長期記憶パラメータの変化点推定

山口　圭子　(東京理科大学工学部)

1　はじめに

Granger and Joyeux (1980), Hosking (1981) により提案された ARFIMA(autoregressive fractionally integrated moving average) モデルは, ARIMA(自己回帰和分移動平均) モデルの和分オーダーを実数に拡張したもので, 自己相関の減衰速度が遅い (hyperbolic) ことが特徴であり, 長期記憶性 (自己相関が絶対総和不可能) を捉えることが可能なモデルとして水文学、ファイナンス、計量経済学等の分野で応用されてきた。

実証分析において、長期記憶モデルが応用されるのは、モデルの性質から明らかなように、比較的標本期間の長いデータに対してである。ならば、そのような長期の間にデータの発生メカニズム自体が変化している可能性を考慮しなくてよいのだろうか？

このような問題意識のもと, Beran and Terrin (1996), Horvath and Shao (1999), Horvath (2001), Ling (2007) では, ARFIMA モデルにおいて自己相関の強さを表す長期記憶パラメータが一定という帰無仮説に対し, 対立仮説は変化あり (ただし変化点は未知) という検定問題を考え, sup-Wald タイプ (Andrews (1993)) の検定を提案している。そして, 帰無仮説における検定統計量の性質が調べられている。

長期記憶パラメータが一定という帰無仮説が棄却された場合, 次に知りたいことは変化点の場所である。そこで本研究では変化点推定方法を提案し, その分布を導出した。

2　sup Wald 検定

以下のような変化点が k_0 の ARFIMA(p,d,q) モデルを考える:

$$y_t = \begin{cases} y_t^{(1)} = (1-L)^{-d_{10}}\phi^{-1}(L)\psi(L)\varepsilon_t = \sum_{j=0}^{\infty}\kappa_j(-\theta_{10})\varepsilon_t & t = 1,\dots,k_0 \\ y_t^{(2)} = (1-L)^{-d_{20}}\phi^{-1}(L)\psi(L)\varepsilon_t = \sum_{j=0}^{\infty}\kappa_j(-\theta_{20})\varepsilon_t & t = k_0+1,\dots,T \end{cases} \tag{1}$$

ここで $\{\varepsilon_t\} \sim IID(0,1)$, $\phi(L) = 1-\sum_{i=1}^{p}\phi_i L^i$, $\psi(L) = 1-\sum_{i=1}^{q}\psi_i L^i$, ただし L はラグオペレータ, $0 < d < 0.5$, $(1-L)^{-d} = \sum_{j=0}^{\infty}\frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}L^j$ である。Ling (2007) の Assumption 4.1 と同様に, 任意の $|z| \le 1$ に対して, $\phi(z) \neq 0$, $\psi(z) \neq 0$ を仮定する。また, $\phi_p \neq 0$, $\psi_q \neq 0$ で, $\phi(z)$ と $\psi(z)$ は共通の根をもたないと仮定する。$\eta = (\phi_1,\dots,\phi_p,\psi_1,\dots,\psi_q)'$, $\theta_{10} = (d_{10},\eta')'$, $\theta_{20} = (d_{20},\eta')'$, $m = p+q+1$ とする。θ_{10}, θ_{20} は未知パラメータとする。Ling (2007) にあるように, パラメータ空間 Θ は R^m のコンパクト部分空間とする。真の break fraction を $\tau_0 = k_0/T$ とし, $0 < \underline{\tau} \le \tau_0 \le \bar{\tau} < 1$ とする。break fraction は一定で未知とする。

Beran and Terrin (1996), Horvath and Shao (1999) は以下の検定問題を考えた:

$$H_0 : d_{10} = d_{20} \text{ for all } t \quad \text{v.s.}$$

$$H_1(k) : y_t = y_t^{(1)} \text{ for } t \le k \quad \text{and} \quad y_t = y_t^{(2)} \text{ for } k < t \quad \text{with } d_{10} \neq d_{20} \text{ for some } k \in [\underline{\tau}T, \bar{\tau}T],$$

ここで $[\cdot]$ は整数部分を表す。 Whittle 推定量をもとにした sup-Wald タイプ検定を提案し, 帰無仮説のもとでの漸近分布を導出した。Horvath (2001), Ling (2007) はすべてのパラメータが変化した場合を考えている。

本研究では長期記憶パラメータの変化に注目し, Ling (2007) で用いられている CSS 推定量をもとにした sup-Wald 検定について考える。

各 $k \in [\underline{\tau}T, \bar{\tau}T]$ に対して, θ_{10} と θ_{20} の CSS 推定量は

$$\hat{\theta}_1(k/T) = \arg\min_{\Theta}\frac{1}{T}\sum_{t=1}^{k}\left[(1-L)^{d_1}\phi(L)\psi^{-1}(L)(y_t I\{t>0\})\right]^2 = \arg\min_{\Theta}\frac{1}{T}\sum_{t=1}^{k}e_t^2(\theta_1),$$

$$\hat{\theta}_2(k/T) = \arg\min_{\Theta} \frac{1}{T} \sum_{t=k+1} \left[(1-L)^{d_2}\phi(L)\psi^{-1}(L)(y_t I\{t>0\})\right]^2 = \arg\min_{\Theta} \frac{1}{T} \sum_{t=k+1} e_t^2(\theta_2),$$

