2つの状態群をもつ潜在マルコフモデル

菊池泰樹（長崎大学）　西晃央（佐賀大学）　野間口謙太郎（高知大学）

1．モデルの紹介

本モデルでは, 時系列データ $y_1, y_2, \ldots, y_{2n}$ を扱う．システム状態 S_t $(t = 1, 2, \ldots, 2n)$ とデータ発生システムに関して以下を仮定する．

- S_t $(1 \le t \le 2n)$ は定常状態に達した斉時的マルコフ連鎖にしたがって推移する．
- S_t $(1 \le t \le 2n)$ がとるシステム状態は次の 2 群に分かれている．

$$E_1 = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{m_1}\},\ E_2 = \{\varepsilon_{m_1+1}, \varepsilon_{m_1+2}, \ldots, \varepsilon_{m_1+m_2}\}$$

- $1 \le t \le n$ に対して, S_{2t-1} は E_1 の, S_{2t} は E_2 の, いずれか 1 つの状態をとる．

$$\sum_{i=1}^{m_1} P(S_{2t-1} = \varepsilon_i) = 1, \quad \sum_{i=m_1+1}^{m_1+m_2} P(S_{2t} = \varepsilon_i) = 1$$

- マルコフ連鎖の推移確率行列

$$\varGamma = \begin{pmatrix} O & \varGamma_1 \\ \varGamma_2 & O \end{pmatrix},$$

ここで, $\varGamma_1$ は $m_1 \times m_2$ 行列, $\varGamma_2$ は $m_2 \times m_1$ 行列であり, O は零行列である．

- データ発生システム
  - ◦ $1 \le t \le n$ に対して

$$S_{2t-1} = \varepsilon_i \text{ のとき}, Y_{2t-1} \sim g(\cdot\,; \theta_i), \quad 1 \le i \le m_1,$$
$$S_{2t} = \varepsilon_i \text{ のとき}, Y_{2t} \sim g(\cdot\,; \theta_i), \quad m_1 + 1 \le i \le m_1 + m_2.$$

  - ◦ パラメータ

$$\varTheta = (\theta_1, \ldots, \theta_{m_1}, \theta_{m_1+1}, \ldots, \theta_{m_1+m_2})$$

$\zeta = (\varGamma_1, \varGamma_2, \varTheta)$ とおくと, 完全データ $x = (s, y)$ に対する, 定常分布を含まない対数尤度関数は次で与えられる．

$$\begin{aligned} \log f_\zeta(s, y) \ =\ & \sum_{i=1}^{m_1} \sum_{j=m_1+1}^{m_1+m_2} \left[\sum_{t=1}^{n} s_{2t-1,i} s_{2t,j} \right] \log \gamma_{ij} + \sum_{i=m_1+1}^{m_1+m_2} \sum_{j=1}^{m_1} \left[\sum_{t=1}^{n-1} s_{2t,i} s_{2t+1,j} \right] \log \gamma_{ij} \\ & + \sum_{i=1}^{m_1} \left[\sum_{t=1}^{n} s_{2t-1,i} \log g(y_{2t-1}\,; \theta_i) \right] + \sum_{i=m_1+1}^{m_1+m_2} \left[\sum_{t=1}^{n} s_{2t,i} \log g(y_{2t}\,; \theta_i) \right] \end{aligned} \tag{1}$$

2．EM algorithm

不完全データに対する推定法として広く利用されている EM algorithm によってパラメータの最尤推定値を求める．反復法である EM algorithm では, 初期値 $\hat{\zeta}^{(0)} = (\hat{\varGamma}_1^{(0)}, \hat{\varGamma}_2^{(0)}, \hat{\varTheta}^{(0)})$ から出発し

て, E(xpectation)-step と M(aximization)-step を, 交互に, 収束するまで繰り返す.

E-step では, ℓ 回の反復後の推定値 $\hat{\zeta}^{(\ell)}$ が得られたとき, 次の Q-function (Dempster, Laird and Rubin, 1977; 小西・越智・大森, 2008)

$$Q(\zeta \,|\, \hat{\zeta}^{(\ell)}) = E\left[\log f_\zeta(S, y) \,|\, y\,; \hat{\zeta}^{(\ell)}\right]$$

を計算する. (1) より, Q-function は次で与えられる.

