第３章「ＡＲＭＡモデルに基づく予測」

１．最良線形不偏予測量 (BLUP: Best Linear Unbiased Predictor)

$\{y_t\}$ を $E(y_t)=\mu$ の定常過程として，$y(T)=(y_1,\cdots,y_T)'$ に基づく y_{T+1} の任意の予測量を $\tilde{y}_{T+1}=f(y(T))$ とする．

定義： $E(y_{T+1}-\tilde{y}_{T+1})=0$ (**予測誤差**の期待値が 0) のとき，$\tilde{y}_{T+1}$ は y_{T+1} の**不偏予測量**であるという．不偏予測量の中で，予測の平均２乗誤差 $MSE(\tilde{y}_{T+1})=E(y_{T+1}-\tilde{y}_{T+1})^2$ を最小にする予測量を**最良不偏予測量**（BUP: Best Unbiased Predictor）という．

Note: 予測量が不偏ならば，予測の MSE は予測誤差の分散と一致する．

定理１ BUP は条件付き期待値 $E(y_{T+1}|y(T))$ である．

[証明]

$$\begin{aligned}
MSE(\tilde{y}_{T+1}) &= E\left[(y_{T+1}-\tilde{y}_{T+1})^2\right] \\
&= E\left[\{y_{T+1}-E(y_{T+1}|y(T))-(\tilde{y}_{T+1}-E(y_{T+1}|y(T)))\}^2\right] \\
&= E\left[\{y_{T+1}-E(y_{T+1}|y(T))\}^2\right]+E\left[\{\tilde{y}_{T+1}-E(y_{T+1}|y(T))\}^2\right]
\end{aligned}$$

を得るから，これを最小にする $\tilde{y}_{T+1}$ は，明らかに $E(y_{T+1}|y(T))$ である．

例１： $\{y_t\}$ が平均 μ の正規定常過程であるとき，y_{T+1} の BUP は

$$E(y_{T+1}|y(T))=\mu+\Sigma_{12}\Sigma_{22}^{-1}(y(T)-\mu e) \tag{1}$$

であり，予測の MSE は

$$\begin{aligned}
E\left[\{y_{T+1}-E(y_{T+1}|y(T))\}^2\right] &= E\left[\left\{y_{T+1}-\mu-\Sigma_{12}\Sigma_{22}^{-1}(y(T)-\mu e)\right\}^2\right] \\
&= \Sigma_{11}-\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}
\end{aligned}$$

ただし，$e=(1,\cdots,1)':T\times 1$ であり，さらに，

$$\Sigma_{11}=V(y_{T+1})=\gamma_0,\quad \Sigma_{21}=Cov(y(T),y_{T+1})=\Sigma_{12}',\quad \Sigma_{22}=V(y(T)):T\times T.$$

Note: 無相関過程ならば，$E(y_{T+1}|y(T))=E(y_{T+1})=\mu$

定義： y_{T+1} の線形予測量 $a_0+a_1y_T+a_2y_{T-1}+\cdots+a_Ty_1$ の中で，予測の MSE を最小にする予測量を**最良線形不偏予測量 (BLUP)** といい，$\hat{y}_{T+1}$ で表す．

定理２ BLUP を与える a_0 と $a=(a_1,\cdots,a_T)'$ は

$$a_0=\mu\left(1-\sum_{t=1}^{T}a_t\right),\qquad a=\Gamma_T^{-1}\gamma(T) \tag{2}$$

である．ここで，

$$\Gamma_T = [((Cov(y_s, y_t)))] = [((\gamma_{s-t}))] : T \times T, \qquad \gamma(T) = (\gamma_1, \cdots, \gamma_T)'$$

[証明]

$$f(a_0, a) = E\left(y_{T+1} - a_0 - a_1 y_T - a_2 y_{T-1} - \cdots - a_T y_1\right)^2$$

