

時系列分析の概要

2024年6月(作成)
倉谷 隆博

1. 時系列分析の統計量

時系列データとは、時間の経過とともに観察されるデータの数列 $\{y(t)\} = \{y(1), y(2), \dots, y(T)\}$ 但し、 $t = 1, 2, \dots, T$ (図1-1) のことであり、時間の推移とともに変化する確率変数の集合である。時系列データとしては、例えば、気温などの天候データ、為替レートなどの経済指標、客数などの営業データなど多岐にわたる。時系列分析の目的は、過去の時系列データを使用して今後の動向を予測したり、時系列データ間の相互関係を把握したりすることである。

時系列データ $\{y(t)\}$ の基本統計量を整理しておく。時系列分析では、時系列データを確率変数の実現値であると見なす。なお、時点 $t = 1 \sim T$ である。

(1) 期待値: $\mu(t) = E[y(t)] = (\sum_{t=1}^T y(t)) / T \quad \dots \text{(式1-1)}$

(2) 分散: $\sigma^2(t) = V[y(t)] = E[(y(t) - \mu(t))^2] = (\sum_{t=1}^T (y(t) - \mu(t))^2) / T \quad \dots \text{(式1-2)}$

(3) 標準偏差: $\sigma(t)$ 分散の平方根 (ファイナンス分野では、標準偏差のことを**ボラティリティ**(volatility)という)

(4) k 次の**自己共分散**(auto covariance) (図1-2)

現時点と k 時点前の時系列データの分散共分散を k 次の自己共分散 $\gamma_k(t)$ という (同一の確率変数を対象とするため、「自己」という名称が付いている)。

時点 t に対して k 時点前の時点は $t - k$ で表す。 k は時点間の時間差を表す。時点 $t - k$ は時点 t に対して k 時点の**遅れ**(k 次**ラグ**, lag) があるという。

従って、 k 次のラグが同じである自己共分散 γ_k と γ_{-k} に関して、 $\gamma_k(t) = \gamma_{-k}(t)$ が成り立つ。 k 次の自己共分散は(式1-3)で表される。

$$\gamma_k(t) = \text{Cov}[y(t), y(t - k)] = E[(y(t) - \mu(t))(y(t - k) - \mu(t - k))] \quad \dots \text{(式1-3)}$$

$k = 0$ のとき、 $\gamma_0(t) = \text{Cov}[y(t), y(t - 0)] = \text{Cov}[y(t), y(t)] = E[(y(t) - \mu(t))^2] = V[y(t)] = \sigma^2(t) \quad \dots \text{(式1-4)}$

(式1-4)は、 0 次の自己共分散 $\gamma_0(t)$ は分散 $\sigma^2(t)$ に等しい、とすることを表している。

(5) k 次の**自己相関係数**(auto correlation)

k 次の自己相関を分散の幾何平均 $(V[y(t)] V[y(t - k)])^{1/2}$ を使用して規格化したものを k 次の自己相関係数 $\rho_k(t)$ という。

これによって、単位の異なる時系列データの自己共分散を互いに比較できるようになる。

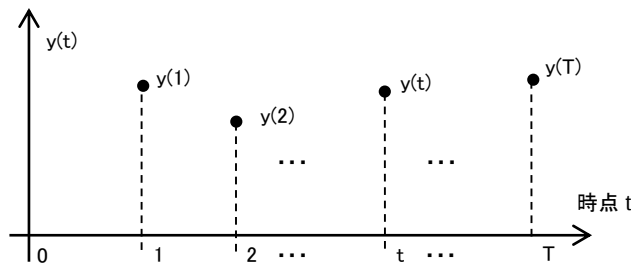
$$\rho_k(t) = \text{Cov}[y(t), y(t - k)] / (V[y(t)] V[y(t - k)])^{1/2} = \gamma_k(t) / (\gamma_0(t) \gamma_0(t - k))^{1/2} \quad \dots \text{(式1-5)}$$

$k = 0$ のとき、 $\rho_0(t) = \gamma_0(t) / (\gamma_0(t) \gamma_0(t - 0))^{1/2} = \gamma_0(t) / (\gamma_0(t) \gamma_0(t))^{1/2} = \gamma_0(t) / \gamma_0(t) = 1 \quad \dots \text{(式1-6)}$

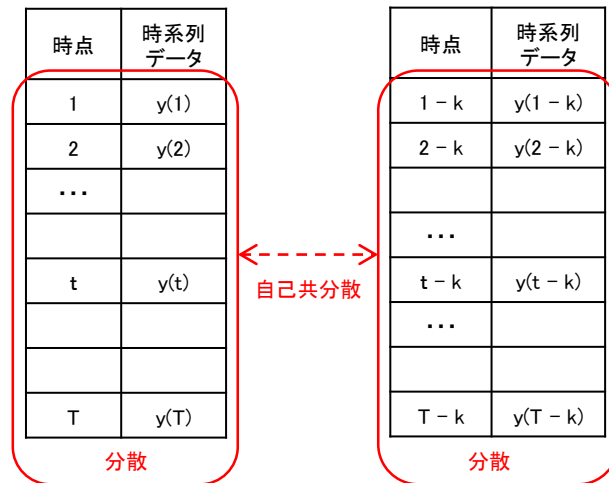
(式1-5)で示すように**自己相関係数 $\rho_k(t)$ は時間差 k の関数になっている。過去のデータが現在のデータに及ぼしている影響の強さを知ることができる。**

(式1-5)を可視化したグラフは、**コレログラム**(correlogram)と呼ばれる。

自己相関係数とコレログラムの計算例を(付録1)に載せる。



(図1-1) 時系列データ



(図1-2) 分散と自己共分散

2. 定常性 (Stationarity)

時系列分析においては、時系列データの定常性という概念が重要である。2種類の定常性が定義されるが、時系列分析では主に、「弱定常性」を前提とした分析が行われる。

(1) 弱定常性 (weakly stationarity) の定義

① $E[y(t)] = \mu \quad \dots$ (式2-1)

この式は、 $y(t)$ の平均 $\mu(t)$ が、時点 t に依存せず一定であることを示している。
つまり、 $\mu(t)$ が時点 t に依存しないならば、 $\mu(t) = \mu$ とおくことができ、(式2-1)になる。

② $Cov[y(t), y(t-k)] = \gamma_k \quad \dots$ (式2-2)

この式は、 $y(t)$ と時間差 k の $y(t-k)$ の自己共分散が、時点 t に依存せず時間差 k のみに依存することを示している (自己相関に関しても同様である)。
つまり、 $\gamma_k(t)$ が時点 t に依存しないならば、 $\gamma_k(t) = \gamma_k$ とおくことができ、(式2-2)になる。

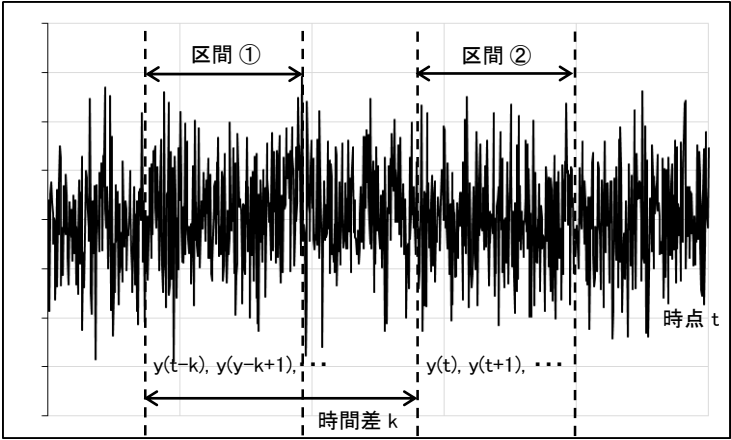
なお、(式1-4)より、分散 $\sigma^2(t)$ は0次の自己共分散 $\gamma_0(t)$ に等しいが、
定常過程であれば、 $\gamma_0(t)$ が時刻 t に依存せず $\gamma_0(t) = \gamma_0$ と表すことができるので、分散も時刻 t に依存せず $\sigma^2(t) = \sigma^2$ (一定)と表すことができる。

また、定常過程であれば $\gamma_k(t)$ も時刻 t に依存せず $\gamma_k(t) = \gamma_k$ 、 $\gamma_0(t) = \gamma_0(t-k) = \gamma_0$ と表すことができるので、自己相関係数 ρ_k は(式2-3)のようになる。
自己相関係数は、時刻 t に依存せず、時間差 k に依存することが分かる。

$\rho_k = \gamma_k(t) / (\gamma_0(t) \gamma_0(t-k))^{1/2} = \gamma_k / (\gamma_0 \gamma_0)^{1/2} = \gamma_k / \gamma_0 \quad \dots$ (式2-3)

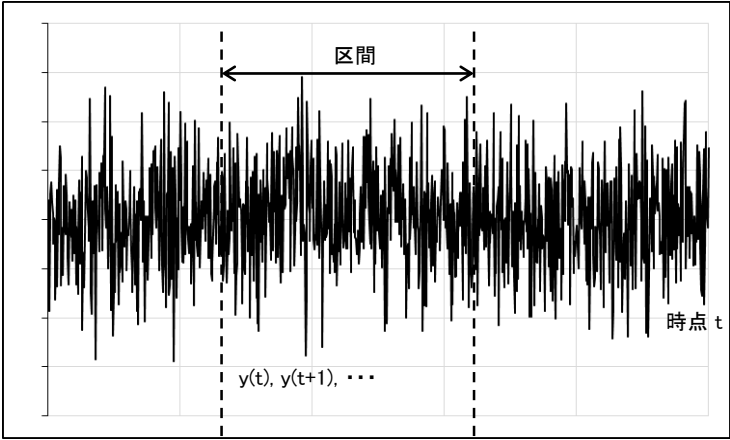
(2) 強定常性 (strongly stationarity 独立同分布) の定義

任意の時点の同時確率分布 $f(y(t), y(t+1), \dots, y(t+i))$ が等しい。つまり、同時確率分布が時間変化しないとき、強定常性であるという。
例えば、i.i.d. (independent and identically distributed : 独立同一分布) に従うデータ系列は、各時点のデータは互いに独立であり、異なる時点のデータ間の相関はないので、強定常性である。



弱定常性のデータでは、任意の区間(例えば、区間①と区間②)の平均と分散が一定であり、(同じ区間幅を対象とする)自己共分散と自己相関係数は時間差 k のみに依存する。

(図2-1) 弱定常性



任意の時点の任意の区間の同時確率分布が等しい。

(図2-2) 強定常性

(3) エルゴード性 (Ergodicity)

空間平均と時間平均が等しいことをいう。例えば、1,000 人が同時にサイコロを振ってでた目の平均(空間平均)と、1 人が 1,000 サイコロを振ってでた目の平均(時間平均)が等しいことをいう。時系列分析は通常、エルゴード性を前提にしている。

3. ホワイトノイズ(White Noise)

ホワイトノイズは、フーリエ変換(別途資料「周波数解析(フーリエ変換)の基礎知識」を参照)を使ってパワースペクトルを求めると、広い周波数域でほぼ一定の値をとることが分かる。そこで、色の領域で全周波数域をカバーする色は白色(ホワイト)であることから、このことに因んでホワイトノイズと言う名前が付けられている。

ホワイトノイズの定義

定常過程の時系列データ $\varepsilon(t)$ に関して、全ての時点において、下記が成立するときホワイトノイズと呼ばれる。

$$\text{期待値 } E[\varepsilon(t)] = 0 \quad \dots \text{ (式3-1)}$$

$$\text{自己共分散 } \gamma_k = \text{Cov}[\varepsilon(t), \varepsilon(t-k)] = E[(\varepsilon(t) - E[\varepsilon(t)])(\varepsilon(t-k) - E[\varepsilon(t-k)])] = E[(\varepsilon(t) - 0)(\varepsilon(t-k) - 0)]$$

$$= E[\varepsilon(t) \varepsilon(t-k)] = \begin{cases} \sigma^2 & k=0 \text{ のとき(分散)} \\ 0 & k \neq 0 \text{ のとき(共分散)} \end{cases} \quad \dots \text{ (式3-2)}$$

(1) ホワイトノイズの分散 $V[\varepsilon(t)]$

定義より、 $V[\varepsilon(t)] = E[(\varepsilon(t) - E[\varepsilon(t)])^2]$ であるが、 $k=0$ のとき(分散)、(式3-1)を代入して(式3-2)より、 $V[\varepsilon(t)] = E[(\varepsilon(t) - 0)^2] = E[\varepsilon(t)^2] = \sigma^2 \quad \dots \text{ (式3-3)}$
つまり、**ホワイトノイズは、どの時点においても、期待値は0であり(式3-1)、分散は一定 σ^2 (式3-3)である。**

(2) ホワイトノイズの自己相関 ρ_k

$$\text{(式3-2)において、} k=0 \text{ のとき、} \gamma_0 = E[\varepsilon(t) \varepsilon(t-0)] = E[\varepsilon(t)^2] = V[\varepsilon(t)] = \sigma^2 \quad \dots \text{ (式3-4)}$$

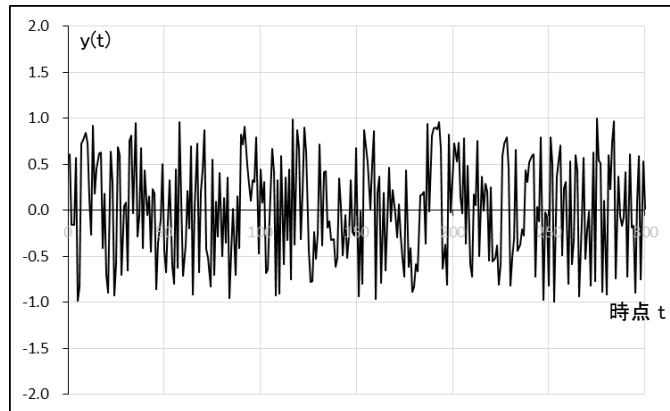
$$\text{(式2-3)より自己相関は、} \rho_k = \gamma_k / \gamma_0 \text{ であるが、} k \neq 0 \text{ のとき、(式3-2)より } \gamma_k = 0 \text{ であるので、} \rho_k = \gamma_k / \gamma_0 = 0 / \sigma^2 = 0 \quad \dots \text{ (式3-5)}$$

つまり、**ホワイトノイズは、どの時点においても、自己相関を持たない**ことが分かる。

(図3-1)にホワイトノイズに模した時系列データのグラフの一例を示す。 $\varepsilon(t)$ がホワイトノイズに従うとき、 $\varepsilon(t) \sim \text{W.N.}(\sigma^2)$ と表記する。

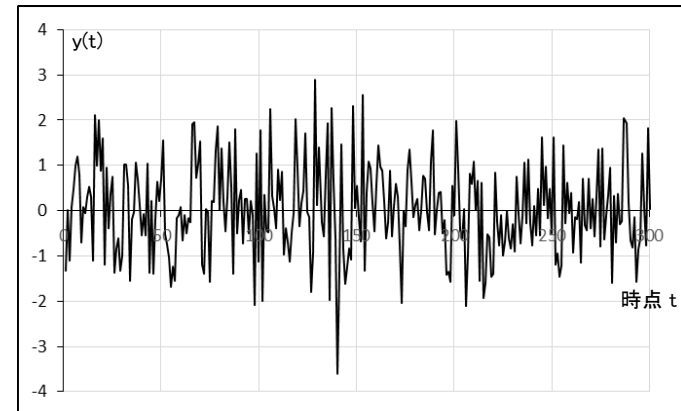
また、(図3-2)に、正規ホワイトノイズに模した時系列データのグラフの一例を示す。

