論文・研究レポート

生態学的予測モデルに対する カルマンフィルタによる近似法の研究

上田　徹，中村　剛

1. まえがき

カルマンフィルタ[1, 5]は予測誤差を調整しながら予測を進めるため短期間の予測では大きな誤差を避けることができる．また，ロジスティックモデルやBassモデルのような生態学的予測モデル[2]はモデルの係数に定性的な意味があり，長期の構造を説明するのに適しているが，短期的には大きな誤差が生じる可能性がある．そこで，これらの方法の長所だけを組み合わせることができないかを検討する．すなわち，カルマンフィルタを適用できるモデルは基本的には線形モデルであるが，カルマンフィルタを非線形モデルの近似として適用できないかを検討する．そして，2種の競合関係を考慮したモデル[3, 4, 6]にカルマンフィルタを適用できるかも検討する．

2. カルマンフィルタの概要

時点nでのトレンドや季節性などに関する状態を表す変数$\mathbf{x}(n)$を推定・予測したいものとする．時点nから時点$(n+1)$への状態変化は，雑音$U(n)$を考慮して構造（システム）方程式

$$\mathbf{x}(n+1)=F(n)\mathbf{x}(n)+G(n)\mathbf{u}(n) \tag{1}$$

で与えられるものとする[$n=1, 2, \cdots, N$]．時点nでの観測値$y(n)$は，トレンド成分$T(n)$と周期成分$S(n)$と雑音$w(n)$[分散ω^2]の和，すなわち観測方程式

$$y(n)=T(n)+S(n)+w(n)=H(n)^t x(n)+w(n) \tag{2}$$

で表されるものとする（式(5)の$\mathbf{x}(n)$，式(6)の$H(n)$参照）．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い，カルマンフィルタは状態$\mathbf{x}(n)$を逐次的に推定するアルゴリズムである．

トレンド成分，周期成分に関しては

$$\nabla^2 T(n) \equiv \{T(n)-T(n-1)\}-\{T(n-1)-T(n-2)\} = u(n) \tag{3}$$

$$\sum_{i=0}^{s-1} S(n-i)=v(n) \tag{4}$$

が成り立つものとする．ただし，$u(n)$, $v(n)$は互いに独立かつ平均0で，それぞれ分散τ^2, σ^2（未知）の正規性ホワイトノイズと仮定し，周期成分の周期はsとする．式(3)のトレンド成分表現を滑らかなトレンドモデルと呼ぶ．このとき，

$$\mathbf{x}(n)=(T(n), T(n-1), S(n), \cdots, S(n-s+2))^t \tag{5}$$

と置くと式(1), (2)の$F(n)$, $G(s)$, $H(n)$, $\mathbf{u}(n)$および$U(n)$の分散共分散行列$Q(n)$は

$$F=\begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & -1 & \cdots & -1 & -1 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

$$G=\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix}, \quad H(n)=(1, 0, 1, 0, \cdots, 0)^t, \quad \mathbf{u}(n)=(u(n), v(n))^t, \quad Q=\begin{bmatrix} \tau^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \tag{6}$$

である．推定すべきパラメタ$\theta=(\omega^2, \tau^2, \sigma^2)$のうち$\omega^2$は1と置くことができ，その他は最尤法を用いて推定できる．

時点n_2までの観測値を用いた$z(n_1)$の推定値を$z(n_1|n_2)$とすると，（順方向では）

うえだ　とおる，なかむら　つよし（現在，NEC）
成蹊大学 工学研究科
〒180-8633 武蔵野市吉祥寺北町 3-3-1
受付02.4.8　採択02.11.7

【時間更新】

$$y(n|n-1)=H(n)^t\mathbf{x}(n|n-1)$$
$$R(n|n-1)\equiv Var[y(n|n-1)]$$
$$=H(n)^tP(n|n-1)H(n)+\omega^2$$
$$\mathbf{x}(n|n-1)=F\mathbf{x}(n-1|n-1)$$
$$P(n|n-1)\equiv Var[\mathbf{x}(n|n-1)]$$
$$=FP(n-1|n-1)F^t+GQG^t \tag{7}$$

