レジーム･スウィッチングモデル†

担当: 長倉大輔

† この資料は私のゼミにおいて使用するために作成した資料です。ゼミのWEBページ上で公開しており、自由に参照して頂いて構いません。ただし、内容について、一応検証してありますが、もし間違いがあった場合でもそれによって生じるいかなる損害、不利益について責任を負いかねますのでご了承ください。間違いは発見次第、継続的に直していますが、まだ存在する可能性があります。

レジーム・スウィチングモデル

- **マルコフ転換モデル(Markov switching model)**

“状態”が複数個あり、それぞれの状態においてモデルのパラメーターが異なるモデルを考えよう。例えば状態が m個あるような AR(p) モデルは以下のようになる。

$$y_t = \begin{cases} c_1 + \phi_{11} y_{t-1} + \ldots + \phi_{p1} y_{t-p} + \sigma_1 \varepsilon_t, & s_t = 1 \\ c_2 + \phi_{12} y_{t-1} + \ldots + \phi_{p2} y_{t-p} + \sigma_2 \varepsilon_t, & s_t = 2 \\ \vdots & \\ c_m + \phi_{1m} y_{t-1} + \ldots + \phi_{pm} y_{t-p} + \sigma_p \varepsilon_t, & s_t = m \end{cases}$$

$$\varepsilon_t \sim \text{iid}\ (0, 1),$$

ここで s_t は “状態変数”を表し、**状態 j にいる時、$s_t = j$ となるような変数**である。

レジーム・スウィチングモデル

状態変数 s_t は直接観測されない変数であり、状態変数 s_t が**マルコフ連鎖(Markov Chain; MC)**と呼ばれる離散型の確率過程に従うとき、そのモデルは**マルコフ転換(MS)モデル**と呼ばれる。先ほどのAR(p)モデルはマルコフ転換AR(p)(MSAR(p))モデルと呼ばれる。

状態は “レジーム (regime)”とも呼ばれ、複数の状態間を行き来するようなモデルはレジーム転換(regime switching)モデルと呼ばれる。マルコフ転換モデルはレジーム転換モデルの特殊な場合と考えることができる(しばしばレジーム転換モデルというとマルコフ転換モデルを指す)。

レジーム・スウィチングモデル

MSモデルは景気などの "状態" を直接観測できないようなものに対して、その状態のもとで、どのようにモデルに変化が起きているかを分析するのに便利である。

Hamilton (1989) がMSモデルによって景気循環をうまくとらえることができることを示して以来、経済やファイナンスのデータの分析にMSモデルがよく用いられるようになってきている。

レジーム・スウィチングモデル

- マルコフ連鎖

MSモデルにおいて s_t はマルコフ連鎖に従う確率過程であるとする。説明の簡単化のために状態は 2 つとする。

t 時点の状態変数の値は $s_t = j,\ j=1, 2$ という 2つの値のどちらかをとるとする。マルコフ連鎖では、過去の全ての状態変数の値に条件づけられた条件付き確率について

$$\Pr(s_t = j \mid s_{t-1} = i,\ s_{t-2} = k, ...) = \Pr(s_t = j \mid s_{t-1} = i)$$

を仮定する。つまり、**過去の全ての状態で条件付けした条件付確率が 1時点前の状態で条件付け下条件付確率と等しい**(1時点前の値がわかれば他は必要ない)という仮定である。

レジーム・スウィチングモデル

この時、j =1, 2について

$$\Pr(s_t = j) = \Pr(s_t = j \mid s_{t-1} = 1)\Pr(s_{t-1} = 1) + \Pr(s_t = j \mid s_{t-1} = 2)\Pr(s_{t-1} = 2)$$

が成り立つ。ここで $\Pr(s_t = j \mid s_{t-1} = i)$, i=1,2 は t に**依存しない**とし、$p_{ji} = \Pr(s_t = j \mid s_{t-1} = i)$、$p_{t,j} = \Pr(s_t = j)$ とすると

$$\begin{bmatrix} p_{t,1} \\ p_{t,2} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} p_{t-1,1} \\ p_{t-1,2} \end{bmatrix}$$