で求める。ここで $I\{\cdot\}$ は定義関数である。

検定統計量を説明するまえに, 記号を定義する。$\tilde{D}_t(\theta) = \frac{\partial}{\partial\theta} e_t^2(\theta)$, $\tilde{P}_t(\theta) = -\frac{\partial^2}{\partial\theta\partial\theta'} e_t^2(\theta)$, $\Sigma_{1\tau}(\theta) = \sum_{t=1}^{k} \tilde{P}_t(\theta)$, $\Sigma_{2\tau}(\theta) = \sum_{t=k+1}^{T} \tilde{P}_t(\theta)$,

$$\hat{\Sigma}_T(k) = \Sigma_{1\tau}(\hat{\theta}_1(k/T)) + \Sigma_{2\tau}(\hat{\theta}_2(k/T)),$$

$$\hat{\Omega}_T(k) = \sum_{t=1}^{k} \tilde{D}_t(\hat{\theta}_1(k/T))\tilde{D}_t'(\hat{\theta}_1(k/T)) + \sum_{t=k+1}^{T} \tilde{D}_t(\hat{\theta}_2(k/T))\tilde{D}_t'(\hat{\theta}_2(k/T))$$

とする。また. $z_1(\theta) = (1-L)^{d_1}\phi(L)\psi^{-1}(L)y_1$ とする。これは $e_1(\theta)$ と似ているが, 定義関数が入っていないところがちがう。$D_1(\theta) = \frac{\partial}{\partial\theta} z_1^2(\theta)$, $P_1(\theta) = -\frac{\partial^2}{\partial\theta\partial\theta'} z_1^2(\theta)$, $\Sigma = E[P_1(\theta_{10})]$, $\Omega = E[D_1(\theta_{10})D_1'(\theta_{10})]$ とする。Ling (2007) ではより一般的な場合も考えているので, Ω と Σ が区別されているが, ここでは, CSS 推定は擬似最尤推定であるので, Robinson (2006) からもわかるように, $\Omega = \Sigma$ である。

各 k に対し, Wald 検定統計量は以下で与えられる。

$$W_T(k/T) = \frac{k(T-k)}{T} \hat{\rho}^{-1} [\hat{d}_1(k/T) - \hat{d}_2(k/T)]^2 \tag{2}$$

ここで, $\hat{\rho}$ は $(\hat{\Sigma}_T(k)/T)^{-1}$ の 1 行 1 列目の要素である。また, ρ を Σ^{-1} の 1 行 1 列目の要素とする。

いま, 変化点は未知であるとしているので, Andrews (1993) にあるように, (2) に対して $\tau \in [\underline{\tau}, \bar{\tau}]$ の範囲で上限をとったものを検定統計量とし. $\sup W_T(\tau)$ と表すことにする。$\sup W_T(\tau)$ の分位点は, 例えば Andrews (1993) の Table 1 に載っている。

3 break fraction の推定量

変化点推定量を

$$\hat{k} = \{k : \max_k W_T(k/T)\},$$

とする。Bai (1997) でコメントされているように, 帰無仮説が棄却された場合, そのまま自動的に変化点推定量が得られるという利点がある。

ここでは break fraction 推定量 $\hat{\tau} = k/T$ の分布について考えていく。Chong (2001) でコメントされているように, 変化の幅が固定されている場合, 極限分布が誤差項の分布に依存してしまう。そこで, 以下の Assumption にあるように, 変化の幅が適当なオーダーで 0 にいくという仮定がよくみられる (Bai (1997), Chong (2001))。本研究でもこの仮定を採用する。

Assumption 3.1. $d_{10} - d_{20} = \delta\nu_T$, where $\nu_T > 0$, $\nu_T \to 0$ and $T^{1/2-\alpha}\nu_T \to \infty$ for some $\alpha \in (0, 1/2)$ and $\delta \neq 0$.

break fraction の識別を保つため, θ の推定量が真の値に収束するオーダーよりも遅くなるように仮定されている。この仮定の下, 次の定理が成立する。

Theorem 3.1. *Under Assumption 3.1, we have: (i) $\hat{\tau} \to_p \tau_0$; and (ii) for every $\eta > 0$ there exists a $C < \infty$ such that for all large T, $P(|T\nu_T^2(\hat{\tau} - \tau_0)| > C) < \eta$.*

さらに, 分布は以下のとおりである。

Theorem 3.2. *Under Assumption 3.1, we have*

$$\rho^{-1}T(d_{20} - d_{10})^2(\tau - \tau_0) \to^d Arg\max_r \left(B^*(r) - \frac{1}{2}|r|\right), \tag{3}$$

where $B^*(r)$ is a two-sided Brownian motion on R.

Bai (1997) の Appendix B には, (3) 式の左辺の分布関数が求められているので, ここから break fraction の信頼区間を計算することも可能である。