【観測値による更新】

$$\mathbf{x}(n|n)=\mathbf{x}(n|n-1)+K_n\{y(n)-H(n)^t\mathbf{x}(n|n-1)\}$$
$$P(n|n)=P(n|n-1)-K_nH(n)^tP(n|n-1)\}$$
$$K_n=\frac{P(n|n-1)H(n)}{H(n)^tP(n|n-1)H(n)+\omega^2}$$

を通じて予測値 $y(n|n-1)$ を逐次，得ることができる．ただし，$Var(X)$ は X の分散を意味する．ここで，初期値 $\mathbf{x}(0|0)$ はカルマンフィルタを最終時点から逆向きに走らせることにより推定できる．逆向きカルマンフィルタでも $F, G, H, \mathbf{u}$ は変わらないが，状態ベクトルは

$$\mathbf{x}(n)=(T(n), T(n+1), S(n), \cdots, S(n+s-2))$$

としなければならない．

3. 各種モデルのカルマンフィルタによる近似

ここでは生態学的予測モデルを用いてトレンド項 $T(n)$ を近似することを考える．

3.1 ロジスティックモデル[2]

需要予測によく使われるモデルであるロジスティックモデルは次式で表される．

$$\frac{dT(t)/dt}{T(t)}=a_1\left(1-\frac{T(t)}{m}\right),\ m：総（最終）需要$$

上式の微分を差分で表現し，時点 n での増加率は時点 $(n-1)$ で決まるとすると，

$$T(n+1)-T(n)=a_1T(n)\left(1-\frac{T(n-1)}{m}\right) \tag{9}$$

とする近似が考えられる．そして

$$c_{n-1}=1-\frac{T(n-1)}{m} \tag{10}$$

と置きかえると

$$T(n+1)=(a_1c_{n-1}+1)T(n) \tag{11}$$

と表すことができる．時点 n で c_{n-1} を確定値として扱うと $T(n)$ を定数倍したものが $T(n+1)$ ということになる．

次に，カルマンフィルタで初期値を求めるために，逆向きカルマンフィルタ用のシステムを変更する．式(9)と同様に

$$T(n)-T(n-1)=a_2T(n)\left(1-\frac{T(n+1)}{m}\right) \tag{12}$$

式(10)と同様に

$$c_{n+1}=1-\frac{T(n+1)}{m} \tag{13}$$

と置き換えて

$$T(n-1)=(1-a_2c_{n+1})T(n) \tag{14}$$

とする．

式(11)，(14)に誤差項を加えたものを，式(3)および逆向きカルマンフィルタのトレンド項表現の代わりに用いる．すなわち，式(6)における行列 $F(n)$ の左上の 2×2 行列の部分