正規分布 $N(0, \sigma^2)$ に従うホワイトノイズ $\varepsilon(t)$ は**正規ホワイトノイズ**と呼ばれ、 $\varepsilon(t) \sim \text{i.i.d. } N(\sigma^2)$ と表記する(まお。正規ホワイトノイズは、i.i.d. 系列でもある)。



Excel の乱数関数 RAND() (値域は、0 ~ 1) を使い、
 $2 * \text{RAND}() - 1.0$ (値域は、-1.0 ~ 1.0) を
平均 = 0 のホワイトノイズに模したデータとした。なお、このグラフの
データ数は 300 (横軸は時点)、平均は 0.014、分散は 0.315 である。

(図3-1) ホワイトノイズ



乱数(0 ~ 1)を確率に見立てて、
それに対応した正規分布(平均=0、分散=1)に従う確率変数の値を、
Excel 関数 NORMSINV(RAND(),0,1) で求め、
正規ホワイトノイズ(平均 = 0、分散 = 1)を模したデータとした。
なお、データ数は 300 (横軸は時点)、平均は 0.017、分散は 1.010 である。

(図3-2) 正規ホワイトノイズ

4. 自己相関の検定

時系列データが自己相関を持っていれば、それを利用して時系列モデルを構成することができる。つまり、時系列分析では、時系列データが自己相関を持っているかの確認は重要である。確認方法としては、時間差ごとの個々の自己相関を検定する方法と、ある時間差までの自己相関をまとめて検定する方法がある。

(1) 統計量の推定 (定常性が前提)

T = データ数、k = 時間差 として、
期待値の推定量 (標本平均): $\bar{y} = (1 / T) \sum_{t=1}^T y(t) \quad \cdots$ (式4-1)

自己共分散の推定量 (標本自己共分散): $\hat{\gamma}_k = (1 / T) \sum_{t=k+1}^T (y(t) - \bar{y})(y(t-k) - \bar{y}) \quad \cdots$ (式4-2)

(注) T-k で割るのが自然であるが、T >> k のことが多いので T で割る

自己相関の推定量 (標本自己相関係数): 定常であれば、時点 t に依らず、時間差 k のみに依存するので、(式2-3)と同様にして、
 $\hat{\rho}_k = \hat{\gamma}_k(t) / (\hat{\gamma}_0(t) \hat{\gamma}_0(t-k))^{1/2} = \hat{\gamma}_k / (\hat{\gamma}_0 \hat{\gamma}_0)^{1/2} = \hat{\gamma}_k / \hat{\gamma}_0 \quad \cdots$ (式4-3)

(2) 時間差ごとの個々の自己相関の検定

帰無仮説: $\rho_k = 0$ 対立仮説: $\rho_k \neq 0$

自己相関係数の推定量 $\hat{\rho}_k$ の漸近分布は、y(t) が i.i.d. であれば、正規分布 N(0, 1 / T) (但し、T は、時系列データ数) であることが知られている。

次いで、この分布を標準正規分布に標準化すれば、 $(\hat{\rho}_k - 0) / (1 / T)^{1/2} = \hat{\rho}_k / (1 / T)^{1/2} \sim$ 標準正規分布 (0, 1)

ここで、標準正規分布を使用した有意水準を両側 5% に設定すれば、信頼度 95% の信頼区間は標準正規分布の下側 2.5% 点 = -1.96, 上側 2.5% 点 = 1.96 であるので

$-1.96 \leq \hat{\rho}_k / (1 / T)^{1/2} \leq 1.96 \rightarrow -1.96 / T^{1/2} \leq \hat{\rho}_k \leq 1.96 / T^{1/2}$
従って、帰無仮説 $\rho_k = 0$ が棄却されるのは、 $|\hat{\rho}_k| > |1.96 / T^{1/2}| \quad \cdots$ (式4-4) が成立するときであることが分かる。

(3) ある時間差までの自己相関をまとめた検定

帰無仮説: $\rho_1 = \rho_2 = \cdots = \rho_m = 0$ 対立仮説: 少なくとも1つの $\rho_k \neq 0$ (k = 1, 2, ..., m)

この検定では、帰無仮説が明確に定められている一方、対立仮説は比較的あいまいに定められており、**かばん検定 (portmanteau test)** と呼ばれている。

時系列データの代表的な、かばん検定が **Ljung - Box (リュング・ボックス) 検定** で、(式4-5) が検定統計量である。

$Q(m) = T(T+2) \sum_{k=1}^m \rho_k^2 / (T-k) \quad \cdots$ (式4-5) は漸近的に自由度 m のカイニ乗分布 $\chi^2(m)$ (付録P) に従うことが知られている。

Ljung - Box 検定の検定統計量 Q(m) は、**Box - Pierce (ボックス・ピアース) 検定** の検定統計量 $R(m) = T \sum_{k=1}^m \rho_k^2$ (付録8) の改良版であり、Box - Pierce 検定の検定統計量を $(T+2) / (T-k)$ (k = 1, 2, ..., m) で重み付した検定統計量になっている。

Ljung - Box 検定の検定統計量 Q(m) が $\chi^2(m)$ 分布の上側 5% 棄却域に入るとき (有意水準 5%)、言い換えれば、この検定統計量 Q(m) の p 値が 0.05 より小さければ帰無仮説は棄却され、対立仮説が採択される。

なお、帰無仮説を棄却できないときは、自己相関係数が全て 0 であることを棄却 (否定) できないときである。これは、時系列がホワイトノイズであることを否定できない、言い換えれば、時系列がホワイトノイズであるという判断になる (ホワイトノイズであれば自己相関係数は、全て 0 になる)。

このように、**Box - Pierce 検定は時系列データがホワイトノイズであるか否かの検定に使用できる。**

(付録9)に自己相関の検定の計算例を、(付録10)に**ホワイトノイズの検定**の計算例を載せる。

5. MA過程(移動平均過程: Moving Average process)

現在の時系列データ $y(t)$ を現在と過去 (期間 q) のホワイトノイズ $\varepsilon(t), \varepsilon(t-1), \dots, \varepsilon(t-q)$ を使用して表す数式モデルが MA 過程である。
ホワイトノイズを介して時系列データ $y(t)$ と時間差 k の時系列データ $y(t-k)$ の間の相関関係を活用する数式モデルである。時系列データの期間が q である MA 過程は q 次 MA 過程と呼ばれる。この過程は $MA(q)$ 過程と表記され、(式5-1)で表される。以下の式変形では、(付録B)を参照のこと。

$y(t) = \mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q) \quad \dots (式5-1)$
但し、 $\varepsilon(t) \sim W.N.(\sigma^2)$: ホワイトノイズ、期待値: $E[\varepsilon(t)] = 0$ 、自己共分散: $k \neq 0$ のときは $\gamma_k(t) = \text{Cov}[\varepsilon(t), \varepsilon(t-k)] = 0$ 、 $k = 0$ のときは $\gamma_0(t) = \text{Cov}[\varepsilon^2(t)] = \sigma^2$

(1) $MA(q)$ 過程の期待値
 $E[y(t)] = E[\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)] = E[\mu] + E[\varepsilon(t)] + \theta_1 E[\varepsilon(t-1)] + \theta_2 E[\varepsilon(t-2)] + \dots + \theta_q E[\varepsilon(t-q)]$
 $= \mu + 0 + \theta_1 \cdot 0 + \theta_2 \cdot 0 + \dots + \theta_q \cdot 0 = \mu \quad \dots (式5-2)$

(2) $MA(q)$ 過程の分散
 $V[y(t)] = V[\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)]$
 $= E[(\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)) - E[\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)]]^2$
 $= E[(\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)) - E[\mu] - E[\varepsilon(t)] - \theta_1 E[\varepsilon(t-1)] - \theta_2 E[\varepsilon(t-2)] - \dots - \theta_q E[\varepsilon(t-q)]]^2$
 $= E[(\mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)) - \mu - 0 - 0 - 0 \dots - 0]^2 = E[(\varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q))]^2$
 $= (E[\varepsilon^2(t)] + \theta_1 E[\varepsilon(t) \varepsilon(t-1)] + \theta_2 E[\varepsilon(t) \varepsilon(t-2)] + \dots + \theta_q E[\varepsilon(t) \varepsilon(t-q)])$
 $+ (\theta_1 E[\varepsilon(t-1) \varepsilon(t)] + \theta_1^2 E[\varepsilon^2(t-1)] + \theta_1 \theta_2 E[\varepsilon(t-1) \varepsilon(t-2)] + \dots + \theta_1 \theta_q E[\varepsilon(t-1) \varepsilon(t-q)] + \dots$
 $+ (\theta_q E[\varepsilon(t-q) \varepsilon(t)] + \theta_q \theta_1 E[\varepsilon(t-q) \varepsilon(t-1)] + \theta_q \theta_2 E[\varepsilon(t-q) \varepsilon(t-2)] + \dots + \theta_q^2 E[\varepsilon^2(t-q)])$
 $= (\sigma^2 + 0 + \dots + 0) + (0 + \theta_1^2 \sigma^2 + 0 + \dots + 0) + \dots + (0 + \dots + 0 + \theta_q^2 \sigma^2) = (1 + \theta_1^2 + \theta_2^2 \sigma^2 + \dots + \theta_q^2) \sigma^2 \quad \dots (式5-3)$

(3) $MA(q)$ 過程の q 次の自己共分散
① $k \leq q$ のとき (図5-1)
 $y(t) = \mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \dots + \theta_k \varepsilon(t-k) + \theta_{k+1} \varepsilon(t-k-1) + \dots + \theta_q \varepsilon(t-q)$
 $y(t-k) = \mu + \varepsilon(t-k) + \theta_1 \varepsilon(t-k-1) + \dots + \theta_{q-k} \varepsilon(t-q) + \dots + \theta_q \varepsilon(t-k-q)$
 $\text{Cov}[\varepsilon(t-k), \varepsilon(t-k)] = \sigma^2$ の項の係数 = θ_k
 $\text{Cov}[\varepsilon(t-k-1), \varepsilon(t-k-1)] = \sigma^2$ の項の係数 = $\theta_{k+1} \theta_1$
 $\dots$
 $\text{Cov}[\varepsilon(t-q), \varepsilon(t-q)] = \sigma^2$ の項の係数 = $\theta_q \theta_{q-k}$
他の Cov の項は全て 0 である。よって、自己共分散 $\gamma_k = \text{Cov}[y(t), y(t-k)] = (\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}) \sigma^2 \quad \dots (式5-4)$

② $k \geq q+1$ のとき (図5-2)
 $y(t) = \mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q)$
 $y(t-k) = \mu + \theta_k \varepsilon(t-k) + \theta_{k+1} \varepsilon(t-k-1) + \dots + \theta_{k+q} \varepsilon(t-k-q)$
全ての Cov の項は 0 である。よって、自己共分散 $\gamma_k = \text{Cov}[y(t), y(t-k)] = 0 \quad \dots (式5-5)$

(4) $MA(q)$ 過程の q 次の自己相関係数
① $k \leq q$ のとき、(式5-3)と(式5-4)より、
 $\rho_k = \gamma_k / \gamma_0 = \text{Cov}[y(t), y(t-k)] / V[y(t)] = (\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}) \sigma^2 / ((1 + \theta_1^2 + \theta_2^2 \sigma^2 + \dots + \theta_q^2) \sigma^2)$
 $= (\theta_k + \theta_{k+1} \theta_1 + \dots + \theta_q \theta_{q-k}) / (1 + \theta_1^2 + \theta_2^2 \sigma^2 + \dots + \theta_q^2) \quad \dots (式5-6)$
② $k \geq q+1$ のとき、(式5-3)と(式5-5)より、
 $\rho_k = \gamma_k / \gamma_0 = 0 / ((1 + \theta_1^2 + \theta_2^2 \sigma^2 + \dots + \theta_q^2) \sigma^2) = 0 \quad \dots (式5-7)$

従って、MA 過程のコレログラムは、
時間差が次数を超えところでは 0 になる。

$y(t)$	t	t-1	...	t-k	t-k-1	...	t-q			
				$y(t-k)$	t-k	t-k-1	...	t-q	...	t-k-q

(5) 定常性
(1)と(2)より、期待値と分散は時点 t に依らず一定値である。
(3)より、自己共分散は時間差 k のみに依存する。
このことから、MA 過程は常に定常であることが分かる。

$y(t)$	t	t-1	...	t-q	...				
				$y(t-k)$	t-k	t-k-1	...	t-k-q	

MA(2) 過程とそのコレログラムの例を(付録2)に載せる。
(図5-1)時系列データの関係 ($k \leq q$ のとき)
(図5-2)時系列データの関係 ($k \geq q+1$ のとき)

6. AR過程(自己回帰過程: Auto Regressive process)

現在の時系列データ $y(t)$ を過去 (期間 p) の時系列データ $y(t-1), y(t-2), \dots, y(t-p)$ と現在のホワイトノイズ $\varepsilon(t)$ を使用して表す数式モデルが AR 過程である。時系列データ $y(t)$ と時間差 k の時系列データ $y(t-k)$ の間の相関関係を利用する数式モデルである。時系列データの期間が p である AR 過程は p 次 AR 過程と呼ばれ、AR(p) 過程と表記され、(式6-1)で表される。今回の資料では、主に AR 過程を取り上げる。なお、AR(p) 過程は定常過程であるとする。

$$y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots (式6-1)$$

但し、 $\varepsilon(t) \sim W.N.(\sigma^2)$: ホワイトノイズ、期待値: $E[\varepsilon(t)] = 0$, 自己共分散: $k \neq 0$ のときは $\gamma_k(t) = \text{Cov}[\varepsilon(t), \varepsilon(t-k)] = 0$, $k = 0$ のときは $\gamma_0(t) = \text{Cov}[\varepsilon^2(t)] = \sigma^2$

(1) AR(p) 過程の期待値 $E[y(t)]$

$E[y(t)] = E[y(t-1)] = E[y(t-2)] = \dots = E[y(t-p)] = \mu$, $E[\varepsilon(t)] = 0$ であるので、(付録B)の期待値の公式を使って、

$$\begin{aligned} E[y(t)] &= E[c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t)] \\ &= E[c] + \phi_1 E[y(t-1)] + \phi_2 E[y(t-2)] + \dots + \phi_p E[y(t-p)] + E[\varepsilon(t)] \\ &= c + \phi_1 \cdot \mu + \phi_2 \cdot \mu + \dots + \phi_p \cdot \mu + 0 = c + \mu (\phi_1 + \phi_2 + \dots + \phi_p) \end{aligned}$$

従って、期待値 $E[y(t)] = \mu = c / (1 - \phi_1 - \phi_2 - \dots - \phi_p)$ $\dots$ (式6-2)

(2) AR(p) 過程の分散 $V[y(t)]$

$$\gamma_0 = V[y(t)] = \text{Cov}[y(t), y(t)] = \text{Cov}[c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t), y(t)]$$

ここで、(付録B)の共分散の公式を使って、

$$\gamma_0 = \text{Cov}[c, y(t)] + \text{Cov}[\phi_1 y(t-1), y(t)] + \text{Cov}[\phi_2 y(t-2), y(t)] + \dots + \text{Cov}[\phi_p y(t-p), y(t)] + \text{Cov}[\varepsilon(t), y(t)]$$

