

補足資料: 遷移行列の推定

最終更新：2022年3月17日

経済セミナー連載「実証ビジネス・エコノミクス」

第7回「将来予想のインパクトを測る：

単独エージェント動学モデルの推定【基礎編】」

(2022年4・5月号)の付録

© 上武康亮・遠山祐太・若森直樹・渡辺安虎

1 はじめに

この補足資料では、状態変数が従うマルコフ過程の遷移行列を推定する方法を解説する。本資料では最尤法に基づき推定値および標準誤差を求める。なお、最尤法の解説については以下の教科書等を参考にされたい。

- 末石直也 (2015) 「計量経済学 ミクロデータ分析へのいざない」 日本評論社
- Greene, W. (2017) *Econometric Analysis*, Pearson

2 走行距離の遷移行列

走行距離 m_t は 21 のビンに分けられ、 $\kappa = (\kappa_1, \kappa_2)$ をパラメータとする以下の 21×21 の遷移行列で表されるマルコフ過程に従う。

$$P_m = \begin{pmatrix} 1 - \kappa_1 - \kappa_2 & \kappa_1 & \kappa_2 & 0 & \cdots & 0 \\ 0 & 1 - \kappa_1 - \kappa_2 & \kappa_1 & \kappa_2 & \cdots & 0 \\ 0 & 0 & 1 - \kappa_1 - \kappa_2 & \kappa_1 & \ddots & 0 \\ 0 & 0 & 0 & 1 - \kappa_1 - \kappa_2 & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 - \kappa_1 - \kappa_2 & \kappa_1 + \kappa_2 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

今回は、真の κ は $(\kappa_1, \kappa_2) = (0.25, 0.05)$ として、データを生成している。

ここで観察されたデータから最尤法を用いて κ を推定しよう。確率 $1 - \kappa_1 - \kappa_2, \kappa_1, \kappa_2, \kappa_1 + \kappa_2$ それぞれに対応する事象を A_1, A_2, A_3, A_4 としよう。そして、データセットにおける各観測を a_i とし、これは状態の遷移 (ある期における走行距離から、次の期における走行距離への遷移) を捉える。そして、 a_i は上記の各事象のどれか一つに該当する。

尤度関数は以下のように与えられる。

$$L_m(\kappa) = \prod_{i=1}^N \underbrace{(1 - \kappa_1 - \kappa_2)^{\mathbf{1}\{a_i=A_1\}} \kappa_1^{\mathbf{1}\{a_i=A_2\}} \kappa_2^{\mathbf{1}\{a_i=A_3\}} (\kappa_1 + \kappa_2)^{\mathbf{1}\{a_i=A_4\}}}_{\equiv f(a_i; \kappa)}$$

なお、 N はデータセットのサンプルサイズである。この尤度関数について対数をとった対数尤度関数は以下のように書き直すことができる。

$$\log L_m(\kappa) = \sum_{i=1}^N \{ \mathbf{1}\{a_i = A_1\} \log(1 - \kappa_1 - \kappa_2) + \mathbf{1}\{a_i = A_2\} \log \kappa_1 + \mathbf{1}\{a_i = A_3\} \log \kappa_2 + \mathbf{1}\{a_i = A_4\} \log(\kappa_1 + \kappa_2) \}$$

確率 $1 - \kappa_1 - \kappa_2, \kappa_1, \kappa_2, \kappa_1 + \kappa_2$ それぞれに対応する事象が実現したケースの数を x_1, x_2, x_3, x_4 とする。すなわち、 $x_j = \sum_{i=1}^N \mathbf{1}\{a_i = A_j\}$ である。このとき対数尤度関数は

$$\log L_m(\kappa) = x_1 \log(1 - \kappa_1 - \kappa_2) + x_2 \log \kappa_1 + x_3 \log \kappa_2 + x_4 \log(\kappa_1 + \kappa_2)$$

となる。この対数尤度関数の最大化の一階条件は

$$\frac{\partial}{\partial \kappa} \log L_m(\kappa) = \begin{pmatrix} -\frac{x_1}{1-\kappa_1-\kappa_2} + \frac{x_2}{\kappa_1} + \frac{x_4}{\kappa_1+\kappa_2} \\ -\frac{x_1}{1-\kappa_1-\kappa_2} + \frac{x_3}{\kappa_2} + \frac{x_4}{\kappa_1+\kappa_2} \end{pmatrix} = 0$$