$$F_C(n)=\begin{bmatrix}2 & -1\\ 1 & 0\end{bmatrix}$$

を，順方向のカルマンフィルタでは n に依存する

$$\begin{bmatrix}1+a_1c_{n-1} & 0\\ 1 & 0\end{bmatrix}$$

に置き換え，逆方向のカルマンフィルタでは

$$\begin{bmatrix}1-a_2c_{n+1} & 0\\ 1 & 0\end{bmatrix}$$

に置き換える．

3.2 Bass モデル I

新製品拡散モデルとして知られている Bass モデル[2, 3]は，m を総（最終）需要とすると

$$dT(t)/dt=\{m-T(t)\}\{\alpha_1+(\beta_1/m)T(t)\}$$

と表現される．ロジスティックモデルと同様に左辺の微分を差分で表現し，右辺の時点調整を行うと

$$T(n+1)-T(n)$$
$$=\{m-T(n-1)\}\{\alpha_1+(\beta_1/m)T(n)\} \tag{15}$$

となる．ここで

$$c_{n-1}=m-T(n-1) \tag{16}$$

と置き換えると

$$T(n+1)=\{1+(\beta_1/m)c_{n-1}\}T(n)+\alpha_1c_{n-1} \tag{17}$$

となる．

次に逆向きカルマンフィルタのシステムも

$$T(n)-T(n-1)$$
$$=\{m-T(n+1)\}\{\alpha_2+(\beta_2/m)T(n)\} \tag{18}$$

と変更し，

$$T(n-1)=\{1-(\beta_2/m)c_{n+1}\}T(n)-\alpha_2c_{n+1} \tag{19}$$

とする．

式(17)，(19)を式(11)，(14)の代わりに用いる．すなわち，$F_c(n)$ の代わりに，順方向のカルマンフィルタでは

$$\begin{bmatrix}1+(\beta_1/m)c_{n-1} & 0\\ 1 & 0\end{bmatrix}$$

を用い，逆方向のカルマンフィルタでは

$$\begin{bmatrix} 1-(\beta_2/m)c_{n+1} & 0 \\ 1 & 0 \end{bmatrix}$$

を用いる．

また，式(17)，(19)における右辺第2項のために，

$$\mathbf{x}(n+1)=F(n)\mathbf{x}(n)+C(n)+G(n)U(n)$$
$$C(n)^t=(d_n, 0, 0, \cdots, 0)$$

であり，式(17)では

$$d_n=\alpha_1 c_{n-1}$$

式(19)では

$$d_n=-\alpha_2 c_{n+1}$$

である．

3.3 Bass モデル II

式(17)の右辺第2項は $T(n-1)$ の1次式なのでそのまま用いることとし，

$$T(n+1) \\ =\{1+(\beta_1/m)c_{n-1}\}T(n)+\alpha_1\{m-T(n-1)\} \quad (20)$$

と置く．

逆向きカルマンフィルタもシステムを変更し，

$$T(n-1) \\ =\{1+(\beta_2/m)c_{n+1}\}T(n)+\alpha_2\{m-T(n+1)\} \quad (21)$$

と置く．

式(20)，(21)を式(17)，(19)の代わりに用いる．すなわち，$F_c(n)$ の代わりに，順方向のカルマンフィルタでは

$$\begin{bmatrix} 1+(\beta_1/m)c_{n-1} & -\alpha_1 \\ 1 & 0 \end{bmatrix}$$

を用い，逆方向のカルマンフィルタでは

$$\begin{bmatrix} 1-(\beta_2/m)c_{n-1} & \alpha_2 \\ 1 & 0 \end{bmatrix}$$

を用いる．

3.4 競合を考慮したモデル

競合の起きる2種の観測値を $T_1(n), T_2(n)$ とすると，競争的2種モデルは

$$dT_1/dt=T_1(a-bT_1-cT_2)$$
$$dT_2/dt=T_2(d-eT_1-fT_2)$$

となる．ここでも微分を差分で表現し，右辺の(　)内は1時点前の値で決まるとすると

$$T_1(n+1)-T_1(n) \\ =T_1(n)\{a-bT_1(n-1)-cT_2(n-1)\} \quad (22)$$
$$T_2(n+1)-T_2(n) \\ =T_2(n)\{d-eT_1(n-1)-fT_2(n-1)\} \quad (23)$$

となる．ここで，

$$\lambda_{1,\,n-1}=a-T_1(n-1)-cT_2(n-1)$$
$$\lambda_{2,\,n-1}=d-eT_1(n-1)-fT_2(n-1)$$

と置くと，

$$T_1(n+1)=T_1(n)\{1+\lambda_{1,\,n-1}\}$$
$$T_2(n+1)=T_2(n)\{1+\lambda_{2,\,n-1}\}$$

となる．そこで，状態ベクトルを

$$\mathbf{x}(n)=[\mathbf{x}_1(n), \mathbf{x}_2(n)]$$
$$\mathbf{x}_i(n)=[T_i(n), T_i(n-1), S_i(n), \cdots, S_i(n-s+2)]$$

のように節3.3までの状態ベクトルを二つ繋げ，$F(n)$ も

$$F(n)=\begin{bmatrix} F_1(n) & \mathbf{0} \\ \mathbf{0} & F_2(n) \end{bmatrix}$$

$$F_i(n)=\begin{bmatrix} 1+\lambda_{i,\,n-1} & 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & -1 & \cdots & -1 & -1 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