また、 $\varepsilon(t)$ がホワイトノイズであり、共分散は 0 つまり、 $\text{Cov}[\varepsilon(t), c] = 0$, $\text{Cov}[\varepsilon(t), \varepsilon(t)] = 0$, $\text{Cov}[\varepsilon(t), y(t-k)] = 0$ であることから、

$$\begin{aligned} \text{Cov}[\varepsilon(t), y(t)] &= \text{Cov}[\varepsilon(t), c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t)] \\ &= \text{Cov}[\varepsilon(t), c] + \text{Cov}[\varepsilon(t), \phi_1 y(t-1)] + \text{Cov}[\varepsilon(t), \phi_2 y(t-2)] + \dots + \text{Cov}[\varepsilon(t), \phi_p y(t-p)] + \text{Cov}[\varepsilon(t), \varepsilon(t)] \\ &= 0 + \phi_1 \text{Cov}[\varepsilon(t), y(t-1)] + \phi_2 \text{Cov}[\varepsilon(t), y(t-2)] + \dots + \phi_p \text{Cov}[\varepsilon(t), y(t-p)] + \sigma^2 = 0 + 0 + \dots + 0 + \sigma^2 = \sigma^2 \end{aligned}$$

また、(付録B)の共分散の公式より、 $\text{Cov}[c, y(t)] = 0$ であるので、(式2-2)と(式2-3)を使って、

$$\begin{aligned} \gamma_0 &= 0 + \phi_1 \text{Cov}[y(t-1), y(t)] + \phi_2 \text{Cov}[y(t-2), y(t)] + \dots + \phi_p \text{Cov}[y(t-p), y(t)] + \sigma^2 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma^2 \\ &= \phi_1 \rho_1 \gamma_0 + \phi_2 \rho_2 \gamma_0 + \dots + \phi_p \rho_p \gamma_0 + \sigma^2 = \gamma_0 (\phi_1 \rho_1 + \phi_2 \rho_2 + \dots + \phi_p \rho_p) + \sigma^2 \end{aligned}$$

従って、 $\gamma_0 (1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p) = \sigma^2$ であるので、

$$\text{分散 } V[y(t)] = \gamma_0 = \sigma^2 / (1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p) \quad \dots (式6-3)$$

(3) AR(p) 過程の p 次の自己共分散

$$\begin{aligned} \gamma_k &= \text{Cov}[y(t), y(t-k)] = \text{Cov}[c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t), y(t-k)] \\ &= \text{Cov}[c, y(t-k)] + \text{Cov}[\phi_1 y(t-1), y(t-k)] + \text{Cov}[\phi_2 y(t-2), y(t-k)] + \dots + \text{Cov}[\phi_p y(t-p), y(t-k)] + \text{Cov}[\varepsilon(t), y(t-k)] \end{aligned}$$

ここで、 $\text{Cov}[c, y(t-k)] = 0$, $\text{Cov}[\varepsilon(t), y(t-k)] = 0$ であることから、

$$\begin{aligned} \gamma_k &= 0 + \phi_1 \text{Cov}[y(t-1), y(t-k)] + \phi_2 \text{Cov}[y(t-2), y(t-k)] + \dots + \phi_p \text{Cov}[y(t-p), y(t-k)] + 0 \\ &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad \dots (式6-4) \end{aligned}$$

(4) AR(p) 過程の p 次の自己相関係数

(式2-3)に(式6-4)を代入して、再度(式2-3)を使って、

$$\rho_k = \gamma_k / \gamma_0 = (\phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}) / \gamma_0 = \phi_1 (\gamma_{k-1} / \gamma_0) + \phi_2 (\gamma_{k-2} / \gamma_0) + \dots + \phi_p (\gamma_{k-p} / \gamma_0) = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad \dots (式6-5)$$

この式は、 p 次の AR(p) 過程を表す(式6-1)と同じ係数を持つ式であり、**ユール・ウォーカー方程式 (Yule - Walker equation)**と呼ばれている。

(5) 定常性

MA 過程が常に定常過程であるのに対して、AR 過程が定常過程であるためには条件が必要になる(定常条件)。

① 対角行列を使って確認する方法(付録5)

特性方程式 $\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_{p-1} \lambda - \phi_p = 0$ $\dots$ (式6-6) の全ての根 λ に関して、 $|\lambda| < 1$ のとき、AR 過程は定常である。

② z 変換(付録D-1、付録D-2)を使って確認する方法(付録6)

特性方程式 $1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0$ $\dots$ (式6-7) の全ての z^{-1} 根に関して、 $|z^{-1}| < 1$ つまり、 $|z| > 1$ のとき、AR 過程は定常である。

AR(2) 過程とそのコレログラムの例を(付録3)に載せる。

7. ARMA過程(自己回帰移動平均過程: Auto Regressive Moving Average process)

ARMA 過程は、AR 過程と MA 過程を組み合わせた過程であり、(式7-1)で表される。

$$y(t) = c + \underbrace{\phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p)}_{\text{AR 過程}} + \underbrace{\varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \cdots + \theta_q \varepsilon(t-q)}_{\text{MA 過程}} \quad \cdots \text{(式7-1)}$$

但し、 $\varepsilon(t) \sim \text{W.N.}(\sigma^2)$: ホワイトノイズ、期待値: $E[\varepsilon(t)] = 0$ 、自己共分散: $k \neq 0$ のときは $\gamma_k(t) = \text{Cov}[\varepsilon(t), \varepsilon(t-k)] = 0$ 、 $k = 0$ のときは $\gamma_0(t) = \text{Cov}[\varepsilon^2(t)] = \sigma^2$

(1) ARMA(p, q) 過程の期待値

$$\begin{aligned} E[y(t)] &= E[c + \phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p) + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \cdots + \theta_q \varepsilon(t-q)] \\ &= E[c] + \phi_1 E[y(t-1)] + \phi_2 E[y(t-2)] + \cdots + \phi_p E[y(t-p)] + E[\varepsilon(t)] + \theta_1 E[\varepsilon(t-1)] + \theta_2 E[\varepsilon(t-2)] + \cdots + \theta_q E[\varepsilon(t-q)] \end{aligned}$$

ここで、定常過程であれば、 $E[y(t)] = E[y(t-1)] = E[y(t-2)] = \cdots = E[y(t-p)] = \mu$ 、 $E[\varepsilon(t)] = E[\varepsilon(t-1)] = E[\varepsilon(t-2)] = \cdots = E[\varepsilon(t-q)] = 0$ であるので、

$$E[y(t)] = c + \phi_1 \mu + \phi_2 \mu + \cdots + \phi_p \mu + 0 + 0 + 0 + \cdots + 0 = c + (\phi_1 + \phi_2 + \cdots + \phi_p) \mu$$

従って、期待値 $E[y(t)] = \mu = c / (1 - \phi_1 - \phi_2 - \cdots - \phi_p)$ $\cdots$ (式7-2)

(2) ARMA(p, q) 過程の自己共分散

(付録B)の共分散の公式から、独立な変数の和の共分散は、変数ごとの共分散の和に等しいこと、また、 $\text{Cov}[c, y(t-k)] = 0$ であることから、

$$\begin{aligned} \gamma_k &= \text{Cov}[y(t), y(t-k)] = \text{Cov}[c + \phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p) + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \cdots + \theta_q \varepsilon(t-q), y(t-k)] \\ &= \text{Cov}[c, y(t-k)] + \text{Cov}[\phi_1 y(t-1), y(t-k)] + \text{Cov}[\phi_2 y(t-2), y(t-k)] + \cdots + \text{Cov}[\phi_p y(t-p), y(t-k)] \\ &\quad + \text{Cov}[\varepsilon(t), y(t-k)] + \text{Cov}[\theta_1 \varepsilon(t-1), y(t-k)] + \text{Cov}[\theta_2 \varepsilon(t-2), y(t-k)] + \cdots + \text{Cov}[\theta_q \varepsilon(t-q), y(t-k)] \end{aligned}$$

また、 q 次までの共分散は MA(q) の影響を受けるので、 q 次までの自己共分散を簡単に表すのは難しい。しかし、 $p > q$ (つまり、 $p \geq q+1$) であれば AR(p) の影響のみになる。つまり、 $q+1$ 次以降($k \geq q+1$)の自己共分散は以下のように表すことができる(なお、 $k \geq q+1$ であり、 $k=0$ とおくことができず、分散 γ_0 を簡単に表現することは難しい)。

$$\begin{aligned} \gamma_k &= 0 + \phi_1 \text{Cov}[y(t-1), y(t-k)] + \phi_2 \text{Cov}[y(t-2), y(t-k)] + \cdots + \phi_p \text{Cov}[y(t-p), y(t-k)] \\ &\quad + \text{Cov}[\varepsilon(t), y(t-k)] + \theta_1 \text{Cov}[\varepsilon(t-1), y(t-k)] + \theta_2 \text{Cov}[\varepsilon(t-2), y(t-k)] + \cdots + \theta_q \text{Cov}[\varepsilon(t-q), y(t-k)] \\ &= \phi_1 \text{Cov}[y(t-1), y(t-k)] + \phi_2 \text{Cov}[y(t-2), y(t-k)] + \cdots + \phi_p \text{Cov}[y(t-p), y(t-k)] + 0 + \theta_1 \cdot 0 + \theta_2 \cdot 0 + \cdots + \theta_q \cdot 0 \end{aligned}$$

従って、自己共分散 $\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}$ 但し、 $k \geq q+1$ $\cdots$ (式7-3)

(3) ARMA(p, q) 過程の自己相関係数

(式7-3)の両辺を γ_0 で割って、 $\rho_k = \gamma_k / \gamma_0 = (\phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}) / \gamma_0 = \phi_1 (\gamma_{k-1} / \gamma_0) + \phi_2 (\gamma_{k-2} / \gamma_0) + \cdots + \phi_p (\gamma_{k-p} / \gamma_0)$

従って、自己相関係数 $\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p}$ 但し、 $k \geq q+1$ $\cdots$ (式7-4)

(4) 定常性と反転可能性

ARMA(p, q) は、AR 過程と MA 過程を組み合わせた過程であるが、MA(q) 過程は常に定常であり、ARMA(P, q) 過程が定常であるには AR(p) 過程が定常でなくてはならない。
また、MA(q) 過程が反転可能であれば、ARMA(p, q) 過程は反転可能となり(後述)、このとき、ARMA(p, q) 過程は AR(∞) 過程で表すことができる。

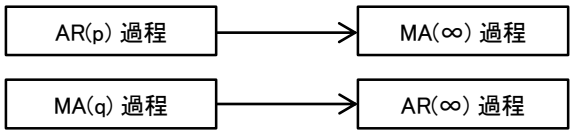
ARMA(2, 2) 過程とそのコレログラムの例を(付録4)に載せる。また、(表7-1)には、ARMA(p, q) 過程をもとにした各種時系列モデル(過程)を挙げておく(個別には別途記載)。

(表7-1)各種時系列モデル

条件	モデル式 (定数 c を割愛)	名称
$p \neq 0, q \neq 0$	$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p) + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \cdots + \theta_q \varepsilon(t-q)$	ARMA 過程
$p = 0, q \neq 0$	$y(t) = \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \cdots + \theta_q \varepsilon(t-q)$	MA 過程
$p \neq 0, q = 0$	$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p) + \varepsilon(t)$	AR 過程
$p = 0, q = 0$	$y(t) = \varepsilon(t)$	ホワイトノイズ
$p = 1, q = 0$	$y(t) = \phi_1 y(t-1) + \varepsilon(t)$	マルコフ過程
$p = 1, q = 0, \phi_1 = 1$	$y(t) = y(t-1) + \varepsilon(t)$	ランダムウォーク

8. 時系列モデルの反転可能性(Invertibility)

MA(q) 過程の時系列モデルを AR(∞) 過程の時系列モデルに置き換えて表すこと、もしくは、AR(p) 過程の時系列モデルを MA(∞) 過程の時系列モデルに置き換えて表すこと、が可能なことを時系列モデルの反転可能性という。
特に、観測できない過去のホワイトノイズを使う MA(q) 過程の時系列モデルを観測可能な過去の時系列データを使う AR(∞) 過程の時系列モデルに置き換えて表すことができれば、実用上、極めて有用である。
これによって MA 過程を AR 過程で考えることが可能になる。時系列モデルの反転の事例は、(付録7)に載せる。



(図(8-1) 反転可能性

(1-1) AR(1) 過程 → MA(∞) 過程

AR(1) 過程は、 $y(t) = \phi_1 y(t-1) + \varepsilon(t)$
この式に、 $y(t-1) = \phi_1 y(t-2) + \varepsilon(t-1)$ を代入すると、 $y(t) = \phi_1 (\phi_1 y(t-2) + \varepsilon(t-1)) + \varepsilon(t) = \phi_1^2 y(t-2) + \phi_1 \varepsilon(t-1) + \varepsilon(t)$
この式変形を繰り返すと、 $y(t) = \phi_1^m y(t-m) + \sum_{k=0}^{m-1} \phi_1^k \varepsilon(t-k)$
ここで、 $\phi_1 < 1$ であれば、 $m \rightarrow \infty$ で、 $\phi_1^m \rightarrow 0$ になるので、 $y(t) = \sum_{k=0}^{\infty} \phi_1^k \varepsilon(t-k)$ となる。この時系列データ $y(t)$ は、MA(∞) 過程である。

(1-2) AR(p) 過程 → MA(∞) 過程

(式6-1)で、 $c = 0$ とおくと、AR(p) 過程は、 $y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t)$ 両辺を Z 変換すると(付録D-1)、
 $Z[y(t)] = \phi_1 Z[y(t-1)] + \phi_2 Z[y(t-2)] + \dots + \phi_p Z[y(t-p)] + Z[\varepsilon(t)] \rightarrow Z[y(t)] = \phi_1 z^{-1} Z[y(t)] + \phi_2 z^{-2} Z[y(t)] + \dots + \phi_p z^{-p} Z[y(t)] + Z[\varepsilon(t)]$
従って、 $(1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p}) Z[y(t)] = Z[\varepsilon(t)] \rightarrow Z[y(t)] = (1 / (1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p})) Z[\varepsilon(t)]$
ここで、特性方程式 $1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0$... (式8-1) の全ての根の絶対値が 1 より大きいとき(つまり、AR(p) 過程が定常であれば)、マクローリン展開した(別途資料「いろいろな確率分布」を参照) $1 / (1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p}) = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots$ (無限級数) は、収束する。
つまり、 $Z[y(t)] = (1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots) Z[\varepsilon(t)]$ と表すことができる。
従って、 $y(t) = \varepsilon(t) + \alpha_1 \varepsilon(t-1) + \alpha_2 \varepsilon(t-2) + \dots$ この式は MA(∞) 過程である。
以上、AR(p) 過程が定常であれば(特性方程式の全ての根の絶対値が 1 より大きければ)、この AR(p) 過程を MA(∞) 過程で表すことができる、反転可能である、ことが分かる。
ちなみに、 $p = 1$ のときの特性方程式は、 $1 - \phi_1 z^{-1} = 0$ で、根は $z^{-1} = 1 / \phi_1$ 従って、 $|z^{-1}| > 1$ は、 $|\phi_1| < 1$ と言い換えることができる。