であり、連立方程式を解いて

$$\hat{\kappa}_1 = \frac{x_2(x_2+x_3+x_4)}{(x_2+x_3)(x_1+x_2+x_3+x_4)}$$

$$\hat{\kappa}_2 = \frac{x_3(x_2+x_3+x_4)}{(x_2+x_3)(x_1+x_2+x_3+x_4)}$$

が κ の最尤推定量になる。

次に漸近分布について考えよう。最尤推定量の漸近正規性より、 $\sqrt{N}(\hat{\kappa} - \kappa) \xrightarrow{d} N(0, I_m(\kappa)^{-1})$ となる。ここで $I_m(\kappa)$ は情報行列 (information matrix) であり、以下として定義される。

$$I_m(\kappa) \equiv -\mathbb{E} \left[\frac{\partial^2 \log f(a_i; \kappa)}{\partial \kappa \partial \kappa'} \right]$$

$$= -\mathbb{E} \left[\begin{pmatrix} -\frac{\mathbf{1}\{a_i=A_1\}}{(1-\kappa_1-\kappa_2)^2} - \frac{\mathbf{1}\{a_i=A_2\}}{\kappa_1^2} - \frac{\mathbf{1}\{a_i=A_4\}}{(\kappa_1+\kappa_2)^2} & -\frac{\mathbf{1}\{a_i=A_1\}}{(1-\kappa_1-\kappa_2)^2} - \frac{\mathbf{1}\{a_i=A_4\}}{(\kappa_1+\kappa_2)^2} \\ -\frac{\mathbf{1}\{a_i=A_1\}}{(1-\kappa_1-\kappa_2)^2} - \frac{\mathbf{1}\{a_i=A_4\}}{(\kappa_1+\kappa_2)^2} & -\frac{\mathbf{1}\{a_i=A_1\}}{(1-\kappa_1-\kappa_2)^2} - \frac{\mathbf{1}\{a_i=A_3\}}{\kappa_2^2} - \frac{\mathbf{1}\{a_i=A_4\}}{(\kappa_1+\kappa_2)^2} \end{pmatrix} \right]$$

漸近共分散行列は $I_m(\kappa)^{-1}$ は母平均を標本平均で置き換えることで以下のように推定する。

$$\hat{I}_m(\hat{\kappa})^{-1} = \left[-\frac{1}{N} \sum_{i=1}^N \left(\frac{\partial^2 \log f(a_i; \hat{\kappa})}{\partial \kappa \partial \kappa'} \right) \right]^{-1}$$

$$= -N \begin{pmatrix} -\frac{x_1}{(1-\hat{\kappa}_1-\hat{\kappa}_2)^2} - \frac{x_2}{\hat{\kappa}_1^2} - \frac{x_4}{(\hat{\kappa}_1+\hat{\kappa}_2)^2} & -\frac{x_1}{(1-\hat{\kappa}_1-\hat{\kappa}_2)^2} - \frac{x_4}{(\hat{\kappa}_1+\hat{\kappa}_2)^2} \\ -\frac{x_1}{(1-\hat{\kappa}_1-\hat{\kappa}_2)^2} - \frac{x_4}{(\hat{\kappa}_1+\hat{\kappa}_2)^2} & -\frac{x_1}{(1-\hat{\kappa}_1-\hat{\kappa}_2)^2} - \frac{x_3}{\hat{\kappa}_2^2} - \frac{x_4}{(\hat{\kappa}_1+\hat{\kappa}_2)^2} \end{pmatrix}^{-1}$$