とし，他の行列，ベクトルも行，列の長さを同様に倍長にすればよい．なお，$\mathbf{0}$ は $F_i(n)$ と同じ大きさで要素がすべて0の行列である．

4. モデル評価

4.1 評価に用いるデータと尺度

本節では総務庁統計局作成の「家計調査年報」の「1世帯当たり1ヶ月の品目別購入数量（全世帯）」の酒類項目における「ビール」と「清酒」のデータを用いる．1983～1997年のデータをパラメタ推定用に用い，1998年のデータを予測誤差評価用に用いる（$n=1$：1983年1月，$n=N\equiv180$：1997年12月）

二乗誤差和が最小となるようにパラメタ推定を行った．例えばロジスティックモデルでは，

$$\varepsilon(n)=T(n+1)-T(n)-a_1T(n)\{1-T(n)/m\}$$

として，1983～1997年のデータに対して

$$E_0=\sum_{n=1}^{N-1}\varepsilon(n)^2$$

が最小となるように a_1, m を決定した．

他のモデルについても同様である．

なお，$T(n)$ は最終需要 m 固定時のカルマンフィルタのトレンド項を用いることもできるが，手順の簡易化のため，中心化12ヶ月移動平均

$$T(n+7)=\frac{1}{24}\sum_{i=1}^{12}\{y(n+i)+y(n+i+1)\}$$

を用いた．

カルマンフィルタ近似によるモデルの適合度を評価

するのに

$$E_1=\sum_{n=1}^{N-1}\{y(n+1|n)-y(n+1)\}^2$$

$$E_2=\sum_{n=1}^{N-2}\{y(n+2|n)-y(n+2)\}^2$$

$$E_f=\sum_{n=1}^{12}\{y(N+n|N)-y(N+n)\}^2$$

を用いる．ただし，

$$y(n+m|n)=H^t\mathbf{x}(n+m|n)=H^tF^m\mathbf{x}(n|n)$$

であり，F の要素には逐次，推定値を用いる．

4.2 評価結果

(1) ビールの場合

ロジスティックモデルでは，最小二乗誤差和 E_0 を実現するパラメタ値は

$$m=6700,\ a_1=0.00501,\ a_2=0.0044$$

であったが，そのとき，$E_0=173384$ であった．最終需要 m として切りのよい $m=10000$ を用いたときには

$$a_1=0.00302,\ a_2=0.0029,\ E_0=173556$$

であり，E_0 の差は小さいのでこれらの値もモデル比較時のパラメタ値として用いる．

節2のモデルをモデル1，節3.1，3.2，3.3のモデルをそれぞれモデル2，3，4と呼ぶ．モデル i の誤差 E_j を E_{ij} とし，各モデルの誤差 E_{i1}, E_{i2}, E_{if} を表1に示す．これから，モデル4（Bassモデル II）の E_2 を除き，モデル1（滑らかなトレンドモデル）よりも誤差が小さい．

モデル1とモデル i とのAICの差は

$$D_j=K_j\log(E_{1j}/E_{ij})+2(p_1-p_i)$$

となる．ただし，K_j は誤差 E_j を求める際のデータ数であり，p_i はモデル i のパラメタ数である．予測に使うパラメタは m と a_1 であるが，a_2 も含めて3個のパラメタが増加，すなわち，$(p_1-p_2=-3)$ と考えると，$(D_j=0)$ となる $R_{ij}=E_{1j}/E_{ij}$ の値は表2のようになる．表1の E_{if} を用いて計算された R_{if} の値を表3に示す．表3の値が表2の値よりも大きければAICの観点からモデル i の方がモデル1よりも優れていることになる．

表1，3より，節3.1～3.3で提案したモデルは滑らかなトレンドモデルに優るとも劣らない予測力を持っているといえる．

（清酒の場合）

清酒の場合には，表4から分かるように，E_{i1}，E_{i2} に関してはどのモデルも滑らかなトレンドモデルより劣っているが，最も重視したい E_{if} に関してはモデル3，4（Bassモデル I，II）ともにモデル1（滑らかなトレンドモデル）より改善されている．また，この例は過去のデータによく合うモデルが予測でも優れているとは限らないことを示している．