(2-1) MA(1) 過程 → AR(∞) 過程

MA(1) 過程は、 $y(t) = \theta_1 \varepsilon(t-1) + \varepsilon(t) \rightarrow \varepsilon(t) = y(t) - \theta_1 \varepsilon(t-1)$
この式に、 $\varepsilon(t-1) = y(t-1) + \theta_1 \varepsilon(t-2)$ を代入すると、 $\varepsilon(t) = y(t) - \theta_1 (y(t-1) + \theta_1 \varepsilon(t-2)) = y(t) - \theta_1 y(t-1) + \theta_1^2 \varepsilon(t-2)$
この式変形を繰り返すと、 $\varepsilon(t) = \sum_{k=0}^{m-1} (-\theta_1)^k y(t-k) + (-\theta_1)^m \varepsilon(t-m)$
ここで、 $\theta_1 < 1$ であれば、 $m \rightarrow \infty$ で、 $(-\theta_1)^m \rightarrow 0$ になるので、 $\varepsilon(t) = \sum_{k=0}^{\infty} (-\theta_1)^k y(t-k) = y(t) + \sum_{k=1}^{\infty} (-\theta_1)^k y(t-k)$ となる。
つまり、 $y(t) = -\sum_{k=1}^{\infty} (-\theta_1)^k y(t-k) + \varepsilon(t)$ となる。この時系列データ $y(t)$ は、AR(∞) 過程である。

(2-2) MA(q) 過程 → AR(∞) 過程

(式5-1)で、 $\mu = 0$ において、MA(q) 過程は、 $y(t) = \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2) + \dots + \theta_q \varepsilon(t-q)$
両辺を Z 変換すると(付録D-1)、
 $Z[y(t)] = Z[\varepsilon(t)] + \theta_1 Z[\varepsilon(t-1)] + \theta_2 Z[\varepsilon(t-2)] + \dots + \theta_q Z[\varepsilon(t-q)]$
 $\rightarrow Z[y(t)] = Z[\varepsilon(t)] + \theta_1 z^{-1} Z[\varepsilon(t)] + \theta_2 z^{-2} Z[\varepsilon(t)] + \dots + \theta_q z^{-q} Z[\varepsilon(t)] \rightarrow Z[y(t)] = (1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_q z^{-q}) Z[\varepsilon(t)]$
従って、 $Z[\varepsilon(t)] = (1 / (1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_q z^{-q})) Z[y(t)]$
ここで、特性方程式 $1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_q z^{-q} = 0$... (式8-2) の全ての根の絶対値が 1 より大きいとき、マクローリン展開した $1 / (1 + \theta_1 z^{-1} + \theta_2 z^{-2} + \dots + \theta_q z^{-q}) = 1 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots$ (無限級数) は、収束する。
つまり、 $Z[\varepsilon(t)] = (1 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots) Z[y(t)]$ と表すことができる。
よって、 $\varepsilon(t) = y(t) + \beta_1 y(t-1) + \beta_2 y(t-2) + \dots$
式変形して、 $y(t) = -\beta_1 y(t-1) - \beta_2 y(t-2) - \dots + \varepsilon(t)$ この式は、AR(∞) 過程である。
以上、MA(q) 過程の特性方程式の全ての根の絶対値が 1 より大きければ、この MA(q) 過程を AR(∞) 過程で表すことができる、反転可能である、ことが分かる。

9. AR(p) 過程のパラメータ推定 (Yule-Walker 法)

(式6-1)で表される AR(p) 過程の係数 $\phi_1, \phi_2, \dots, \phi_p$ を推定する。
ただ、推定にあたっては、(式6-1)の代わりに、定数 c を除いた(式9-1)を使う。
 $y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots$ (式6-1)
 $y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots$ (式9-1)

なお、定数 c は、推定した係数と $y(t)$ の期待値を使って、以下のようにして推定することができる。
 $z(t) = y(t) - \mu$ とおくと、 $E[z(t)] = E[y(t) - \mu] = E[y(t)] - \mu = \mu - \mu = 0, E[z(t-1)] = E[y(t-1) - \mu] = 0, \dots$
(式9-1)を書き換えて、これらの関係を使うと、

$$\begin{aligned} y(t) &= z(t) + \mu = c + \phi_1 (z(t-1) + \mu) + \phi_2 (z(t-2) + \mu) + \dots + \phi_p (z(t-p) + \mu) + \varepsilon(t) \\ \rightarrow z(t) &= \phi_1 z(t-1) + \phi_2 z(t-2) + \dots + \phi_p z(t-p) + \varepsilon(t) + c - \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p) \\ \rightarrow E[z(t)] &= E[\phi_1 z(t-1) + \phi_2 z(t-2) + \dots + \phi_p z(t-p)] + E[\varepsilon(t)] + E[c] - E[\mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)] \\ \rightarrow 0 &= 0 + 0 + \dots + 0 + 0 + c - \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p) \\ \rightarrow c &= \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p) \quad \dots \text{(式9-2)} \end{aligned}$$

以下、AR(p) 過程の係数 $\phi_1, \phi_2, \dots, \phi_p$ を推定する方法として、Yule-Walker 法と最尤法を取り上げる。

1) Yule-Walker 法

AR(p) 過程の自己共分散の式(Yule-Walker の式)(式6-5)を再掲する。なお、 $\phi_m(p)$ と $\sigma(p)$ は、 p 次のAR(p) 過程での値であることを示す。

$$\rho_k = \sum_{m=1}^p \phi_m(p) \rho_{k-m} \quad \dots \text{(式6-5)}$$

この式を、 $k=1, \dots, p$ において展開する。 $\phi_m = \phi_{-m}$ であることを利用して、

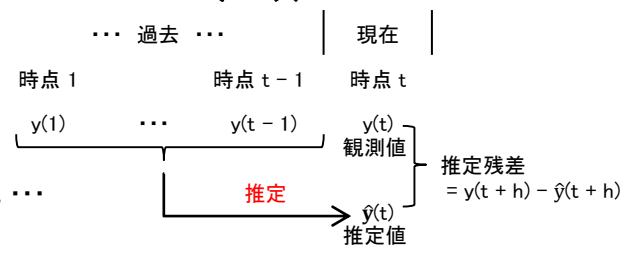
$$\begin{aligned} \rho_1 &= \sum_{m=1}^p \phi_m(p) \rho_{1-m} = \phi_1(p) \rho_0 + \phi_2(p) \rho_{-1} + \dots + \phi_p(p) \rho_{1-p} = \phi_1(p) \rho_0 + \phi_2(p) \rho_1 + \dots + \phi_p(p) \rho_{p-1} \\ \rho_2 &= \sum_{m=1}^p \phi_m(p) \rho_{2-m} = \phi_1(p) \rho_1 + \phi_2(p) \rho_0 + \dots + \phi_p(p) \rho_{2-p} = \phi_1(p) \rho_1 + \phi_2(p) \rho_0 + \dots + \phi_p(p) \rho_{p-2} \\ &\vdots \\ \rho_p &= \sum_{m=1}^p \phi_m(p) \rho_{p-m} = \phi_1(p) \rho_{p-1} + \phi_2(p) \rho_{p-2} + \dots + \phi_p(p) \rho_{p-p} = \phi_1(p) \rho_{p-1} + \phi_2(p) \rho_{p-2} + \dots + \phi_p(p) \rho_0 \end{aligned}$$

これらの式を行列で表すと、 $\rho_0 = 1$ であるので、

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} \rho_0 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & \rho_0 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & \rho_0 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} \quad \dots \text{(式9-3)}$$
$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad \dots \text{(式9-4)}$$

この(式9-3)を使って、係数 $\phi_1, \phi_2, \dots, \phi_p$ を推定することができる。しかし、この式を直接解く代わりに、(式9-3)を起点にして再帰的に係数を求める方法がある。それが(付録13)に載せる**レヴィンソン・ダービンのアルゴリズム**である。(付録15-1)にレヴィンソン・ダービンのアルゴリズムを使った AR(2) 過程の係数推定の計算例を載せる。
また、AR(2) 過程(式6-1)を重回帰式と見なして係数を推定する計算例を(付録15-2)に載せる。
下記の(注)で示すように、**レヴィンソン・ダービンのアルゴリズムは最小二乗法を使用するといった点で、AR(2) 過程を重回帰式と見なして係数を推定する方法と同じである。**

(注) 時点 t の時系列データ $y(t)$ の推定値は、 $\sum_{i=1}^p \phi_i y(t-i)$ である。その推定残差は、 $y(t) - \sum_{i=1}^p \phi_i y(t-i)$ になる。よって、推定残差の平方和の期待値は、
 $E[\varepsilon^2(t)] = E[(y(t) - \sum_{i=1}^p \phi_i y(t-i))^2] = E[y^2(t)] - 2E[\sum_{i=1}^p \phi_i y(t) y(t-i)] + E[\sum_{i=1}^p \sum_{j=1}^p \phi_i \phi_j y(t-i) y(t-j)] = \gamma_0 - 2 \sum_{i=1}^p \phi_i \gamma_i + \sum_{i=1}^p \sum_{j=1}^p \phi_i \phi_j \gamma_{i-j}$
この両辺を ϕ_i で偏微分し 0 とおき、 γ_0 で割る。 $\partial E[\varepsilon^2(t)] / \partial \phi_i = -2 \gamma_i + 2 \sum_{j=1}^p \phi_j \gamma_{i-j} = 0 \rightarrow -\gamma_i / \gamma_0 + \sum_{j=1}^p \phi_j \gamma_{i-j} / \gamma_0 = -\rho_i + \sum_{j=1}^p \phi_j \rho_{i-j} = 0$
この式は、(式9-3)に等しい。つまり、Yule-Walker の式は、(図9-1)に示す推定残差の期待値を最小二乗法を使って最小にするパラメータの式を表していると言える。
つまり、AR(p) の係数 $\phi_1, \phi_2, \dots, \phi_p$ を Yule-Walker 法で推定する方法(レヴィンソン・ダービンのアルゴリズムで推定する方法)は、AR(p) 過程の式(9-1)を目的変数が $y(t)$, 説明変数が $y(t-1), y(t-2), \dots, y(t-p)$, 偏回帰係数が $\phi_1, \phi_2, \dots, \phi_p$ である重回帰式であると見なせば、最小二乗法を使って重回帰式の偏回帰係数を求める方法と等しくなる(別途資料「重回帰分析の概要」を参照)。(付録12)に、(式9-1)を重回帰式と見なしたときに、(式9-3)が導かれることを示しておく。
また、参考までに、(式9-1)を重回帰式と見なしたときに、単回帰式のパラメータとして AR(1) 過程のパラメータを推定する方法を(付録14)に載せておく。



(図9-1) AR(p) 過程の推定

10. AR(p) 過程のパラメータ推定（最尤法）と AIC(赤池情報量基準)

最尤法によるモデルパラメータの推定方法は、観測データを最もよく説明できるモデルのパラメータを推定値とする方法である（別途資料「重回帰分析の概要」を参照）。つまり、パラメータを変動させて観測データと推定データの誤差を最小にするようなパラメータを求め、それをパラメータの最尤推定量であるとする方法である。

(1) 最尤法によるパラメータ推定

AR(p) 過程を定常過程として、重回帰式に見立てて(式10-1)と(式10-2)で表す（別途資料「重回帰分析の概要」を参照）。ここでは、 $\phi_0 = c$ とおく。

$$y = Y \phi + \varepsilon \quad \cdots \text{(付式10-1)}$$

$$y(t) = \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-T+1) \end{bmatrix}, Y = \begin{bmatrix} 1 & y(t-1) & y(t-2) & \cdots & y(t-p) \\ 1 & y(t-2) & y(t-3) & \cdots & y(t-p-1) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & y(t-T) & y(t-T-1) & \cdots & y(t-p-T+1) \end{bmatrix}, \phi = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \vdots \\ \phi_p \end{bmatrix}, \varepsilon(t) = \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \vdots \\ \varepsilon(t-T+1) \end{bmatrix}, E[\varepsilon(t)] = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, V[\varepsilon(t)] = \sigma^2 = \begin{bmatrix} \sigma^2 \\ \sigma^2 \\ \vdots \\ \sigma^2 \end{bmatrix} \quad \cdots \text{(式10-2)}$$

但し、 $\varepsilon(t)$ は正規分布に従うとする。つまり、 $\varepsilon(t) \sim N(0, \sigma^2) \quad \cdots \text{(式10-3)}$
 なお、 $\hat{y}(t)$ を予測値とし、 $E[y(t)] = \hat{y}(t)$, $V[y(t)] = V[\varepsilon(t)] = \sigma^2$ であるので、 $y(t)$ の分布は $\varepsilon(t)$ と同じ正規分布に従う。従って、 $y(t) \sim N(\hat{y}(t), \sigma^2) \quad \cdots \text{(式10-4)}$
 ところで、集合に関する乗法定理、つまり、同時確率密度 = 条件付き確率密度 × 周辺確率密度 が成り立つので(付録M と別途資料「ベイズ統計の概要」を参照)、

$$\begin{aligned} T \leq p \text{ では、} \overbrace{f(y(T), y(T-1), \cdots, y(0))}^{\text{同時確率密度}} &= \overbrace{f(y(T) | y(T-1), y(T-2), \cdots, y(0))}^{\text{条件付確率密度}} \underbrace{f(y(T-1), y(T-2), \cdots, y(0))}_{\text{同時確率密度}} \\ &= f(y(T) | y(T-1), y(T-2), \cdots, y(0)) \underbrace{f(y(T-1) | y(T-2), y(T-3), \cdots, y(0))}_{\text{条件付確率密度}} \underbrace{f(y(T-2), y(T-3), \cdots, y(0))}_{\text{同時確率密度}} \\ &= f(y(T) | y(T-1), y(T-2), \cdots, y(0)) \underbrace{f(y(T-1) | y(T-2), y(T-3), \cdots, y(0))}_{\text{条件付確率密度}} \cdots \underbrace{f(y(1) | y(0))}_{\text{条件付確率密度}} f(y(0)) \\ &= \prod_{t=1}^T f(y(t) | y(t-1), y(t-2), \cdots, y(0)) f(0) \quad \cdots \text{(式10-5)} \end{aligned}$$

この(式10-5)を、パラメータ ϕ, σ^2 がある値を取るときの時系列データの観測値が得られる確率を求める式であるとみて、尤度関数として捉える（なお、定数項 $f(0)$ は省略）。

$$L(\phi, \sigma^2) = \prod_{t=1}^T f(y(t) | y(t-1), y(t-2), \cdots, y(0)) \quad \cdots \text{(式10-6)}$$

時系列データ $y(t)$ は独立であり、(式10-4)の示すように正規分布に従うので、確率密度関数を以下のように表すことができる（別途資料「いろいろな確率分布」を参照）。

$$f(y(T) | y(T-1), \cdots, y(0)) = (1 / (2 \pi \sigma^2)^{1/2}) \exp(- (y(t) - \hat{y}(t))^2 / (2 \sigma^2)) \quad \cdots \text{(式10-7)}$$

従って、尤度関数(式10-6)に(式10-7)を代入して、さらに計算を簡便にするために対数をとると、

$$\ln L(\phi, \sigma^2) = \ln \left\{ \prod_{t=1}^T (1 / (2 \pi \sigma^2)^{1/2}) \exp(- (y(t) - \hat{y}(t))^2 / (2 \sigma^2)) \right\} = - (T/2) \ln (2 \pi \sigma^2) - (1 / 2 \sigma^2) \sum_{t=1}^T (y(t) - \hat{y}(t))^2 \quad \cdots \text{(式10-8)}$$

