Ⅳ－7

カルマン・フィルター理論の人口推計問題への適用

信州大学工学部　学生員　○百瀬　輝幸
信州大学工学部　正　員　　奥谷　巌

1. はじめに

人口の将来推計において、実用にはいまだ最小二乗法による直線回帰が一般的である。実際長野県においては、県人口の推計値を直線回帰により出し、その結果を各生活圏別にシェア・トレンドしている。近年に至りようやくSystem-Dynamics等の手法により各種の要因を取り込む試みがなされるようになってきた。本研究においては、近隣数地区の人口を包括的に表現するモデルを仮定し、そのシステムパラメータを同定しているのだが、その手法として現代制御理論の分野で注目されているカルマン・フィルターを適用している。このカルマン・フィルターの適用により、状態推定系に誤差の観念を含有し、新しいデータが入るたびに逐次システムパラメータを「誤差分散最小」の意味で最適化し、変化する社会構造にモデルを追従させることが可能であると思われる。

2. モデル構造

N個の地区の過去m期の人口を用い現在の人口を表現するモデルを考える。すなわち、i地区のt期の人口はi地区ふびその回りの（N−1）地区の過去の人口の線形結合で現わされるとする。

$$ {}_ix_t=\sum_{k_1=1}^{m}a_{ik_1 1}x_{t-k_1}+\sum_{k_2=1}^{m}b_{ik_2 2}x_{t-k_2}+\cdots+\sum_{k_N=1}^{m}g_{ik_N N}x_{t-k_N}+\varepsilon_{it} \qquad (1)$$

(1)式において、${}_ix_t$：i地区t期人口，$a_{ik}\cdots g_{ik}$：システム・パラメータ，ε_{it}：i地区t期雑音である。(1)式はN地区まとめてマトリックス表示するのが簡便である。

$$\mathbb{X}_t=\mathbb{H}\cdot\mathbb{Y}_t+\mathbb{E}_t \qquad (2)$$

(2)式において

$$\mathbb{X}_t=[x_{1t},x_{2t},\cdots x_{Nt}]^T \qquad \mathbb{Y}_t=[x_{1t-1},x_{1t-2},\cdots x_{2t-1}\cdots x_{N,t-m}]^T$$

$$\mathbb{H}=\begin{bmatrix} a_{11}\,a_{12}\cdots a_{1m}\,b_{11}\cdots g_{1m}\\ a_{21}\,a_{22}\cdots a_{2m}\,b_{21}\cdots g_{2m}\\ \vdots \qquad\quad \vdots \quad \vdots \qquad \vdots \\ a_{N1}\,a_{N2}\cdots a_{Nm}\,b_{N1}\cdots g_{Nm}\end{bmatrix} \qquad \mathbb{E}_t=[\varepsilon_{1t},\varepsilon_{2t},\cdots,\varepsilon_{Nt}]^T$$

であって，$E\{\mathbb{E}_t\}=0$，$E\{\mathbb{E}_t^T\mathbb{E}_t\}=\mathbb{R}_{1t}$　とする。

以上モデル構造としては、人口の過去の値のみを取り上げてきたが、同様の形に、産業生活環境等の状態変量をシステムの入力として採用して取りあつかうことが可能である。

3. カルマン・フィルターによるシステムパラメータの同定と将来推計

(2)式をシステムパラメータに着目して書き改めると次式を得る。

$$\mathbb{X}_t=\mathbb{M}_t\cdot\mathbb{h}_t+\mathbb{E}_t \qquad (3)$$

(3)式において

$$\mathbb{M}_t=\begin{bmatrix}\mathbb{Y}_t^T & & \\ & \mathbb{Y}_t^T & \\ & & \ddots \\ & & & \mathbb{Y}_t^T\end{bmatrix} \qquad \mathbb{h}_t=[a_{11},a_{12},\cdots b_{11}\cdots g_{1m},a_{21}\cdots g_{Nm}]^T$$