表1　各モデルの誤差（ビール）

	m	E_{i1}	E_{i2}	E_{if}
節2(i=1)		4918	4882	386
節3.1 (i=2)	6700	3699	3751	211
	10000	3705	3705	213
節3.2 (i=3)	6700	3718	3776	364
	10000	3713	3770	293
節3.3 (i=4)	6700	3720	4994	219
	10000	3713	4839	215

表2　$D_j=0$ となる R_{ij} の値

j	1	2	f
Rij	1.0341	1.0343	1.0869

表3　R_{ij} の値

j	1	2	f
R_{2j} (m=6700)	1.3295	1.3015	1.8294
R_{2j} (m=10000)	1.3274	1.3177	1.8122
R_{3j} (m=6700)	1.3228	1.2929	1.0604
R_{3j} (m=10000)	1.3245	1.2950	1.3174
R_{4j} (m=6700)	1.3220	0.9776	1.7626
R_{4j} (m=10000)	1.3245	1.0089	1.7953

表4　各モデルの誤差（清酒）

	E_{i1}	E_{i2}	E_{if}
節2 (i=1)	204	205	26.6
節3.1 (i=2)	211	221	26.8
節3.2 (i=3)	206	216	23.9
節3.3 (i=4)	218	255	23.1

表5　節3.4のモデルの場合

	E_1	E_f
ビール	3676	313.3
清酒	206.1	20.79

(ビールと清酒の競合考慮時)

ビールと清酒が競合しているものとして，節3.4の方法を用い，ビールと清酒のトレンド成分をそれぞれ $T_1(n)$, $T_2(n)$ とした結果，表5が得られた．

ビール単独の場合と比較すると，節2の（滑らかなトレンド）モデルよりはよいが，節3.1の（ロジスティック）モデルとは E_f に関して劣っている．清酒単独の場合と比較すると，節3.1の（ロジスティック）モデルよりはよいが，節2の（滑らかなトレンド）モデルとは E_1 に関して劣っている．

5. まとめ

パラメタ固定時の二乗誤差の例 $E_0=173556$ と比べると E_1 の値ははるかに小さくカルマンフィルタによる精度向上が明白となった．また，節3で提案したモデルは既存の滑らかなトレンドモデルに優るとも劣らない結果が得られたのではないかと思う．特に，E_2, E_f に関してトレンドモデルよりも誤差が小さく，長期の予測への対応力がトレンドモデルよりも大きかったといえる．しかし，最後に提案した競合を考慮にいれたモデルは，モデルの複雑性が増す割には予測力，適合度が圧倒的に優れているわけではなく，それぞれ独立に扱ったモデルで十分と思われる．今回は「ビール」と「清酒」のデータでモデルを比較したが，他のデータでも適合度や予測力を比較する必要がある．また，競合型 Bass モデル[3, 4]についても同様のカルマンフィルタ近似が可能であるが，今回は比較対象にしなかった．同様のカルマンフィルタ近似は様々な非線形モデルに適用可能だと考えている．

なお，本研究は科学技術融合振興財団助成金による支援を受けたことを付記する．

参考文献

[1] 上田徹：「予測手法(1)：時系列予測法」オペレーションズ・リサーチ (1994.6).

[2] 上田徹：「予測手法(2)：生態学モデル」オペレーションズ・リサーチ (1994.7).

[3] Tohru Ueda: "Demand Forecasting Models for Markets with Competition", *ITC*-13, Jensen and V. B. Iversen (Editors), Elsevier Science Publishers, pp. 261-265 (1991).

[4] Tohru Ueda: "A Srudy of a Competitive Bass Model Which Takes into Account Competition among Firms", *Journal of the Operations Research Society of Japan*, Vol. 33, No. 4, pp. 319-334 (1990).

[5] 阿部威郎，上田徹：「状態空間モデルを用いた電話収入予測」，電子通信学会論文誌，Vol. J 68-A, No. 5, pp. 437-443 (1985).

[6] R. ハーバーマン著，稲垣宣生訳：「個体群成長の数学モデル」，pp. 130-138, 現代数学社 (1981).

[7]「家計調査年報」，総務庁統計局.