この式は、別途資料「重回帰分析の概要」の対数尤度関数と同じであり、 ϕ と σ^2 の最尤推定量 $\hat{\phi}$ と $\hat{\sigma}^2$ は以下のように求まる。

$$\hat{\phi} = (Y^T Y)^{-1} Y^T y \quad \cdots \text{(式10-9)} \quad \hat{\sigma}^2 = \varepsilon^T(t) \varepsilon(t) / T = (\sum_{t=1}^T (y(t) - \hat{y}(t))^2) / T \quad \cdots \text{(式10-10)}$$

(式10-10)が示すように、最尤法によるパラメータの推定値は最小二乗法で求まる式になっており、Yule-Walker 法で求める結果と同じである。

(2) AIC(赤池情報量基準)

(式10-8)の対数尤度関数を使用して、AIC を求めることができる（別途資料「重回帰分析の概要」を参照）。

$$AIC = T (\ln (2 \pi \hat{\sigma}^2)^{1/2}) + 1 + 2 p + 4 \quad \cdots \text{(式10-11)}$$

AIC は値の大きさに意味はなく、相対的な大小を比較する指標であるため、使用にあたっては、(式10-11)の定数項である 4 を割愛しても差し支えない。

時系列過程では、AIC は、AR(p), MA(q), ARMA(p, q) のモデル選択と、次数 p, q の推定に使用されるが、(付録16)には、時系列過程が AR(p) 過程であることを前提に、AIC を使用して AR(p) 過程の次数 p を求めるという事例を載せる。

なお、AR(p) 過程のモデル次数 p は、偏自己相関係数(付録11)が p + 1 次以降 0 になることを利用して推定することもできる(付録15-1)。

11. AR(p) 過程の点予測と区間予測

AR(p) 過程で過去の時系列データを使用して将来を予測する方法について整理する。計算事例として、AR(1) 過程の点予測は(付録17)、AR(p) 過程の点予測は(付録18)、AR(p) 過程の区間予測は(付録19)に示す。

(1) 準備

(図11-1)に示すように、現時点 t 以前の過去の時系列データ $\Omega(t) = \{y(t), y(t-1), \dots, y(1)\}$ を使って、現時点 t より h 期先の時点 $t+h$ の $y(t+h)$ を予測する。その最適な予測値を $\hat{y}(t+h)$ とする。

まず、過去の時系列データ $\Omega(t)$ を所与とした h 期先の $y(t+h)$ の条件付き期待値(付録M)を次のように定義する。

$$E[y(t+h) | \Omega(t)] = \mu(t+h|t) \quad \dots \text{(式11-1)} \quad (\text{定常過程であれば、} E[y(t+h) | \Omega(t)] = \mu)$$

ここで、下記の条件のもとで、 h 期先の $y(t+h)$ の条件付き期待値を求める。

- ① $t \leq t$ のとき、 $E[y(t) | \Omega(t)] = y(t)$ $\dots$ (式11-2) 時点 t 以前の過去の $y(t)$ の条件付き期待値は、 $y(t)$ である。(図11-1) AR(p) 過程の予測
- ② $k > 0$ のとき、 $E[\varepsilon(t+k) | \Omega(t)] = 0$ $\dots$ (式11-3) $\varepsilon(t)$ は正規ホワイトノイズであり、時点 t より将来の $\varepsilon(t)$ の条件付き期待値は 0 である。

観測データ $y(t+h)$ と 予測値 $\hat{y}(t+h)$ の平均二乗誤差 MSE (Mean Squared Error) を最小とする予測を最適予測とする。

$$\begin{aligned} \text{ここで、} \text{MSE}(\hat{y}(t+h|t)) &= E[(y(t+h) - \hat{y}(t+h|t) | \Omega(t))^2] = E[(y(t+h) - \mu(t+h|t) + (\mu(t+h|t) - \hat{y}(t+h|t)) | \Omega(t))^2] \\ &= E[(y(t+h) - \mu(t+h|t) | \Omega(t))^2 - 2(y(t+h) - \mu(t+h|t))(\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t)) + ((\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t))^2)] \\ &= E[(y(t+h) - \mu(t+h|t) | \Omega(t))^2] - 2E[(y(t+h) - \mu(t+h|t))(\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t))] + E[(\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t))^2] \end{aligned}$$

$\mu(t+h|t)$ と $\hat{y}(t+h|t)$ の条件付き条件は、時点 t のデータに依存するという条件であるが、この条件は $\Omega(t)$ のデータに依存するという条件に含まれている。

$$\text{つまり、} E[(y(t+h) - \mu(t+h|t))(\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t))] = (\mu(t+h|t) - \hat{y}(t+h|t)) E[y(t+h) - \mu(t+h|t) | \Omega(t)]$$

$$\text{ここで、(式11-1)より、} E[y(t+h) - \mu(t+h|t) | \Omega(t)] = E[y(t+h) | \Omega(t)] - E[\mu(t+h|t) | \Omega(t)] = \mu(t+h|t) - \mu(t+h|t) = 0$$

$$\text{従って、} \text{MSE}(\hat{y}(t+h|t)) = E[(y(t+h) - \hat{y}(t+h|t) | \Omega(t))^2] = 0 + E[(\mu(t+h|t) - \hat{y}(t+h|t) | \Omega(t))^2] \quad \dots \text{(式11-4)}$$

定常過程では分散 $V[y(t+h|t) | \Omega(t)]$ は時間に依らず一定であるから、 h 期先の時点 $t+h$ の最適予測値 $\hat{y}(t+h|t)$ は、(式11-4)を最小化する、

$$\hat{y}(t+h|t) = \mu(t+h|t) \quad (\text{条件付き期待値}) \quad \dots \text{(式11-5)} \quad \text{であることが分かる。}$$

(2) AR(p) 過程の点予測 (point forecast)

AR(p) 過程では、 $y(t+1)$ は、(式6-1)より、次のように表すことができる。 $y(t+1) = c + \phi_1 y(t) + \phi_2 y(t-1) + \dots + \phi_p y(t-p) + \varepsilon(t+1) \quad \dots \text{(式11-6)}$

(式11-6)から分かることは、将来の $y(t+1)$ が、現在 $y(t)$ と過去の $y(t-1), \dots, y(t-p)$ と将来の $\varepsilon(t+1)$ で表されていることである。現在までの $y(t)$ は(式11-2)より実際の値を使用し、将来の ε は(式11-3)より 0 で置き換える。

具体的には、1 期先の最適予測値は、 $h=1$ とした(式11-5)に(式11-6)を代入し、(式11-2)と(式11-3)を使って求める。2 期先以降の最適予測値も同様にして求める。

$$\begin{aligned} \hat{y}(t+1|t) &= \mu(t+1|t) = E[y(t+1) | \Omega(t)] = E[c + \phi_1 y(t) + \phi_2 y(t-1) + \dots + \phi_p y(t-p) + \varepsilon(t+1) | \Omega(t)] \\ &= c + \phi_1 E[y(t) | \Omega(t)] + \phi_2 E[y(t-1) | \Omega(t)] + \dots + \phi_p E[y(t-p) | \Omega(t)] + E[\varepsilon(t+1) | \Omega(t)] = c + \phi_1 y(t) + \phi_2 y(t-1) + \dots + \phi_p y(t-p) + 0 \quad \dots \text{(式11-7)} \end{aligned}$$

$$\begin{aligned} \hat{y}(t+2|t) &= \mu(t+2|t) = E[y(t+2) | \Omega(t)] = E[c + \phi_1 y(t+1) + \phi_2 y(t) + \dots + \phi_p y(t+1-p) + \varepsilon(t+2) | \Omega(t)] \\ &= c + \phi_1 E[y(t+1) | \Omega(t)] + \phi_2 E[y(t) | \Omega(t)] + \dots + \phi_p E[y(t+1-p) | \Omega(t)] + E[\varepsilon(t+2) | \Omega(t)] = c + \phi_1 \hat{y}(t+1|t) + \phi_2 y(t) + \dots + \phi_p y(t+1-p) + 0 \\ &= c + \phi_1 E[c + \phi_1 y(t) + \phi_2 y(t-1) + \dots + \phi_p y(t-p) + \varepsilon(t+1) | \Omega(t)] + \phi_2 y(t) + \dots + \phi_p y(t+1-p) \\ &= c + \phi_1 (c + \phi_1 y(t) + \phi_2 y(t-1) + \dots + \phi_p y(t-p) + 0) + \phi_2 y(t) + \dots + \phi_p y(t+1-p) = c + \phi_1 \hat{y}(t+1|t) + \phi_2 y(t) + \dots + \phi_p y(t+1-p) \quad \dots \text{(式11-8)} \end{aligned}$$

$$\text{同様に、} h \text{ 期先の時点 } t+h \text{ の最適予測値は、} \hat{y}(t+h|t) = c + \phi_1 \hat{y}(t+h-1|t) + \phi_2 \hat{y}(t+h-2|t) + \dots + \phi_p \hat{y}(t+h-p|t) \quad \dots \text{(式11-9)}$$

しかし、(式11-9)の代わりに、簡便な計算手順で h 期先の時点 $t+h$ の最適予測値を求める方法がある。AR(1) 過程を例にとれば(付録17)、

$$y(t+1) = c + \phi_1 y(t) + \varepsilon(t+1), \quad y(t+2) = c + \phi_1 y(t+1) + \varepsilon(t+2) = c + \phi_1 (c + \phi_1 y(t) + \varepsilon(t+1)) + \varepsilon(t+2) = (1 + \phi_1) c + \phi_1^2 y(t) + \varepsilon(t+2) + \phi_1 \varepsilon(t+1), \dots$$

$$\text{同様に、} h \text{ 期先の時点 } t+h \text{ の } y(t+h) \text{ は、} y(t+h) = (1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1}) c + \phi_1^h y(t) + \varepsilon(t+h) + \phi_1 \varepsilon(t+h-1) + \dots + \phi_1^{h-1} \varepsilon(t+1)$$

$$\text{(式11-3)が成立するならば、} \hat{y}(t+h|t) = (1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1}) c + \phi_1^h y(t) + 0 + 0 + \dots + 0$$

$$\text{従って、(付録A)の等比数列の公式(初項=1, 公比=}\phi_1\text{)を使用して、} \hat{y}(t+h|t) = (1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1}) c + \phi_1^h y(t) = (1 - \phi_1^h) c / (1 - \phi_1) + \phi_1^h y(t) \quad \dots \text{(式11-10)}$$

(式11-10)は、最適予測値 $\hat{y}(t+h|t)$ は、現在の値 $y(t)$ のみに依存すること、つまり、 $y(t-1), y(t-2), \dots$ など過去の値は予測に影響しないことを示している

(3) AR(p) 過程の区間予測 (interval forecast)

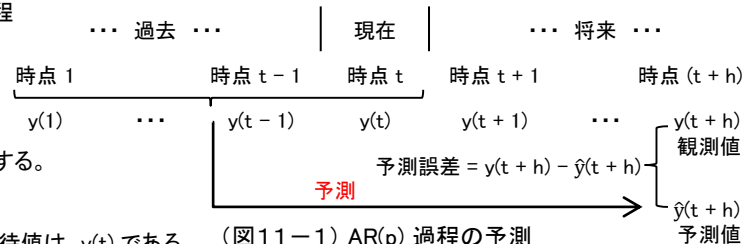
h 期先の予測を $\hat{y}(t+h|t)$ の 1 点で行う点予測に対して、 $\hat{y}(t+h|t)$ を $\alpha\%$ の確率で含むような区間で予測する区間予測についてまとめる。

(式11-6)の $\varepsilon(t+1)$ と同じく $\varepsilon(t+h)$ が正規分布 $(0, \sigma^2)$ に従うとすれば、 $y(t+h|t)$ は、平均 $= \hat{y}(t+h)$ 、分散 $= \sigma^2 = \text{MSE}(\hat{y}(t+h|t))$ の正規分布に従う。

ここで、標準化した $(y(t+h|t) - \text{平均}) / \text{分散}^{1/2}$ は、平均 $= 0$ 、分散 $= 1$ の標準正規分布に従い、標準正規分布の上側 2.5 % 点は 1.960、下側 2.5 % 点は -1.960 であるので、
 $-1.960 \leq (y(t+h|t) - \hat{y}(t+h|t)) / (\text{MSE}(\hat{y}(t+h|t)))^{1/2} \leq 1.960$

よって、 h 期先の $\hat{y}(t+h|t)$ を 95 % の確率で含むような予測区間は、以下の通りになる。

$$\hat{y}(t+h|t) - 1.96 (\text{MSE}(\hat{y}(t+h|t)))^{1/2} \leq y(t+h|t) \leq \hat{y}(t+h|t) + 1.96 (\text{MSE}(\hat{y}(t+h|t)))^{1/2} \quad \dots \text{(式11-11)}$$



12. MA(q) 過程の点予測

MA(q) 過程の点予測を考えるにあたって、(式11-3)に加えて以下の条件を加える。

③ $\tau \leq t$ のとき、 $E[\varepsilon(\tau) | \Omega(t)] = \varepsilon(\tau) \quad \dots$ (式12-1) 時点 t 以前の過去の $\varepsilon(t)$ の条件付き期待値は、 $\varepsilon(t)$ である。

MA(q) 過程では、 $\hat{y}(t+1)$ と $\hat{y}(t+1)$ は、(式5-1)に(式11-1)、(式11-3)、(式12-1)を反映して、それぞれ (式12-2) と (式12-3) のように表すことができる。

$$\begin{aligned} \hat{y}(t+1|t) &= \mu(t+1|t) = E[y(t+1|t)] = E[\mu + \varepsilon(t+1) + \theta_1 \varepsilon(t) + \theta_2 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+1) | t] \\ &= \mu + E[\varepsilon(t+1)|t] + \theta_1 E[\varepsilon(t)|t] + \theta_2 E[\varepsilon(t-1)|t] + \dots + \theta_q E[\varepsilon(t-q+1)|t] \\ &= \mu + 0 + \theta_1 \varepsilon(t) + \theta_2 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+1) = \mu + \theta_1 \varepsilon(t) + \theta_2 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+1) \quad \dots \text{(式12-2)} \end{aligned}$$

$$\begin{aligned} \hat{y}(t+2|t) &= \mu(t+2|t) = E[y(t+2|t)] = E[\mu + \varepsilon(t+2) + \theta_1 \varepsilon(t+1) + \theta_2 \varepsilon(t) + \dots + \theta_q \varepsilon(t-q+2) | t] \\ &= \mu + E[\varepsilon(t+2)|t] + \theta_1 E[\varepsilon(t+1)|t] + \theta_2 E[\varepsilon(t)|t] + \dots + \theta_q E[\varepsilon(t-q+2)|t] \\ &= \mu + 0 + 0 + \theta_2 \varepsilon(t) + \theta_3 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+2) = \mu + \theta_2 \varepsilon(t) + \theta_3 \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+2) \quad \dots \text{(式12-3)} \end{aligned}$$

同様に、 h 期先の最適予測値が求まる。

$$\left. \begin{aligned} h \leq q-1 \text{ のとき、} \quad \hat{y}(t+h|t) &= \mu + \theta_h \varepsilon(t) + \theta_{h+1} \varepsilon(t-1) + \dots + \theta_q \varepsilon(t-q+h) \\ h = q \text{ のとき、} \quad \hat{y}(t+h|t) &= \mu + \theta_q \varepsilon(t) \\ h \geq q+1 \text{ のとき、} \quad \hat{y}(t+h|t) &= \mu \end{aligned} \right\} \quad \dots \text{(式12-4)}$$