$\hat{\kappa}$ の漸近標準誤差は、 $\hat{I}_m(\hat{\kappa})^{-1} / N$ の対角成分の平方根として与えられる。

3 価格の遷移行列

価格 p は $\{2000, 2100, 2200, 2300, 2400, 2500\}$ のうち 1 つの値を取りうる。 p は λ をパラメータとする以下の遷移行列で表されるマルコフ過程に従う。

$$P_p = \begin{pmatrix} \lambda_{1,1} & \lambda_{1,2} & \lambda_{1,3} & \lambda_{1,4} & \lambda_{1,5} & \lambda_{1,6} \\ \lambda_{2,1} & \lambda_{2,2} & \lambda_{2,3} & \lambda_{2,4} & \lambda_{2,5} & \lambda_{2,6} \\ \lambda_{3,1} & \lambda_{3,2} & \lambda_{3,3} & \lambda_{3,4} & \lambda_{3,5} & \lambda_{3,6} \\ \lambda_{4,1} & \lambda_{4,2} & \lambda_{4,3} & \lambda_{4,4} & \lambda_{4,5} & \lambda_{4,6} \\ \lambda_{5,1} & \lambda_{5,2} & \lambda_{5,3} & \lambda_{5,4} & \lambda_{5,5} & \lambda_{5,6} \\ \lambda_{6,1} & \lambda_{6,2} & \lambda_{6,3} & \lambda_{6,4} & \lambda_{6,5} & \lambda_{6,6} \end{pmatrix}$$

ここで $\sum_{k=1}^6 \lambda_{j,k} = 1, j = 1, 2, 3, 4, 5, 6$ が成立することに注意する。 $j = 1, 2, 3, 4, 5, 6, k = 1, 2, 3, 4, 5, 6$ について $\lambda_{j,k}$ に対応する事象を $B_{j,k}$ としよう。そして、データセットにおける各観測を b_i とし、これは状態の遷移 (ある期における価格から、次の期における価格への遷移) を捉える。そして、 b_i は上記の各事象のどれか一つに該当する。

尤度関数は以下のように与えられる。

$$L_p(\lambda) = \underbrace{\prod_{i=1}^N \prod_{j=1}^6 \prod_{k=1}^6 \lambda_{j,k}^{\mathbf{1}\{b_i=B_{j,k}\}}}_{\equiv f(b_i; \lambda)}$$

$j = 1, 2, 3, 4, 5, 6, k = 1, 2, 3, 4, 5, 6$ について $\lambda_{j,k}$ に対応する事象が実現した観察の数を $y_{j,k}$ とする。すなわち、 $y_{j,k} = \sum_{i=1}^N \mathbf{1}\{b_i = B_{j,k}\}$ である。このとき対数尤度関数は

$$\log L_p(\lambda) = \sum_{j=1}^6 \sum_{k=1}^6 y_{j,k} \log(\lambda_{j,k})$$

となる。ここで

$$L_{pj}(\lambda_j) = \underbrace{\prod_{i=1}^N \prod_{k=1}^6 \lambda_{j,k}^{\mathbf{1}\{b_i=B_{j,k}\}}}_{\equiv f_j(b_i; \lambda_j)}, j = 1, 2, 3, 4, 5, 6$$

とすると $\log L_{pj}(\lambda_j) = \sum_{k=1}^6 y_{j,k} \log(\lambda_{j,k}), j = 1, 2, 3, 4, 5, 6$ が成り立ち、

$$\log L_p(\lambda) = \sum_{j=1}^6 \log L_{pj}(\lambda_j)$$

となる。 $\lambda_{j,j} = 1 - \sum_{k \neq j} \lambda_{j,k}$ であることを利用して、各 j について $\log L_{pj}(\lambda_j)$ を $\lambda_j = (\lambda_{j,k})_{k \neq j}$ により最大化すればよい。最大化の一階条件は

$$\frac{\partial}{\partial \lambda_j} \log L_{pj}(\lambda_j) = 0$$

で、計算すると $-\frac{y_{j,j}}{\lambda_{j,j}} + \frac{y_{j,k}}{\lambda_{j,k}} = 0, k \neq j$ となる。このとき最尤推定量は $\hat{\lambda}_{j,k} = \frac{y_{j,k}}{\sum_{k=1}^6 y_{j,k}}$ となる。