を a_0 と a で偏微分して 0 とおく（正規方程式）ことにより，解を得る．なお，予測量の不偏性は，a_0 による偏微分

$$E\left[y_{T+1} - a_0 - \sum_{t=1}^{T} a_t y_{T+1-t}\right] = E\left[y_{T+1} - \hat{y}_{T+1}\right] = 0$$

から必然的に成り立つ（回帰モデルの LSE が不偏であることと同様）．

BLUP $\hat{y}_{T+1}$ の性質

(a) $\hat{y}_{T+1} = \mu + \sum_{t=1}^{T} a_t(y_{T+1-t} - \mu)$, $a = \Gamma_T^{-1}\gamma(T)$

(b) $V(y_{T+1} - \hat{y}_{T+1}) = \gamma_0 - \gamma(T)'\Gamma_T^{-1}\gamma(T)$

(c) 正規定常過程では BUP となる．

(d) $E[(y_{T+1} - \hat{y}_{T+1})y_t] = 0 \ (t = 1, \cdots, T)$

(e) BLUP は y_{T+1} を $1, y_T, \cdots, y_1$ で張られるベクトル空間に射影したものである（性質 $E[(y_{T+1} - \hat{y}_{T+1}) \times 1] = 0$ および (d) による）．

上の性質 (e) より，$\Pi(x|y)$ を射影演算子（Projection Operator) とするとき，次のことが成り立つ．

$$\hat{y}_{T+1} = \Pi(y_{T+1}|1, y_T, \cdots, y_1) \tag{3}$$

射影演算子の性質

Q, R を確率変数（分散有限），W を $T \times 1$ の確率ベクトル（共分散行列 Γ），$\alpha_1, \cdots, \alpha_T$, β を定数とする．

1. $\Pi(Q|W) = E(Q) + a'(W - E(W))$, ただし $\Gamma a = Cov(W, Q)$
2. $E[(Q - \Pi(Q|W))^2] = V(Q) - a'Cov(W, Q)$
3. $\Pi(\alpha_1 Q + \alpha_2 R + \beta|W) = \alpha_1\Pi(Q|W) + \alpha_2\Pi(R|W) + \beta$
4. $\Pi(\sum_{t=1}^{T} \alpha_t W_t + \beta|W) = \sum_{t=1}^{T} \alpha_t W_t + \beta$
5. $Cov(W, Q) = 0$ ならば $\Pi(Q|W) = E(Q)$

例2： AR(p) モデル $y_t = m + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$ では，(2) 式より, $T > p$ ならば

$$a = \Gamma_T^{-1}\gamma(T) = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad a_0 = \mu(1 - \sum_{t=1}^{T} a_t) = \mu(1 - \sum_{j=1}^{p} \phi_j) = m$$

を得る．他方，射影演算子を使うことにより，

$$\begin{aligned} \hat{y}_{T+1} &= \Pi(y_{T+1}|1,\, y_T,\, \cdots,\, y_1) \\ &= \Pi(m + \phi_1 y_T + \phi_2 y_{T-1} + \cdots + \phi_p y_{T+1-p} + \varepsilon_t|1,\, y_T,\, \cdots,\, y_1) \\ &= m + \phi_1 y_T + \phi_2 y_{T-1} + \cdots + \phi_p y_{T+1-p} \end{aligned}$$

このとき，

$$MSE(\hat{y}_{T+1}) = \gamma_0 - \gamma(T)'\Gamma_T^{-1}\gamma(T) = V(\varepsilon_{T+1}) = \sigma^2$$