(式12-4)より、 h 期先の $\hat{y}(t+h|t)$ を求めるには、観測できない $\varepsilon(t)$ の値が必要であることが分かる。

観測可能な無限個の $y(t)$ が存在すれば、MA(q) 過程を AR(∞) 過程で表現することができるので(反転可能性)、 $\varepsilon(t)$ の値を求めることができる。

しかし、観測可能な無限個の $y(t)$ は存在しないため、やはり $\varepsilon(t)$ の値を求めることはできない。

そこで、観測できない $\varepsilon(t)$ の代わりに、下記に示す $\hat{\varepsilon}(t)$ を $\varepsilon(t)$ の近似値として使用する。 $\hat{\varepsilon}(t)$ の計算には、 $y(t)$ を使用して、下記に示すように逐次 計算で求める。

この $\hat{\varepsilon}(t)$ を $\varepsilon(t)$ の代わりとして(式5-1)に代入して、 h 期先の最適予測値 $\hat{y}(t+h|t)$ を求める。

・逐次計算補法

(式5-1)より、 $\varepsilon(t) = y(t) - \mu - \theta_1 \varepsilon(t-1) - \theta_2 \varepsilon(t-2) - \dots - \theta_q \varepsilon(t-q)$

であるので、この式を使って、また、(式11-3)と(式12-1)を反映して、逐次、 $\varepsilon(t)$ の近似値 $\hat{\varepsilon}(t)$ を逐次求める。

$$\left. \begin{aligned} \hat{\varepsilon}(1) &= y(1) - \mu \\ \hat{\varepsilon}(2) &= y(2) - \mu - \theta_1 \hat{\varepsilon}(1) \\ \hat{\varepsilon}(3) &= y(3) - \mu - \theta_1 \hat{\varepsilon}(2) - \theta_2 \hat{\varepsilon}(1) \\ \hat{\varepsilon}(4) &= y(4) - \mu - \theta_1 \hat{\varepsilon}(3) - \theta_2 \hat{\varepsilon}(2) - \theta_3 \hat{\varepsilon}(1) \\ &\dots \\ \hat{\varepsilon}(q) &= y(q) - \mu - \theta_1 \hat{\varepsilon}(q-1) - \theta_2 \hat{\varepsilon}(q-2) - \dots - \theta_{q-1} \hat{\varepsilon}(1) \\ \hat{\varepsilon}(q+1) &= y(q+1) - \mu - \theta_1 \hat{\varepsilon}(q) - \theta_2 \hat{\varepsilon}(q-1) - \dots - \theta_q \hat{\varepsilon}(1) \\ \hat{\varepsilon}(q+2) &= y(q+2) - \mu - \theta_1 \hat{\varepsilon}(q+1) - \theta_2 \hat{\varepsilon}(q) - \dots - \theta_q \hat{\varepsilon}(2) \end{aligned} \right\} \quad \dots \text{(式12-5)}$$

(付録20)に、MA(2) 過程の点予測の計算例を載せる。

13. VAR (ベクトル回帰過程: Vector Auto Regression) モデル

VAR 過程は、弱定常性を前提としたARモデルを多変量に拡張した過程である。
予測精度の向上と変数間の動的関係の分析(後述)に威力を発揮する。(表13-1)に、AR(p) 過程との対比で、VAR(p) 過程を整理しておく。

(表13-1) AR(p) 過程と VAR(p) 過程の対比

	AR(p) 過程	VAR(p) 過程
モデル式	$y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \dots$ (式6-1) $\varepsilon(t) \sim \text{W.N.}(\sigma^2)$: ホワイトノイズ 期待値 $E[\varepsilon(t)] = 0$ 自己共分散: $k \neq 0$ のときは $\gamma_k(t) = \text{Cov}[\varepsilon(t), \varepsilon(t-k)] = 0$ $k = 0$ のときは $\gamma_0(t) = \text{Cov}[\varepsilon^2(t)] = V[\varepsilon(t)] = \sigma^2$	$\mathbf{y}(t) = [y_1(t) \ y_2(t) \ \dots \ y_n(t)]^T$: 時点 t における、 n 個の変数(n 変量) $\mathbf{y}(t) = \mathbf{c} + \Phi_1 \mathbf{y}(t-1) + \Phi_2 \mathbf{y}(t-2) + \dots + \Phi_p \mathbf{y}(t-p) + \boldsymbol{\varepsilon}(t) \dots$ (式 $\mathbf{c}: n \times 1$ の定数ベクトル, $\Phi_i: n \times n$ の係数行列 $\boldsymbol{\varepsilon}(t) \sim \text{W.N.}(\boldsymbol{\Sigma})$: 攪乱項ベクトル(ホワイトノイズ) 期待値 $E[\boldsymbol{\varepsilon}(t)] = \mathbf{0}$ 自己共分散 $E[\boldsymbol{\varepsilon}(t) \boldsymbol{\varepsilon}(t-k)^T]: k=0$ のとき $\boldsymbol{\Sigma}$, $k \neq 0$ のとき $\mathbf{0}$
期待値	$E[y(t)] = \mu = c / (1 - \phi_1 - \phi_2 - \dots - \phi_p) \dots$ (式6-2)	$\boldsymbol{\mu} = E[\mathbf{y}(t)] = [E[y_1(t)] \ E[y_2(t)] \ \dots \ E[y_n(t)]]^T$ $\mu_i = (E[\mathbf{y}(t)])_i = E[y_i(t)]$ $\boldsymbol{\mu} = (\mathbf{I}_n - \Phi_1 - \Phi_2 - \dots - \Phi_p)^{-1} \mathbf{c}$, $\mathbf{I}_n$ は n 次の単位行列 $\dots$ (式13-2)
分散	$V[y(t)] = \gamma_0 = \sigma^2 / (1 - \phi_1 \rho_1 - \phi_2 \rho_2 - \dots - \phi_p \rho_p) \dots$ (式6-3)	$V[\mathbf{y}(t)] = [V[y_1(t)] \ V[y_2(t)] \ \dots \ V[y_n(t)]]^T$
p 次の自己共分散	$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \dots$ (式6-4) $\gamma_k = \gamma_{-k}$	$\Gamma(k) = \text{Cov}[\mathbf{y}(t), \mathbf{y}(t-k)]$ $\Gamma(k)_{ij} = (\text{Cov}[\mathbf{y}(t), \mathbf{y}(t-k)])_{ij} = \text{Cov}[y_i(t) y_j(t-k)] = \sigma_{ij}(k)$ $\dots$ k 次の自己共分散行列の i, j 成分は、 $y_i(t)$ と $y_j(t-k)$ の自己共分散 対角成分は、 k 次の分散 $\Gamma(k) = \Phi_1 \Gamma(k-1) + \Phi_2 \Gamma(k-2) + \dots + \Phi_p \Gamma(k-p) \dots$ (式13-3) $\Gamma(k) = \Gamma^T(-k) \dots$ (式13-4)
p 次の自己相関	$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \dots$ (式6-5) $\rho_k = \rho_{-k}$	$\rho(k) = \text{Corr}[\mathbf{y}(t), \mathbf{y}(t-k)]$, Corr は相関係数を求める関数 $\rho(k)_{ij} = (\text{Corr}[\mathbf{y}(t), \mathbf{y}(t-k)])_{ij} = \text{Corr}[y_i(t) y_j(t-k)]$ $\dots$ k 次の自己相関行列の i, j 成分は、 $y_i(t)$ と $y_j(t-k)$ の相関係数 $\rho(k) = \rho^T(-k) \dots$ (式13-5)
自己共分散と自己相関の関係	$\rho_k = \gamma_k / \gamma_0 \dots$ (式6-5)	$\rho(k) = \mathbf{D}^{-1/2} \Gamma(k) \mathbf{D}^{-1/2} \dots$ (式13-6) ただし、 $\mathbf{D}$ は、 σ_i^2 を対角成分とする $n \times n$ の対角行列
定常性条件	特性方程式 $1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0 \dots$ (式6-7) (式6-7)の z^{-1} の根 λ が全て $ \lambda > 1$ のとき、AR(p) 過程は定常	特性方程式 行列式 $ \mathbf{I}_n - \Phi_1 z^{-1} - \Phi_2 z^{-2} - \dots - \Phi_p z^{-p} = 0 \dots$ (式13-7) (式13-7)の z^{-1} の根 λ が全て $ \lambda > 1$ のとき、VAR(p) 過程は定常
反転可能性	AR(p) 過程が定常であれば、AR(p) 過程 $\rightarrow$ MA(∞) 過程 の反転可能	VAR(p) 過程が定常であれば、VAR(p) 過程 $\rightarrow$ VMA(∞) 過程 の反転可能

(式13-6)の導出方法は、(付録F)に示す。

(式13-4)は、以下のようにして求めるとことができる。

$\Gamma(k)_{ij} = \text{Cov}[y_i(t), y_j(t-k)] \rightarrow \Gamma(-k)_{ij} = \text{Cov}[y_i(t), y_j(t+k)]$ 従って、定常過程であれば、 $\Gamma^T(-k)_{ij} = \text{Cov}[y_i(t+k), y_j(t)] = \text{Cov}[y_i(t), y_j(t-k)] = \Gamma(k)_{ij}$
また、(式13-6)に示す通り、 $\rho(k)$ の時間差 k は $\Gamma(k)$ の時間差に依存しているので、(式13-4)と同じように(式13-5)が成立する。

14. VAR モデル：グレンジャー因果性(Granger Causality)

グレンジャー因果性は「因果性」と言う名称が付いているが、通常の原因 → 結果という因果律を示すのではなくて、ある変数の過去の時系列データが他の変数の将来の予測に役立つときグレンジャー因果性があると言われる。

例えば、ある店舗の将来の売上高を予測する場合、ある店舗の過去の売上高データに加えて他の店舗の売上高データを加味した場合に予測精度が上がる場合、他の店舗の売り上げデータから、ある店舗の売り上げデータに対してグレンジャー因果性がある、というように表現される。

ここでは、グレンジャー因果性が存在するか否かを検定するための検定統計量を導出する。

(1) まず、見通しをよくするために、2 変量 ($n = 2$) の VAR(p) で考える。データ数 = T とする。(式13-1)から、

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} \phi_{11}(1) & \phi_{12}(1) \\ \phi_{21}(1) & \phi_{22}(1) \end{bmatrix} \begin{bmatrix} y_1(t-1) \\ y_2(t-1) \end{bmatrix} + \begin{bmatrix} \phi_{11}(2) & \phi_{12}(2) \\ \phi_{21}(2) & \phi_{22}(2) \end{bmatrix} \begin{bmatrix} y_1(t-2) \\ y_2(t-2) \end{bmatrix} + \cdots + \begin{bmatrix} \phi_{11}(p) & \phi_{12}(p) \\ \phi_{21}(p) & \phi_{22}(p) \end{bmatrix} \begin{bmatrix} y_1(t-p) \\ y_2(t-p) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{bmatrix} \quad \cdots \text{(式14-1)}$$

(2) $y_1(t)$ に着目する。

$$\begin{aligned} y_1(t) &= c_1 + \phi_{11}(1) y_1(t-1) + \phi_{12}(1) y_2(t-1) \\ &\quad + \phi_{11}(2) y_1(t-2) + \phi_{12}(2) y_2(t-2) \\ &\quad \cdots \\ &\quad + \phi_{11}(p) y_1(t-p) + \phi_{12}(p) y_2(t-p) + \varepsilon_1(t) \quad \cdots \text{(式14-2)} \end{aligned}$$

(式14-2)のモデルのパラメータ $c_1, \phi_{11}(1), \phi_{12}(1), \cdots, \phi_{12}(p)$ の $2p+1$ 個のパラメータを最小二乗法で推定し、その残差平方和を SSR_1 とする。

(3) 一方、(式14-1)において、 $y_2(t)$ が $y_1(t)$ に影響しないモデル(式14-3)を考える。つまり、(式14-2)で、 $\phi_{12}(1) = \phi_{12}(2) = \cdots = \phi_{12}(p) = 0$ とする。

$$\begin{aligned} y_1(t) &= c_1 + \phi_{11}(1) y_1(t-1) + \textcolor{red}{0} \times y_2(t-1) & \text{よって、} & y_1(t) = c_1 + \phi_{11}(1) y_1(t-1) \\ &\quad + \phi_{11}(2) y_1(t-2) + \textcolor{red}{0} \times y_2(t-2) & & + \phi_{11}(2) y_1(t-2) \\ &\quad \cdots & & \cdots \\ &\quad + \phi_{11}(p) y_1(t-p) + \textcolor{red}{0} \times y_2(t-p) + \varepsilon_1(t) & & + \phi_{11}(p) y_1(t-p) + \varepsilon_1(t) \quad \cdots \text{(式14-3)} \end{aligned}$$

(式14-3)のモデルのパラメータ $c_1, \phi_{11}(1), \phi_{11}(2), \cdots, \phi_{11}(p)$ の $p+1$ 個のパラメータを最小二乗法で推定し、その残差平方和を SSR_0 とする。

(4) 2 変量 $y_1(t), y_2(t)$ を、以下の k 個の変量 $y_1(t), y_2(t), \cdots, y_k(t)$ に一般化して考える。なお、 σ^2 は $\varepsilon_i(t)$ の分散で時点に依らず一定値をとる。

帰無仮説 H_0 は、 $\phi_{1i}(1) = \phi_{1i}(2) = \cdots = \phi_{1i}(r) = 0$ (パラメータ数 = r) であるとする。

(2)の結果を参考にして、 SSR_0 / σ^2 は、自由度 $T - (k p + 1 - r)$ のカイニ乗分布に従うことが分かる(別途資料「統計的推定の手順」を参照)。

(1)の結果を参考にして、 SSR_1 / σ^2 は、自由度 $T - (k p + 1)$ のカイニ乗分布に従うことが分かる(別途資料「統計的推定の手順」を参照)。

また、カイニ乗分布の再生性(別途資料「統計的推定の手順」を参照)より、

$(SSR_0 / \sigma^2) - (SSR_1 / \sigma^2)$ は、自由度 $(T - (k p + 1 - r)) - (T - (k p + 1)) = T - k p - 1 + r - T + k p + 1 = r$ のカイニ乗分布に従う。

(5) 従って、

$$U = \{[(SSR_0 / \sigma^2) - (SSR_1 / \sigma^2)] / r\} / \{(SSR_1 / \sigma^2) / (T - (k p + 1))\} = \{(SSR_0 - SSR_1) / r\} / \{SSR_1 / (T - k p - 1)\}$$

は、自由度 $(r, T - k p - 1)$ の F 分布に従うことが分かる(付録R)。

ここで、データ数 T が十分大きいとき、 $T - k p - 1$ の値も十分大きな値をとるので、

下記の(式14-4)に示す統計量 $r U$ は、自由度 r のカイニ乗分布 $\chi^2(r)$ に従う(付録L)。この統計量 $r U$ がグレンジャー因果性の検定に使用される。

$$r U = r \{[(SSR_0 - SSR_1) / r] / \{SSR_1 / (T - k p - 1)\}\} = (SSR_0 - SSR_1) / \{SSR_1 / (T - k p - 1)\} \quad \cdots \text{(式14-4)}$$

(付録21)にグレンジャー因果性の検定の計算例を載せる。

15. VAR モデル：インパルス応答関数(Impulse Response Function)

グレンジャー因果性は変数間の関係性を示すが、その**影響の大きさ**と**か影響を及ぼす期間**とは把握できない。これを補う方法の一つがインパルス応答関数である。

(1) 非直交化インパルス応答関数 (Generalized Impulse Response Function)