という関係がある事がわかる。

レジーム・スウィチングモデル

この時、p_{ji} を**推移確率、**

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

で定義される行列を**推移確率行列**という。確率の性質より、$p_{1i} + p_{2i} = 1$, i =1, 2, であるので $\mathbf{P}$ は

$$\mathbf{P} = \begin{bmatrix} p_{11} & 1 - p_{22} \\ 1 - p_{11} & p_{22} \end{bmatrix}$$

とも表せる。この場合、未知パラメーターは p_{11} と p_{22} となる。

レジーム・スウィチングモデル

(注意)文献によっては、

$$p_{ij} = \Pr(s_t = j \mid s_{t-1} = i)$$

と表記しているものもよくある(**下付き文字の *i* と *j* の順序が逆**)。その場合は先ほどのPは

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix}$$

となる。MS関連の文献を読む際は、条件付確率が p_{ij} か p_{ji} かどちらで表記されているか注意が必要。
　ここでは $p_{ji} = \Pr(s_t = j \mid s_{t-1} = i)$ を用いる。

レジーム・スウィチングモデル

状態の数が m 個の場合、$p_{t,i}$, p_{ji} , i, j=1,…,m を同様に定義すると、**推移確率**は、

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

となる。ここでそれぞれの i (=1,…., mのどれか) について

$$p_{1i} + p_{2i} + \ldots + p_{mi} = 1$$

という制約を満たすことに注意(列を足すと 1 になる)。

レジーム・スウィチングモデル

- マルコフ連鎖の定常分布

推移確率行列 P に対して、ベクトル $\boldsymbol{\pi} = [\,\pi_1, \ldots, \pi_m\,]'$が

$$\boldsymbol{\pi} = \mathbf{P}\boldsymbol{\pi}\ , \qquad \pi_1 + \ \ldots\ + \pi_m = 1$$

を満たす時、$\boldsymbol{\pi}$ はマルコフ連鎖の**定常確率**と呼ばれる。

定常確率が存在するためには **P** が特定の条件を満たさなくてはならない。例えば状態数が 2 つの場合、$p_{11} < 1$, $p_{22} < 1$, $p_{11} + p_{22} > 0$ であれば $\boldsymbol{\pi}$ が存在し、$\boldsymbol{\pi}$ は

$$\boldsymbol{\pi} = \left[\frac{1 - p_{22}}{2 - p_{11} - p_{22}}, \frac{1 - p_{11}}{2 - p_{11} - p_{22}} \right]'$$

で与えられる。

レジーム・スウィチングモデル

一般的には m 状態 マルコフ連鎖の場合、定常確率は

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_m - \mathbf{P} \\ \mathbf{1}'_m \end{bmatrix}$$

という$(m+1)\times m$ 行列 $\mathbf{A}$ に対して

$$(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$$

という $m\times(m+1)$ 行列の第 $m+1$ 列目によって与えられる。

　定常確率はマルコフ連鎖の初期状態の確率(無条件確率)として推定に用いられることが多い。

レジーム・スウィチングモデル

- 過去の2時点の状態に依存したモデル

推移確率は1時点前の状態がわかればそれ以前の状態には依存しないという仮定は、制約的に思えるが実はそうではない。

例えば状態数は2つだが、推移確率が過去の2時点前までの状態に依存している場合を考えよう。

$$\Pr(s_t = k \mid s_{t-1} = j, s_{t-2} = i) , k, j, i = 1, 2$$

この場合、連続した2時点が取りうるパターンは全部で $\{s_t, s_{t-1}\} = \{1,1\}, \{2,1\}, \{1,2\}, \{2,2\}$ の4つであるので、

レジーム・スウィチングモデル

新しく状態変数を

$$
\begin{aligned}
s_t^* &= 1, \quad \text{if} \quad s_t = 1, \quad s_{t-1} = 1 \\
s_t^* &= 2, \quad \text{if} \quad s_t = 2, \quad s_{t-1} = 1 \\
s_t^* &= 3, \quad \text{if} \quad s_t = 1, \quad s_{t-1} = 2 \\
s_t^* &= 4, \quad \text{if} \quad s_t = 2, \quad s_{t-1} = 2
\end{aligned}
$$