$$\begin{aligned} Q(\zeta \,|\, \hat{\zeta}^{(\ell)}) \;=\; & \sum_{i=1}^{m_1} \sum_{j=m_1+1}^{m_1+m_2} \left[\sum_{t=1}^{n} \hat{v}_{2t-1,ij}^{(\ell)}\right] \log \gamma_{ij} + \sum_{i=m_1+1}^{m_1+m_2} \sum_{j=1}^{m_1} \left[\sum_{t=1}^{n-1} \hat{v}_{2t,ij}^{(\ell)}\right] \log \gamma_{ij} \\ & + \sum_{i=1}^{m_1} \left[\sum_{t=1}^{n} \hat{u}_{2t-1,i}^{(\ell)} \log g(y_{2t-1}\,; \theta_i)\right] + \sum_{i=m_1+1}^{m_1+m_2} \left[\sum_{t=1}^{n} \hat{u}_{2t,i}^{(\ell)} \log g(y_{2t}\,; \theta_i)\right] \end{aligned}$$

ここで, 条件付き期待値を次のようにおいた.

$$\begin{aligned} \hat{u}_{2t-1,i}^{(\ell)} &= E\left[S_{2t-1,i} \,|\, y; \hat{\theta}^{(\ell)}\right], \quad \hat{u}_{2t,i}^{(\ell)} = E\left[S_{2t,i} \,|\, y; \hat{\theta}^{(\ell)}\right], \\ \hat{v}_{2t-1,ij}^{(\ell)} &= E\left[S_{2t-1,i} S_{2t,j} \,|\, y; \hat{\theta}^{(\ell)}\right], \quad \hat{v}_{2t,ij}^{(\ell)} = E\left[S_{2t,i} S_{2t+1,j} \,|\, y; \hat{\theta}^{(\ell)}\right]. \end{aligned}$$

M-step では, $Q(\zeta \,|\, \hat{\zeta}^{(\ell)})$ を最大にする $\zeta = (\mathit{\Gamma}_1, \mathit{\Gamma}_2, \mathit{\Theta})$ を求め, 次の段階の推定値 $\hat{\zeta}^{(\ell+1)} = (\hat{\mathit{\Gamma}}_1^{(\ell+1)}, \hat{\mathit{\Gamma}}_2^{(\ell+1)}, \hat{\mathit{\Theta}}^{(\ell+1)})$ とする.

- $\sum_{j=m_1+1}^{m_1+m_2} \gamma_{ij} = 1$ のもとで, $Q(\zeta \,|\, \hat{\zeta}^{(\ell)})$ の第 1 項を最大にする γ_{ij} は次で与えられる.

$$\hat{\gamma}_{ij}^{(\ell+1)} \;=\; \frac{\sum_{t=1}^{n-1} \hat{v}_{2t-1,ij}^{(\ell)}}{\sum_{j'=m_1+1}^{m_1+m_2} \sum_{t=1}^{n-1} \hat{v}_{2t-1,ij'}} = \frac{\sum_{t=1}^{n-1} \hat{v}_{2t-1,ij}^{(\ell)}}{\sum_{t=1}^{n-1} \hat{u}_{2t-1,i}} \quad (1 \le i \le m_1,\, m_1+1 \le j \le m_1+m_2)$$

- $\sum_{j=1}^{m_1} \gamma_{ij} = 1$ のもとで, $Q(\zeta \,|\, \hat{\zeta}^{(\ell)})$ の第 2 項を最大にする γ_{ij} は, 次で与えられる.

$$\hat{\gamma}_{ij}^{(\ell+1)} \;=\; \frac{\sum_{t=1}^{n-1} \hat{v}_{2t,ij}^{(\ell)}}{\sum_{j'=1}^{m_1} \sum_{t=1}^{n-1} \hat{v}_{2t,ij'}} = \frac{\sum_{t=1}^{n-1} \hat{v}_{2t,ij}^{(\ell)}}{\sum_{t=1}^{n-1} \hat{u}_{2t,i}} \quad (m_1+1 \le i \le m_1+m_2,\, 1 \le j \le m_1)$$

- $Q(\zeta \,|\, \hat{\zeta}^{(\ell)})$ の第 3 項を最大にする θ_i は, 次を最大化することによって得られる.

$$\sum_{t=1}^{n} \hat{u}_{2t-1,i}^{(\ell)} \log g(y_{2t-1}\,; \theta_i) \quad (1 \le i \le m_1)$$

- $Q(\zeta \,|\, \hat{\zeta}^{(\ell)})$ の第 4 項を最大にする θ_i は, 次を最大化することによって得られる.

$$\sum_{t=1}^{n} \hat{u}_{2t,i}^{(\ell)} \log g(y_{2t}\,; \theta_i) \quad (m_1+1 \le i \le m_1+m_2)$$

参考文献

[1] Dempster, A. P., Laird, N. M. and Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion), *Journal of the Royal Statistical Society, Series B* **39**, 1–38.

[2] 小西貞則・越智義道・大森裕浩 (2008). 計算統計学の方法 ーブートストラップ・EM アルゴリズム・MCMC ー, 朝倉書店.