3 変数の VAR(1) のモデルを例にして考える。(式14-1)の定数項を省略して、また攪乱項を第1項において展開していくと、

$$\begin{aligned} y(t+k) &= \varepsilon(t+k) + \Phi_1 y(t+k-1) = \varepsilon(t+k) + \Phi_1 (\varepsilon(t+k-1) + \Phi_1 y(t+k-2)) = \varepsilon(t+k) + \Phi_1 \varepsilon(t+k-1) + (\Phi_1)^2 y(t+k-2) = \dots \\ &= \varepsilon(t+k) + \Phi_1 \varepsilon(t+k-1) + (\Phi_1)^2 \varepsilon(t+k-2) + \dots + (\Phi_1)^{k-1} \varepsilon(t+1) + (\Phi_1)^k y(t) \quad \dots \quad (\text{式15-1}) \end{aligned}$$

(式15-1)から、例えば、 $y_2(t)$ の変化(インパルス)を与えたとき、 k 時点後の $y_1(t+k)$ に与える影響の大きさを $IRF_{12}(k)$ で表せば、下記の式で計算できる。

$$IRF_{12}(k) = [1 \ 0 \ 0] (\Phi_1)^k [0 \ 1 \ 0]^T \quad \dots \quad (\text{式15-2})$$

この式では、ベクトル $[0 \ 1 \ 0]^T$ が $y_2(t)$ の大きさ1の変化を表し、 $(\Phi_1)^k$ が受けた影響の大きさを表し、ベクトル $[1 \ 0 \ 0]$ が k 時点後の $y_1(t+k)$ を抽出していると見なせる。さらに、(式15-1)の展開を続けると、

$$\begin{aligned} y(t+k) &= \varepsilon(t+k) + \Phi_1 \varepsilon(t+k-1) + (\Phi_1)^2 y(t+k-2) + \dots + (\Phi_1)^{k-1} \varepsilon(t+1) + (\Phi_1)^k \varepsilon(t) + (\Phi_1)^{k+1} y(t-1) = \dots \\ &= \varepsilon(t+k) + \Phi_1 \varepsilon(t+k-1) + (\Phi_1)^2 y(t+k-2) + \dots + (\Phi_1)^{k-1} \varepsilon(t+1) + (\Phi_1)^k \varepsilon(t) + (\Phi_1)^{k+1} \varepsilon(t-1) + \dots \quad \dots \quad (\text{式15-3}) \end{aligned}$$

(式15-3)では、影響の大きさを表す $(\Phi_1)^k$ は $\varepsilon(t)$ の係数であることが分かる。つまり、(式15-2)の $y(t)$ に変化を与えて影響の大きさを把握する代わりに、(式15-3)の $\varepsilon(t)$ に変化を与えて影響の大きさを把握できることを示している。攪乱項 $\varepsilon(t)$ にはモデルに組み込めなかった因子が集約されていると考えれば、インパルス応答はその因子に変化を与えて影響を把握することだと考えてもいい。結局、 $y_1(t)$ に大きさ1の変化を与えたときに k 時点後に $y_1(t+k)$ に与える影響の大きさは、(式15-4)で把握することができる。この式を、非直交化インパルス応答関数という。計算例を(付録22)に示す。

$$IRF_{ij}(k) = \partial y_i(t+k) / \partial \varepsilon_j(t) \quad \dots \quad (\text{式15-4})$$

(2) 直交化インパルス応答関数 (Orthogonal Impulse Response Function)

非直交化インパルス応答関数の問題点は、実際には攪乱項 $\varepsilon(t)$ の要素間に相関関係があるにも関わらずないものとして(式15-4)を使うことにある。そこで、 $\varepsilon(t)$ を無相関な攪乱項になるように分解し、それに変化を与えてその影響の大きさを求める方法が提案されている。これが直交化インパルス応答関数である。計算例を(付録23)に示す。

1) 分解方法の一つは、 $\varepsilon(t)$ の分散行列 $\Sigma = V [\varepsilon(t)]$ を**コレスキー分解**して、下三角行列 P の積に分解する方法である(付録1-1、付録1-2)。

$$\Sigma = P P^T \quad \dots \quad (\text{式15-5})$$

この行列 P (コレスキー因子)を使って、新しい攪乱項を $v(t) = P^{-1} \varepsilon(t) \quad \dots \quad (\text{式15-6})$ で定義する。

$$\begin{aligned} \text{ここで、} V[v(t)] &= V[P^{-1} \varepsilon(t)] = E[(P^{-1} \varepsilon(t) - E[P^{-1} \varepsilon(t)])(P^{-1} \varepsilon(t) - E[P^{-1} \varepsilon(t)])^T] = E[(P^{-1} \varepsilon(t) - P^{-1} E[\varepsilon(t)])(\varepsilon(t)^T - (E[\varepsilon(t)])^T)(P^{-1})^T] \\ &= E[P^{-1} (\varepsilon(t) - E[\varepsilon(t)])(\varepsilon(t)^T - (E[\varepsilon(t)])^T)(P^{-1})^T] = P^{-1} E[(\varepsilon(t) - E[\varepsilon(t)])(\varepsilon(t)^T - (E[\varepsilon(t)])^T)(P^{-1})^T] \\ &= P^{-1} E[(\varepsilon(t) - E[\varepsilon(t)])(\varepsilon(t) - E[\varepsilon(t)])^T](P^{-1})^T = P^{-1} V[\varepsilon(t)](P^{-1})^T = P^{-1} \Sigma (P^{-1})^T = P^{-1} (P P^T) (P^{-1})^T = (P^{-1} P) (P^{-1} P)^T = I \quad (\text{単位行列}) \end{aligned}$$

つまり、 $V[v(t)]$ は対角行列(単位行列)であるので、 $v(t)$ は互いに無相関である。

従って、 $v(t)$ を使って、(式15-7)に示すように、直交化インパルス応答関数を求めることができる。 $IRF_{ij}(k) = \partial y_i(t+k) / \partial v_j(t) \quad \dots \quad (\text{式15-7})$

2) また、 $\varepsilon(t)$ の分散行列 $\Sigma = V[\varepsilon(t)]$ を**修正コレスキー分解**して、下三角行列 A と対角行列 D の積に分解する方法もある(付録1-1、付録1-2)。

$$\Sigma = A D A^T \quad \dots \quad (\text{式15-8})$$

この行列 A を使って、新しい錯乱項を $u(t) = A^{-1} \varepsilon(t) \quad \dots \quad (\text{式15-9})$ で定義する。

$$\text{上記と同様に、} V[u(t)] = V[A^{-1} \varepsilon(t)] = E[(A^{-1} \varepsilon(t) - E[A^{-1} \varepsilon(t)])(A^{-1} \varepsilon(t) - E[A^{-1} \varepsilon(t)])^T] = A^{-1} V[\varepsilon(t)](A^{-1})^T = A^{-1} \Sigma (A^{-1})^T = A^{-1} \Sigma (A^T)^{-1} = D \quad \dots \quad (\text{式15-10})$$

つまり、 $V[u(t)]$ は対角行列 D と等しく、 $u(t)$ は互いに無相関である。

従って、この $u(t)$ を使って、(式15-11)に示すように、直交化インパルス応答関数を求めることができる。 $IRF_{ij}(k) = \partial y_i(t+k) / \partial u_j(t) \quad \dots \quad (\text{式15-11})$

なお、(15-5)と(15-8)から、 $P = A D^{1/2}$ が求まり、(式15-6)と(式15-9)から、 $P v(t) = A u(t)$ が求まる。従って、

$$v(t) = P^{-1} A u(t) = (A D^{1/2})^{-1} A u(t) = (D^{1/2})^{-1} A^{-1} A u(t) = (D^{1/2})^{-1} u(t) = D^{-1/2} u(t) \quad \dots \quad (\text{式15-12})$$

行列 D は対角要素が $u(t)$ の分散である対角行列なので、(式15-12)より、 $v(t)$ は $u(t)$ をその標準偏差(=分散^{1/2})で割ったものになっている。言い換えれば、 $v_j(t)$ に1単位の変化を与えて計算した直交化インパルス応答は、 $u_j(t)$ に1標準偏差の変化を与えて計算したインパルス応答に等しい。変化の与え方としては1標準偏差が使用されることが多く、直交化インパルス応答関数では、 $u_j(t)$ つまり、修正コレスキー分解がよく使用されている。

また、(式15-9)の計算例である(付式23-6)から伺うことができるが、 $y_k(t)$ に対応している $\varepsilon_k(t)$ は $u_m(t)$ の影響を受けないことが分かる(ただし、 $k < m$ で)。これは、**ベクトル $y(t)$ の変数の並びは他から影響を受けづらい(外生性の高い)順に並んでいることを示している。インパルス応答関数を計算するにあたってはこのことを配慮する必要がある。**

(付式23-6)では、 $u_1(t)$ は $\varepsilon_1(t)$ 、 $\varepsilon_2(t)$ 、 $\varepsilon_3(t)$ 全てに影響、 $u_2(t)$ は $\varepsilon_1(t)$ に影響せず $\varepsilon_2(t)$ 、 $\varepsilon_3(t)$ に影響、 $u_3(t)$ は $\varepsilon_1(t)$ 、 $\varepsilon_2(t)$ に影響せず $\varepsilon_3(t)$ に影響していることが分かる。

16. VAR モデル：分散分解(Variance Decomposition)

時系列の予測データの平均二乗誤差(MSE : Mean Squared Error)は各変数の攪乱項に起因するとして、各変数ごとの寄与割合を定量化する。これを、分散分解という。結果として、ある変数の(予測できない)変動を説明するためにどの変数が重要であるかを明らかにすることができる。

ここでは、VAR(1) 過程を例に挙げて、分散分解の仕組みを確認する。VAR(1) 過程の式は(式13-1)より、また、(式15-9)より $\varepsilon(t) = A u(t)$ であるので、

$$y(t) = c + \Phi y(t-1) + \varepsilon(t) = c + \Phi y(t-1) + A u(t) \quad \dots \text{(式16-1)}$$

(式16-1)を使って、1 時点先の $y(t+1)$ の値は、 $y(t+1) = c + \Phi y(t) + A u(t+1)$

よって、k 時点先の $y(t+k)$ の値は、

$$\begin{aligned} y(t+k) &= c + \Phi y(t+k-1) + A u(t+k) = c + \Phi (c + \Phi y(t+k-2) + A u(t+k-1)) + A u(t+k) \\ &= c + \Phi c + \Phi^2 y(t+k-2) + \Phi A u(t+k-1) + A u(t+k) = c + \Phi c + \Phi^2 (c + \Phi y(t+k-3) + A u(t+k-2)) + \Phi A u(t+k-1) + A u(t+k) \\ &= c + \Phi c + \Phi^2 c + \Phi^3 y(t+k-3) + \Phi^2 A u(t+k-2) + \Phi A u(t+k-1) + A u(t+k) \\ &\dots \\ &= c + \Phi c + \Phi^2 c + \Phi^3 y(t+k-3) + \dots + \Phi^k y(t) + \Phi^{k-1} A u(t+1) + \Phi^{k-2} A u(t+2) + \dots + \Phi^2 A u(t+k-2) + \Phi A u(t+k-1) + A u(t+k) \end{aligned}$$

t+k 時点の予測は攪乱項を無視して、 $\hat{y}(t+k) = c + \Phi c + \Phi^2 c + \Phi^3 y(t+k-3) + \dots + \Phi^k y(t) \quad \dots \text{(式16-2)}$

t+k 時点の予測誤差は、 $\hat{e}(t+k) = \Phi^{k-1} A u(t+1) + \Phi^{k-2} A u(t+2) + \dots + \Phi^2 A u(t+k-2) + \Phi A u(t+k-1) + A u(t+k) \quad \dots \text{(式16-3)}$

(式15-9)から $E[u(t)] = E[A^{-1} \varepsilon(t)]$ であり、 $E[\varepsilon(t)] = 0$ より、 $E[u(t)] = 0$ となる。

また、定常過程であるから、 $E[u(t+1)] = E[u(t+2)] = \dots = E[u(t+k)] = 0$ であり、(付録B)を参考にして、

$$\begin{aligned} E[\hat{e}(t+k)] &= E[\Phi^{k-1} A u(t+1) + \Phi^{k-2} A u(t+2) + \dots + \Phi^2 A u(t+k-2) + \Phi A u(t+k-1) + A u(t+k)] \\ &= \Phi^{k-1} A E[u(t+1)] + \Phi^{k-2} A E[u(t+2)] + \dots + \Phi^2 A E[u(t+k-2)] + \Phi A E[u(t+k-1)] + A E[u(t+k)] = 0 \quad \dots \text{(式16-4)} \end{aligned}$$

また、 $V[\hat{e}(t+k)]$ は2通りの表現方法があり、(式16-4)を用いて、1つは、

$$V[\hat{e}(t+k)] = E[(\hat{e}(t+k) - E[\hat{e}(t+k)]) (\hat{e}(t+k) - E[\hat{e}(t+k)])^T] = E[(\hat{e}(t+k) - 0) (\hat{e}(t+k) - 0)^T] = E[\hat{e}(t+k) (\hat{e}(t+k))^T] \quad \dots \text{(式16-5)}$$

もう1つは、(式16-3)の右辺は互いに独立な $u(t+1)$, $\dots$, $u(t+k-1)$, $u(t+k)$ の線形結合であり、独立な変数の和の分散は変数ごとの分散の和に等しいので(付録B)、

$$\begin{aligned} V[\hat{e}(t+k)] &= V[\Phi^{k-1} A u(t+1) + \Phi^{k-2} A u(t+2) + \dots + \Phi^2 A u(t+k-2) + \Phi A u(t+k-1) + A u(t+k)] \\ &= V[\Phi^{k-1} A u(t+1)] + V[\Phi^{k-2} A u(t+2)] + \dots + V[\Phi^2 A u(t+k-2)] + V[\Phi A u(t+k-1)] + V[A u(t+k)] \quad \dots \text{(式16-6)} \end{aligned}$$

(1) t+k 時点の予測 $\hat{y}(t+k)$ の平均二乗誤差(MSE)

$E[u(t)] = 0$ であるので、 $V[u(t)] = E[(u(t) - E[u(t)]) (u(t) - E[u(t)])^T] = E[(u(t) - 0) (u(t) - 0)^T] = E[u(t) u(t)^T]$

(式16-5)と(式16-6)より、また、定常過程であるから $V[u(t)] = V[u(t+1)]$, $\dots$, $V[u(t+k)]$ であるので、(付録B)を参考にして、

$$\begin{aligned} \text{MSE}(\hat{y}(t+k)) &= E[(\hat{e}(t+k)) (\hat{e}(t+k))^T] \\ &= V[\hat{e}(t+k)] \\ &= V[\Phi^{k-1} A u(t+1)] + \dots + V[\Phi^2 A u(t+k-2)] + V[\Phi A u(t+k-1)] + V[A u(t+k)] \\ &= \Phi^{k-1} A V[u(t+1)] (\Phi^{k-1} A)^T + \dots + \Phi^2 A V[u(t+k-2)] (\Phi^2 A)^T + \Phi A V[u(t+k-1)] (\Phi A)^T + A V[u(t+k)] A^T \\ &= \Phi^{k-1} A V[u(t)] (\Phi^{k-1} A)^T + \dots + \Phi^2 A V[u(t)] (\Phi^2 A)^T + \Phi A V[u(t)] (\Phi A)^T + A V[u(t)] A^T \quad \dots \text{(式16-7)} \end{aligned}$$