と定義すれば **4状態マルコフ連鎖**として表すことができる。
例えば

$$s_1 = 1,\ s_2 = 1,\ s_3 = 2,\ s_4 = 1,\ s_5 = 2\quad s_6 = 2$$

は

$$s_2^* = 1,\ s_3^* = 2,\ s_4^* = 3,\ s_5^* = 2,\ s_6^* = 4,$$

と表すことができる。

レジーム・スウィチングモデル

この時、s_t^* についての推移確率は

$$\Pr(s_t^* = 1 \mid s_{t-1}^* = 1) = \Pr(s_t = 1 \mid s_{t-1}=1, s_{t-2} =1),$$
$$\Pr(s_t^* = 2 \mid s_{t-1}^* = 1) = \Pr(s_t = 2 \mid s_{t-1}=1, s_{t-2} =1),$$
$$\Pr(s_{t*} = 3 \mid s_{t-1}^* = 1) = 0$$
$$\Pr(s_t^* = 4 \mid s_{t-1}^* = 1) = 0$$
$$\Pr(s_t^* = 1 \mid s_{t-1}^* = 2) = 0$$
$$\vdots$$
$$\Pr(s_t^* = 4 \mid s_{t-1}^* = 4) = \Pr(s_t = 2 \mid s_{t-1}=2, s_{t-2} =2)$$

のように求めることができる。$s_{t-1}^* =1$ $(s_{t-1}=1, s_{t-2} =1)$の時、$s_t^* = 3$ $(s_t =1,\ s_{t-1} = 2)$ は起こりえないので確率が 0になっている。このように推移確率に制約が入ることに注意。

レジーム・スウィチングモデル

- マルコフ連鎖の吸収状態

推移確率に制約を置くことによって、状態の推移をコントロールすることもできる。

例えば2状態モデルにおいて $p_{22}=1$ とすると、いったん状態2になると、以後は常に状態2をとるようになる(状態2から状態1に行く確率が 0 になるため)。この時、状態2は**吸収状態**と呼ばれる。このようなモデルはMSモデルで構造変化の分析を行う時などに使用される。

レジーム・スウィチングモデル

- フィルター化確率、平滑化確率

$\{y_1,\ldots,y_T\}$ というデータが与えられたとし、情報集合 $\Omega_t=\{y_1,\ldots,y_t\}$ を定義する。この時、**Ω_t という条件付き**で、$s_t=j$ となる確率(すなわち t 時点の状態が j である確率)

$$\Pr(s_t=j\,|\,\Omega_t)$$

を**フィルター化確率**という。また **Ω_T という条件付き**で、$s_t=j$ となる確率

$$\Pr(s_t=j\,|\,\Omega_T)$$

を**平滑化確率**という。

レジーム・スウィチングモデル

フィルター化確率は t 時点の状態を **t 時点までの観測値**に基づいて推測しているのに対して、平滑化確率は**全ての観測値**が与えられた時に、過去のある t 時点においてどの状態だったかを推測している。

フィルター化確率や平滑化確率は、直接観測できない状態について、与えられた情報から t 時点において、どの状態にどれくらいの確率でいるかを分析するときに有用である。例えば景気循環の分析などでは、t 時点において、経済の状態が景気拡大期と景気縮小期のどちらなのか（どちらだったのか）ということが分析できる。

レジーム・スウィチングモデル

- フィルター化確率および平滑化確率の計算

y_t の Ω_{t-1} という条件付きの密度関数を $f(y_t \mid \Omega_{t-1})$ とする。
フィルター化確率は

$$\Pr(s_t = j \mid \Omega_t) = \frac{\Pr(s_t = j \mid \Omega_{t-1}) f(y_t \mid \Omega_{t-1}, s_t = j)}{f(y_t \mid \Omega_{t-1})}$$

と求められる（詳しい説明は省略）。ここで分子の $f(y_t$
$|\Omega_{t-1})$ は

$$f(y_t \mid \Omega_{t-1}) = \sum_{i=1}^{m} \Pr(s_t = i \mid \Omega_{t-1}) f(y_t \mid \Omega_{t-1}, s_t = i)$$

