおける第二項は $\hat{\theta}$ を含まないので，第一項のみでリスクの最小を考えればよい。リスクを最小とする $\hat{\theta}$ は

$$\begin{aligned} & \partial(\hat{\theta}-\xi\theta)' Q(\hat{\theta}-\xi\theta) / \partial \hat{\theta} = 0 \\ & 2Q(\hat{\theta}-\xi\theta) = 0 \\ & \hat{\theta} = \xi\theta \quad \therefore Q \geq 0 \end{aligned}$$

と求められる。このことは，推定量 $\hat{\theta}$ は事後密度（事後分布）の平均値として与えられることを示している。

Bayes 推定量は損失関数を推定誤差の二次形式で与えた場合，事後密度の平均値として与えられ，かつ最小分散（(B-9) の最後のラインの第一項は推定誤差の分散である）推定量，すなわち最良推定量となる。また，推定量が事後密度の平均で与えられることから，不偏性を保持することになる。

参 考 文 献

- [1] Athans, M., "The Importance of Kalman Filtering Methods for Economic Syetems," Annals of Economic and Social Measurement, 3(1974), 49—64.
- [2] Duncan, D. B. and S. D. Horn, "Linear Dynamic Regression Estimation from the Viewpoint of Regression Analysis," Journal of the American Statistical Association, 67 (1972), 815—821.
- [3] Hatanaka, M., "A Note on the Application of the Kalman Filter to

時変パラメーター回帰モデルに対するカルマン・フィルターの応用 (原田)

Regression Models with Some Parameters Varying Over Time and Others Unvarying," Australian Journal of Statistics, 22 (1980), 298—306.

- [4] Ho, Y. C. and R. C. K. Lee, "A Bayesian Approach to Problems in Stochastic Estimation and Control," IEEE Transactions on Automatic Control, 9 (1964), 333—339.
- [5] Jazwinski, A. H., Stochastic Processes and Filtering Theory, Academic Press, 1970.
- [6] Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problem," Transactions of ASME, Journal of Basic Engineering, 85 D (1960), 35—45.
- [7] Maybeck, P. S., Stochastic Models, Estimation and Control, Vol. 1, Academic Press, 1979.
- [8] Mood, A. M. and F. A. Graybill, Introduction to the Theory of Statistics, McGraw-Hill, 1963.
- [9] Pagan, A., "Some Identification and Estimation Results for Regression Models with Stochastically Varying Coefficients," Journal of Econometrics, 13 (1980), 341—363.
- [10] Sant, D. T., "Generalized Least Squares Applied to Time Varying Parameter Models," Annals of Economic and Social Measurement, 6(1977), 301—311.
- [11] Sarris, A. H., "A Bayesian Approach to Estimation of Time-Varying Regression Coefficients," Annals of Economic and Social Measurement, 2(1973), 501—523.
- [12] Theil, H., Principles of Econometrics, North-Holland, 1971.

(1982年11月)