である。

(3)式はカルマンフィルター適用における観測方程式と考えることができる。次にシステム方程式としてモデル(2)式が定常と仮定し、h_tに関する次式を得る。

$$h_{t+1} = \mathbb{I} \cdot h_t + w_t \tag{4}$$

(4)式において、$\mathbb{I}$:単位マトリックス，w_t:は期雑音であり、(2)式と同様に$E\{w_t\}=0$, $E\{w_t^T w_t\} = \mathbb{R}_{2t}$とする。以上(3)、(4)式によりカルマンフィルターの適用が可能である。すなわち、

$$\hat{h}_{t+1} = \mathbb{I} \cdot \hat{h}_t + \mathbb{K}_{t+1}[x_t - \mathbb{M}_t \cdot \hat{h}_t] \tag{5}$$

(5)式と推定器を仮定すると、(5)式の改良ゲイン（カルマンゲイン）$\mathbb{K}_{t+1}$は次の逐次式により計算されるのである。

$$\mathbb{Q} = \mathbb{P}_t + \mathbb{R}_{1t} \tag{6}$$

$$\mathbb{K}_{t+1} = \mathbb{Q}\mathbb{M}_t^T[\mathbb{M}_t \mathbb{Q} \mathbb{M}_t^T + \mathbb{R}_{2t}]^{-1} \tag{7}$$

$$\mathbb{P}_{t+1} = [\mathbb{I} - \mathbb{K}_{t+1} \cdot \mathbb{M}_t]\mathbb{Q} \tag{8}$$

(6)(7)(8)式において$\mathbb{P}_t$は予測誤差の分散共分散行列であり。簡単には$\mathbb{P}_{t0}=\mathbb{I}$, $\hat{h}_{t0}=0$ とし、$\mathbb{R}_1, \mathbb{R}_2$を適当に仮定し計算を進める。将来推計は以上によって同定したパラメータを以後不変として次式により行う。すなわち、現在よりk期先の人口は(3)式を用いて、

$$\hat{x}_{t+k} = \hat{\mathbb{M}}_{t+k} \cdot \hat{h}_t \tag{9}$$

で現わされ、$k=1$より(9)式を繰り返し用いることにより$\{\hat{x}_{t+1}, \hat{x}_{t+2} \cdots \hat{x}_{t+n}\}$の推定列が求められる。

4．実際の適用

カルマンフィルターは(2)(4)式のような線形なシステムに対し適用可能であるが、実際の適用においては、適当な基準値x_t^0 を導入し、その回りで非線形方程式を動的線形化した、残差分$\Delta x_t = x_t - x_t^0$に対するシステムに対し、上記カルマンフィルターを適用するのがより妥当である。

$$\Delta x_t = \mathbb{H}' \cdot \Delta \mathbb{Y}_t + \mathbb{E}'_t \tag{10}$$

厳密なx_t^0を導入することは、困難であるので${}_i x_t^0 = a_i t + b_i$，あるいは${}_i x_t^0 = a_i \mathrm{EXP}\{b_i t\}$の回帰式を用いる。また、逐次式の(5)式の初期値として、数期のデータにより最小二乗法を適用し以後フィルタリングを行うことも可能であり、このフィルタリングにより、社会構造の変化にもモデルを追従させる可能性を見いだすことができる。詳細については当日発表する予定であるが、これら各法を長野県内に適用した例を紹介し、包括的な人口推計を試みる。

5．おわりに

以上のようにカルマンフィルターは社会構造のようなソフトなシステムに適用可能であり、その効果も大きいと思われるのだが、反面問題も多く含んでいるのであって、すなわち才1には適用可能なデータが時系列に少ない事であり、才2にはカルマンフィルタの不安定性であり、才3には雑音分散の与え方である。しかしながら、これらの問題は今後時系列にデータが多く得られることにより、またカルマンフィルターを改良して用いる(Square Root法)により解決可能である。

参考　文献

相良「同定問題」計測と制御 Vol.8 No.4

宮武等「マクロ経済システムの同定と予測」計測自動制御学会論文集 11-6

日野・森・吉川「カルマンフィルターによる大気汚染の予測」土木学会論文報告集 224