と表すことができる。

レジーム・スウィチングモデル

$\Pr(s_t = j \mid \Omega_t), j = 1, \ldots, m$ が与えられれば $\Pr(s_{t+1} = k \mid \Omega_t)$は

$$\Pr(s_{t+1} = k \mid \Omega_t) = \sum_{j=1}^{m} \Pr(s_{t+1} = k, s_t = j \mid \Omega_t)$$

$$= \sum_{j=1}^{m} \Pr(s_{t+1} = k \mid s_t = j, \Omega_t) \Pr(s_t = j \mid \Omega_t)$$

$$= \sum_{j=1}^{m} \Pr(s_{t+1} = k \mid s_t = j) \Pr(s_t = j \mid \Omega_t)$$

$$= \sum_{j=1}^{m} p_{k,j} \Pr(s_t = j \mid \Omega_t)$$

と計算することができる。また、$\Pr(s_{t+1} = k \mid \Omega_t), k = 1, \ldots, m$ が計算できれば、$\Pr(s_{t+1} = k \mid \Omega_{t+1})$が前頁のように計算できる。

レジーム・スウィチングモデル

よって 初期値として $\Pr(s_j \mid \Omega_0) = \pi_j$ と与えてあげれば

$$
\begin{aligned}
\Pr(s_1 = j \mid \Omega_0),\ j = 1,\ldots,m \to\ & \Pr(s_1 = j \mid \Omega_1),\ j = 1,\ldots,m \\
\to\ & \Pr(s_2 = j \mid \Omega_1),\ j = 1,\ldots,m \\
\to\ & \Pr(s_2 = j \mid \Omega_2),\ j = 1,\ldots,m \\
& \vdots \\
\to\ & \Pr(s_T = j \mid \Omega_T),\ j = 1,\ldots,m
\end{aligned}
$$

と逐次的にフィルター化確率 $\Pr(s_t = j \mid \Omega_t), t = 1,\ldots,T,$ $j = 1,\ldots, m$ が計算できる。

ここで、副次的に計算される $\Pr(s_t = j \mid \Omega_{t-1})$ は**(1期先)予測確率(one step ahead prediction probability)**と呼ばれるもので、これはこれ自体で非常に有用である。

レジーム・スウィチングモデル

平滑化確率について（詳しい説明は省略）、

$$\Pr(s_{t-1}=k\,|\,\Omega_T)=\sum_{j=1}^{m}\frac{\Pr(s_t=j\,|\,\Omega_T)\Pr(s_{t-1}=k\,|\,\Omega_{t-1})p_{j,k}}{\Pr(s_t=j\,|\,\Omega_{t-1})}$$

と計算することができるので、これを用いて、$\Pr(s_T=j\,|\,\Omega_T)$ から逐次的に $\Pr(s_{T-1}=j\,|\,\Omega_T)$, $\Pr(s_{T-2}=j, \Omega_{T-1})$, …., $\Pr(s_1=j\,|\,\Omega_T)$, $j=1,\ldots,m$ を計算する事ができる。

レジーム・スウィチングモデル

- MSモデルの推定

MSモデルの推定には通常最尤法を用いる。対数尤度関数は、フィルター化確率のところで出てきた $f(y_t \mid \Omega_{t-1})$ を用いて

$$\log L(\boldsymbol{\theta}) = \sum_{t=1}^{T} \log f(y_t \mid \Omega_{t-1})$$

と計算できるので、これを未知パラメーターベクター $\boldsymbol{\theta}$ について最大化して $\boldsymbol{\theta}$ の最尤推定値を求める。

レジーム･スウィチングモデル

- 時変推移確率モデル(Time Varying Transition Probability Model)

ここまでは推移確率行列は時間を通じて一定であるとしてきた。しかしながら、推移確率も説明変数に依存して変化するとすることもできる。そのようなモデルを**時変推移確率モデル**という。

事変推移確率モデルでは確率をモデル化する必要がある。よく用いられるのはロジット分析で用いられる関数である。

レジーム・スウィチングモデル

例えば、推移確率 $\Pr(s_t = j \mid s_{t-1} = i)$ は

$$p_{ij} = \Lambda(\beta_{0,i,j} + \beta_{1,i,j} X_{1,t-1} + \beta_{2,i,j} X_{2,t-1} + \ldots + \beta_{K,i,j} X_{K,t-1})$$

とモデル化できる。ここで $\Lambda(x) = \dfrac{\exp(x)}{1+\exp(x)}$ である。

$\Lambda(x)$は $(0, 1)$の間の値を取り、単調増加関数であることが確かめられる。確率をモデル化する際によく用いられる。

係数が正であれば、その変数が増加すると推移確率が増加することとなる。