例3： MA(1) モデル $y_t = m + \varepsilon_t - \theta_1 \varepsilon_{t-1}$ では，(2) を使って

$$\begin{aligned} \hat{y}_2 &= a_0 + a_1 y_1 \\ &= m(1 - a_1) + \frac{\gamma_1}{\gamma_0} y_1 = m(1 - \rho_1) + \rho_1 y_1 \\ &= m\left(1 - \frac{-\theta_1}{1 + \theta_1^2}\right) + \frac{-\theta_1}{1 + \theta_1^2} y_1 \\ &= m - \frac{\theta_1}{1 + \theta_1^2}(y_1 - m) \end{aligned}$$

$$MSE(\hat{y}_2) = \gamma_0 - \frac{\gamma_1^2}{\gamma_0} = \frac{1 + \theta_1^2 + \theta_1^4}{1 + \theta_1^2}\,\sigma^2 \tag{4}$$

注： MA モデルに関する予測は，推定の場合も同様であるが，AR モデルの場合よりも複雑である．予測の場合には，議論を簡単にするために，あたかも無限の過去からの観測値があるとして，$\hat{y}_{T+1} = \Pi(y_{T+1}|1,\, y_T,\, \cdots,\, y_1,\, \cdots)$ を考えることが多い．このとき，次の事実が成立する．

$$\Pi(\varepsilon_t|1,\, y_T,\, \cdots,\, y_1,\, \cdots) = \begin{cases} 0 & (t > T) \\ \varepsilon_t & (t \le T) \end{cases}$$

したがって，

$$\begin{aligned} \Pi(y_2|1,\, y_1,\, y_0,\, \cdots) &= \Pi(m + \varepsilon_2 - \theta_1 \varepsilon_1|1,\, y_1,\, \cdots) \\ &= m - \theta_1 \varepsilon_1 \end{aligned}$$

を得る．このとき，予測誤差分散は

$$MSE(\Pi(y_2|1,\, y_1,\, y_0,\, \cdots)) = V(\varepsilon_2) = \sigma^2$$

となり，上の (4) よりも小さくなる．

2．h 期先の予測 (h-step ahead prediction)

定理３ $y(T)=(y_1,\cdots,y_T)'$ に基づく h 時点先の値 y_{T+h} の BLUP は

$$\hat{y}_{T+h}=\Pi(y_{T+h}|1,\,y_T,\,\cdots,\,y_1)$$

で与えられる．

例４：AR(1) モデル $y_t=m+\phi_1 y_{t-1}+\varepsilon_t$ の場合

$$\begin{aligned} y_{T+h} &= m+\phi_1 y_{T+h-1}+\varepsilon_{T+h} \\ &= m+\phi_1(m+\phi_1 y_{T+h-2}+\varepsilon_{T+h-1})+\varepsilon_{T+h} \\ &= m(1+\phi_1+\cdots+\phi_1^{h-1})+\phi_1^h y_T+\varepsilon_{T+h}+\phi_1\varepsilon_{T+h-1}+\cdots+\phi_1^{h-1}\varepsilon_{T+1} \end{aligned}$$

であるから，

$$\begin{aligned} \hat{y}_{T+h} &= m(1+\phi_1+\cdots+\phi_1^{h-1})+\phi_1^h y_T \\ MSE(\hat{y}_{T+h}) &= V(\varepsilon_{T+h}+\phi_1\varepsilon_{T+h-1}+\cdots+\phi_1^{h-1}\varepsilon_{T+1}) \\ &= \sigma^2(1+\phi_1^2+\phi_1^4+\cdots+\phi_1^{2(h-1)}) \\ &\to \sigma^2\frac{1}{1-\phi_1^2}=\gamma_0 \quad (h\to\infty) \end{aligned}$$

例５： MA(1) モデル $y_t=m+\varepsilon_t-\theta_1\varepsilon_{t-1}$ では，$h\geq 2$ ならば，明らかに

$$\begin{aligned} \hat{y}_{T+h} &= \Pi(m+\varepsilon_{T+h}-\theta_1\varepsilon_{T+h-1}|1,\,y_T,\,\cdots,\,y_1) \\ &= m \quad (h\geq 2) \\ MSE(\hat{y}_{T+h}) &= V(\varepsilon_{T+h}-\theta_1\varepsilon_{T+h-1}) \\ &= \sigma^2(1+\theta_1^2)=\gamma_0 \quad (h\geq 2) \end{aligned}$$