続いて漸近分布について考えよう。最尤推定量の漸近正規性より、 $\sqrt{N}(\hat{\lambda} - \lambda) \xrightarrow{d} N(0, I_p(\lambda)^{-1})$ となる。ここで情報行列 (information matrix) $I_p(\lambda)$ は以下のようなブロック対角行列として与えられる。

$$I_p(\lambda) \equiv -\mathbb{E} \begin{bmatrix} \frac{\partial^2 \log f_1(b_i; \lambda_1)}{\partial \lambda_1 \partial \lambda_1'} & & & & & O \\ & \frac{\partial^2 \log f_2(b_i; \lambda_2)}{\partial \lambda_2 \partial \lambda_2'} & & & & \\ & & \frac{\partial^2 \log f_3(b_i; \lambda_3)}{\partial \lambda_3 \partial \lambda_3'} & & & \\ & & & \frac{\partial^2 \log f_4(b_i; \lambda_4)}{\partial \lambda_4 \partial \lambda_4'} & & \\ & & & & \frac{\partial^2 \log f_5(b_i; \lambda_5)}{\partial \lambda_5 \partial \lambda_5'} & \\ O & & & & & \frac{\partial^2 \log f_6(b_i; \lambda_6)}{\partial \lambda_6 \partial \lambda_6'} \end{bmatrix}$$

対角要素における行列 $\frac{\partial^2 \log f_j(b_i; \lambda_j)}{\partial \lambda_j \partial \lambda_j'}$ は、 (5×5) の行列であり、以下として与えられる。

$$\frac{\partial^2 \log f_j(b_i; \lambda_j)}{\partial \lambda_j \partial \lambda_j'} = -\text{diag} \left(\left(\frac{\mathbf{1}\{b_i = B_{j,k}\}}{\lambda_{j,k}^2} \right)_{k \neq j} \right) - \frac{\mathbf{1}\{b_i = B_{j,j}\}}{\lambda_{j,j}^2} \cdot \mathbf{1}_{5 \times 5}$$

ここで $\mathbf{1}_{5 \times 5}$ は全ての要素が 1 であるような 5×5 の行列である。また、 $\text{diag} \left(\left(\frac{\mathbf{1}\{b_i = B_{j,k}\}}{\lambda_{j,k}^2} \right)_{k \neq j} \right)$ は対角要素が $\frac{\mathbf{1}\{b_i = B_{j,k}\}}{\lambda_{j,k}^2}$ ($k = 1, \dots, j-1, j+1, \dots, 6$) の値を取る 5×5 の対角行列である。

漸近共分散行列 $I_p(\lambda)^{-1}$ は母平均を標本平均で置き換えることで以下のように推定する

$$\hat{I}_p(\hat{\lambda})^{-1} = - \left[\text{diag} \left(\left(\frac{1}{N} \sum_{i=1}^N \frac{\partial^2 \log f_j(b_i; \hat{\lambda}_1)}{\partial \lambda_j \partial \lambda_j'} \right)_{j=1, \dots, 6} \right) \right]^{-1}$$

ここで $\text{diag} \left(\left(\sum_{i=1}^N \frac{\partial^2 \log f_j(b_i; \hat{\lambda}_1)}{\partial \lambda_j \partial \lambda_j'} \right)_{j=1, \dots, 6} \right)$ はブロック対角行列である。^{*1}また、このブロック対角要素における行列は

$$*1 \text{ 具体的には、} \text{diag} \left((A_j)_{j=1, \dots, 6} \right) = \begin{bmatrix} A_1 & & & & & O \\ & A_2 & & & & \\ & & A_3 & & & \\ & & & A_4 & & \\ & & & & A_5 & \\ O & & & & & A_6 \end{bmatrix} \text{ となる。}$$

$$\sum_{i=1}^N \frac{\partial^2 \log f_j(b_i; \hat{\lambda}_1)}{\partial \lambda_j \partial \lambda_j'} = -\text{diag} \left(\left(\frac{y_{j,k}}{\lambda_{j,k}^2} \right)_{k \neq j} \right) - \frac{y_{j,j}}{\lambda_{j,j}^2} \cdot \mathbf{1}_{5 \times 5}$$

という形で、上で定義した $\lambda_{j,k}$ に対応する事象が実現した観察の数 $y_{j,k}$ を用いて書くことができる。

以上を踏まえて、最終的に $\hat{\lambda}$ の漸近標準誤差は、 $\hat{I}_p(\hat{\lambda})^{-1} / N$ の対角成分の平方根として与えられる。