3．射影の応用 (Application of projection)

再帰的射影（Recursive projection）

まず，y を $1,\,x_1,\,x_2$ に射影する（＝ $1,\,x_1,\,x_2$ に基づいて y を予測する）ことにより，次の表現が得られる．

$$y=\Pi(y|1,x_1,x_2)+\varepsilon=a_0+a_1x_1+a_2x_2+\varepsilon \tag{5}$$

ここで，

$$E(\varepsilon)=0,\qquad E(x_1\varepsilon)=0,\qquad E(x_2\varepsilon)=0$$

次に，(5) の両辺を 1 と x_1 に射影して，

$$\Pi(y|1,x_1)=a_0+a_1x_1+a_2\Pi(x_2|1,x_1) \tag{6}$$

を得る．(5) から (6) を差し引いて，

$$y - \Pi(y|1,x_1) = a_2\left(x_2 - \Pi(x_2|1,x_1)\right) + \varepsilon$$

右辺第１項と第２項は直交するから，

$$\Pi\left(y - \Pi(y|1,x_1)|x_2 - \Pi(x_2|1,x_1)\right) = a_2\left(x_2 - \Pi(x_2|1,x_1)\right)$$

となり，したがって

$$y = \Pi(y|1,x_1) + \Pi\left(y - \Pi(y|1,x_1)|x_2 - \Pi(x_2|1,x_1)\right) + \varepsilon \tag{7}$$

を得る．(5) と (7) から，

$$\Pi(y|1,x_1,x_2) = \Pi(y|1,x_1) + \Pi\left(y - \Pi(y|1,x_1)|x_2 - \Pi(x_2|1,x_1)\right)$$

一般に，次の recursive projection に関する定理が成り立つ．

定理４ Ω を確率変数の集合とするとき，

$$\Pi(y|\Omega,x) = \Pi(y|\Omega) + \Pi\left(y - \Pi(y|\Omega)|x - \Pi(x|\Omega)\right) \tag{8}$$

定理４から，次の**繰り返し射影の法則**（Law of iterated projections）が得られる

定理５

$$\Pi(\Pi(y|\Omega,x)|\Omega) = \Pi(y|\Omega)$$

[証明] (8) は

$$\Pi(y|\Omega,x) = \Pi(y|\Omega) + a(x - \Pi(x|\Omega))$$

と表すことができるから，この両辺を Ω に射影することにより結果を得る．ここで，次の事実を使った．

$$\Pi(x - \Pi(x|\Omega)|\Omega) = 0, \qquad \Pi(\Pi(y|\Omega)|\Omega) = \Pi(y|\Omega)$$

射影の考え方を使って，偏自己相関に関して次のことが成り立つ．

定理６ 定常過程 $\{y_t\}$ における時差 h の自己相関は，２つの予測誤差 $y_t - \Pi(y_t|y_{t+1},\ \cdots,\ y_{t+h-1})$ と $y_{t+h} - \Pi(y_{t+h}|y_{t+1},\ \cdots,\ y_{t+h-1})$ の自己相関係数に等しい．
[証明の方針] 時差 h の偏自己相関を ϕ_{hh} とするとき，次のことを証明すればよい．

$$\phi_{hh} = \frac{Cov(\tilde{y}_t, \tilde{y}_{t+h})}{\sqrt{V(\tilde{y}_t)V(\tilde{y}_{t+h})}}$$

ここで，

$$\begin{aligned} \tilde{y}_t &= y_t - \Pi(y_t|y_{t+1},\ \cdots,\ y_{t+h-1}) \\ \tilde{y}_{t+h} &= y_{t+h} - \Pi(y_{t+h}|y_{t+1},\ \cdots,\ y_{t+h-1}) \end{aligned}$$