例えば、 $y(t) = [y_1(t) \ y_2(t) \ y_3(t)]^T$ であれば、 $y_1(t)$ に関して、 $V[u(t)] = D$ (対角行列)であることに注意すれば、k=3 のとき、

(注) 下記の式で、例えば、 $A_{1,2}$ は行列 A の (1,2) 要素を示す。

$$\text{MSE}(\hat{y}(t+3)) = \Phi^2 A V[u(t)] (\Phi^2 A)^T + \Phi A V[u(t)] (\Phi A)^T + A V[u(t)] A^T \quad \dots \text{(式16-8)} \quad \text{であるから、}$$

$$\text{MSE}(\hat{y}_1(t+3)) = (((\Phi^2 A)_{1,1})^2 + ((\Phi A)_{1,1})^2 + (A_{1,1})^2) V[u(t)]_{1,1} + (((\Phi^2 A)_{1,2})^2 + ((\Phi A)_{1,2})^2 + (A_{1,2})^2) V[u(t)]_{2,2} + (((\Phi^2 A)_{1,3})^2 + ((\Phi A)_{1,3})^2 + (A_{1,3})^2) V[u(t)]_{3,3} \quad \dots \text{(式16-9)}$$

$$\text{MSE}(\hat{y}_2(t+3)) = (((\Phi^2 A)_{2,1})^2 + ((\Phi A)_{2,1})^2 + (A_{2,1})^2) V[u(t)]_{1,1} + (((\Phi^2 A)_{2,2})^2 + ((\Phi A)_{2,2})^2 + (A_{2,2})^2) V[u(t)]_{2,2} + (((\Phi^2 A)_{2,3})^2 + ((\Phi A)_{2,3})^2 + (A_{2,3})^2) V[u(t)]_{3,3} \quad \dots \text{(式16-10)}$$

$$\text{MSE}(\hat{y}_3(t+3)) = (((\Phi^2 A)_{3,1})^2 + ((\Phi A)_{3,1})^2 + (A_{3,1})^2) V[u(t)]_{1,1} + (((\Phi^2 A)_{3,2})^2 + ((\Phi A)_{3,2})^2 + (A_{3,2})^2) V[u(t)]_{2,2} + (((\Phi^2 A)_{3,3})^2 + ((\Phi A)_{3,3})^2 + (A_{3,3})^2) V[u(t)]_{3,3} \quad \dots \text{(式16-11)}$$

(2) 相対的分散寄与率(RCV : Relative Variance Contribution)

t+k 時点の予測 $\hat{y}_i(t+k)$ の MSE に対して、 $y_j(t+k)$ の直交化攪乱項 $u_j(t+1)$, $u_j(t+2)$, $\dots$, $u_j(t+k)$ が寄与する割合は、(式16-12)で表される。

$$\text{RCV}_{ij}(k) = \{ \text{MSE}(\hat{y}_i(t+k)) \text{ において } y_j(t+k) \text{ が寄与する部分} \} / \text{MSE}(\hat{y}_i(t+k)) \quad \dots \text{(式16-12)}$$

(付録24)に、分散分解の計算例を載せる。

17. 単位根過程 (Unit Root process) の特徴

物価、GDP (Gross Domestic Product, 国内総生産)、株価、為替レートなどの経済に関わる時系列データは、平均に回帰するなどの定常過程の要件を満たさない非定常過程であることが多い。このような非定常過程のモデル化に単位根過程モデルは使用される。単位根の名前は過程の特性方程式(式6-7)の根が1になることに由来している。

(1) 単位根過程 (Unit Root Process) の定義

原系列 $y(t)$ が非定常過程であり、その差分系列 $\Delta y(t) = y(t) - y(t-1)$ が定常過程であるとき、その過程は単位根過程であると言われる。

単位根過程は、定義に従い、差分定常過程 (Difference Stationary Process) とも言われる。下記に示すランダムウォークは単位根過程の一例である。

(2) 和分過程 (Integrated Process) の定義

原系列 $y(t)$ の $d-1$ 階差分系列 $y(t) - y(t-d+1)$ が非定常過程であり、その d 階差分系列 $y(t) - y(t-d)$ が定常過程であるとき、この過程を d 次和分過程、もしくは、 $I(d)$ 過程と言われる。また、 $I(0)$ 過程は定常過程であると定義される。

(3) ARIMA 過程 (Auto Regressive Integrated Moving Average Process)

原系列 $y(t)$ の d 階差分系列 $y(t) - y(t-d)$ が定常かつ反転可能な $ARMA(p, q)$ 過程に従うとき、この過程は $ARIMA(p, d, q)$ 過程、または、次数 (p, d, q) の自己回帰和分移動平均過程であると言われる。

(4) ランダムウォーク (Random Walk) ... (表17-1)のなかの(図17-1)に例を示す。

$y(t) = \delta + y(t-1) + \varepsilon(t)$... (式17-1) 但し、 $y(0) = 0$, $\varepsilon \sim \text{i.i.d.}(0, \sigma^2)$ で定義される。

・トレンド $y(t) = \delta + y(t-1) + \varepsilon(t) = \delta + (\delta + y(t-2) + \varepsilon(t-1)) + \varepsilon(t) = 2\delta + y(t-2) + \varepsilon(t-1) + \varepsilon(t) = \dots = \delta t + \varepsilon(1) + \varepsilon(2) + \dots + \varepsilon(t)$... (式17-2)

ここで、 δt は時間の経過に伴い変化する。つまり、トレンドを表す。 δ はドリフト率と呼ばれる。

・予測 $\delta = 0$ では、1期先は、 $y(t+1) = y(t) + \varepsilon(t+1)$ 、2期先は、 $y(t+2) = y(t+1) + \varepsilon(t+2) = y(t) + \varepsilon(t+1) + \varepsilon(t+2)$, ...

h 期先は、 $y(t+h) = y(t) + \varepsilon(t+1) + \varepsilon(t+2) + \dots + \varepsilon(t+h)$ よって、 h 期先の最適予測は、 $\hat{y}(t+h) = y(t)$... (式17-3)

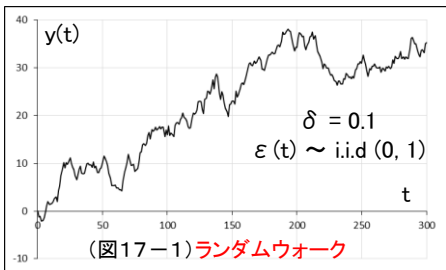
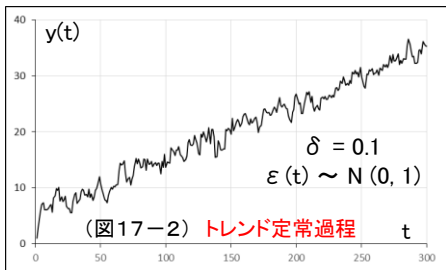
・インパルス応答関数 (式15-4)より、 $\partial y(t+h) / \partial \varepsilon(t) = \partial (y(t) + \varepsilon(t+1) + \dots + \varepsilon(t+h)) / \partial \varepsilon(t)$

$= \partial ((\delta + y(t-1) + \varepsilon(t)) + \varepsilon(t+1) + \dots + \varepsilon(t+h)) / \partial \varepsilon(t) = 1$... (式17-4)

(5) トレンド定常過程 ... (表17-1)のなかの(図17-2)に例を示す。

$y(t) = \delta t + x(t)$ (但し、 $x(t)$ は定常過程) ... (式17-5) で定義される。ランダムウォークと同じく、トレンド項 δt を持つが、定常過程である。

(表17-1) 単位根過程の性質(トレンド定常過程との比較)

	時系列グラフの例	トレンドとの差の分散	h 期先の最適予測	インパルス応答関数
単位根過程 (ランダムウォーク)	 (図17-1) ランダムウォーク	トレンドライン δt からの $y(t)$ の乖離の程度を δt と $y(t)$ の差の分散で表せば、(式17-2)より、 $V[y(t) - \delta t]$ $= V[\varepsilon(1) + \varepsilon(2) + \dots + \varepsilon(t)]$ $= V[\varepsilon(1)] + \dots + V[\varepsilon(t)]$ $= \sigma^2 + \sigma^2 + \dots + \sigma^2$ $= \sigma^2 t$ 時間とともに $y(t)$ のトレンドラインからの乖離が大きくなる(不確実性が増大する)。	(式17-3)より、 $\hat{y}(t+h) = y(t)$ であるので、 期間 h 期先の最適予測 $\hat{y}(t+h)$ は $y(t)$ にしか依存せず、 期間 h が長くなっても $y(t)$ の影響は消えない。	(式17-4)より、 $\partial y(t+h) / \partial \varepsilon(t) = 1$ であるから、インパルスの影響は恒久的に残る。
トレンド 定常過程	 (図17-2) トレンド定常過程	例えば、トレンド定常過程では、 トレンドライン δt からの $y(t)$ の乖離の程度を δt と $y(t)$ の差の分散で表せば、 (式17-5)より、 $V[y(t) - \delta t] = V[x(t)]$ $x(t)$ は定常過程であるので、時間が経っても、$V[x(t)]$ の値は有限である。 つまり、$y(t)$ のトレンドライン δt からの乖離は時間が経っても一定の範囲内に収まる。	例えば、$AR(p)$ 過程では、 (式11-10)より、 $\hat{y}(t+h)$ $= (1 - \phi_1^h) c / (1 - \phi_1) + \phi_1^h y(t)$ h 期先の最適予測 $\hat{y}(t+h)$ は $y(t)$ にしか依存しないが、 定常過程では $\phi_1 < 1$ であり、期間 h が長くなるにつれ、$y(t)$ 項の影響は減衰していく。	例えば、$AR(p)$ 過程では、 (付録23)の計算例で示すように、インパルスの影響は時間が経つにつれてなくなっていく。

18. 単位根過程の検定(DF検定とADF検定)

時系列分析では、単位根過程は非定常過程であるが実務上重要な分析対象になる。そして、時系列データが単位根過程であるのか定常過程であるのかを判断することは、時系列モデルを作成する上では極めて重要である。AR(1) 過程を対象とした DF 検定(Dickey-Fuller Test) と AR(p) 過程を対象とした ADF 検定(Augmented Dickey-Fuller Test, 拡張型 DF 検定)は、単位根過程の検定に使用される。ここでは、DF 検定と ADF 検定の概要をまとめる。

(1) DF 検定(ディッキー・フラー検定)

定数項(c)とトレンド項(δt)を持つ、AR(1) 過程 $y(t) = c + \delta t + \phi_1 y(t-1) + \varepsilon(t)$... (式18-1)

AR(1) 過程の特性方程式は(式7-8)より、 $1 - \phi_1 z^{-1} = 0$ であるが、この特性方程式の根が $z^{-1} = 1$ (単位根)となるのは、 $\phi_1 = 1$ のときである。これが、AR(1) 過程が単位根過程となる条件である。従って、AR(1) 過程の単位根検定は、(式18-1)で帰無仮説 H_0 を $\phi_1 = 1$ (単位根過程)、対立仮説 H_1 を $\phi_1 < 1$ (定常過程)であると置く検定になる。ここで、単位根過程の差分系列は定常過程であること(定義)を利用して、(式18-1)の AR(1) 過程を差分系列を使って表すことを考える。

まず、 $\Delta y(t) = y(t) - y(t-1)$, $\rho = \phi_1 - 1$ とおく。

$$\Delta y(t) = y(t) - y(t-1) = (c + \delta t + \phi_1 y(t-1) + \varepsilon(t)) - y(t-1) = c + \delta t + (\phi_1 - 1)y(t-1) + \varepsilon(t) = c + \delta t + \rho y(t-1) + \varepsilon(t)$$

よって、以下の差分系列が求まる。 $\Delta y(t) = c + \delta t + \rho y(t-1) + \varepsilon(t)$... (式18-2)

(式18-2)を使った単位根検定は、帰無仮説 H_0 を $\rho = \phi_1 - 1 = 1 - 1 = 0$ (単位根過程)、対立仮説 H_1 を $\rho = \phi_1 - 1 < 0$ (定常過程)であると置く検定になる。

つまり、(式18-2)を、目的変数 $\Delta y(t)$, 説明変数 $y(t-1)$, 偏回帰係数 ρ の回帰式であるとし、最小二乗法で推定した偏回帰係数の推定値 $\hat{\rho}$ を使用した検定統計量は、(式18-3)で表すことができる(別途資料「重回帰分析の概要」を参照)。ここで、SE は標準誤差を表す(別途資料「重回帰分析の概要」を参照)。

$$T = \hat{\rho} / SE(\hat{\rho}) \quad \dots \text{(式18-3)}$$

ただ、この検定統計量 T は、回帰分析のときのように t 分布に従わない。つまり、この検定統計量を使って t 検定を行うことはできない。何故ならば、帰無仮説が成立するとき、 $y(t)$ は単位根過程、つまり非定常過程であり、(式18-2)に示すように、攪乱項の和の分散は、 t が大きくなると発散する。

$$V[\varepsilon(1) + \varepsilon(2) + \dots + \varepsilon(t)] = V[\varepsilon(1)] + V[\varepsilon(2)] + \dots + V[\varepsilon(t)] = \sigma^2 + \sigma^2 + \dots + \sigma^2 = \sigma^2 t$$

従って、大数の法則も中心極限定理(別途資料「いろいろな確率分布」を参照)も成立せず、この検定統計量が t 分布に従うことはない。

(式18-3)の検定統計量 T が従う分布は、ディッキー・フラー分布と言われる特殊な分布になる。この検定統計量を使った単位根検定では、ディッキー・フラー分布の、例えば下側 5% 棄却域を示す臨界値との比較を行い、検定統計量 T がこの臨界値より小さな値をとるときは、帰無仮説(単位根過程であるという仮説)を棄却し、対立仮説(定常過程であるという仮説)を採択する。一方、検定統計量 T がこの臨界値より大きな値をとるときは、帰無仮説を受容する、つまり、単位根仮説であると判断する。

DF 検定の計算例は、(付録25)に載せる。

(2) ADF 検定(拡張型ディッキー・フラー検定)

定数項(c)とトレンド項(δt)を持つ、AR(p) 過程 $y(t) = c + \delta t + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t)$... (式18-4)

AR(p) 過程の特性方程式は(式7-8)より、 $1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0$ であり、この特性方程式の根が $z^{-1} = 1$ (単位根)となるのは、 $1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$ のときである。これが、AR(p) 過程が単位根過程となる条件である。ここで、DF 検定の場合と同じく、AR(p) 過程の差分系列を求める。

例えば、AR(3) 過程の場合を考えてみる。

$$y(t) = c + \delta t + \phi_1 y(t-1) + \phi_2 y(t-2) + \phi_3 y(t-3) + \varepsilon(t)$$

$$\begin{aligned} \Delta y(t) &= y(t) - y(t-1) = (c + \delta t + \phi_1 y(t-1) + \phi_2 y(t-2) + \phi_3 y(t-3) + \varepsilon(t)) - y(t-1) \\ &= c + \delta t + (\phi_1 - 1)y(t-1) + \phi_2 y(t-2) + \phi_3 y(t-3) + \varepsilon(t) \\ &= c + \delta t + (\phi_1 - 1)y(t-1) + \phi_2 y(t-2) + \phi_3 y(t-2) - \phi_3 y(t-3) + \phi_3 y(t-3) + \varepsilon(t) \\ &= c + \delta t + (\phi_1 - 1)y(t-1) + (\phi_2 + \phi_3)y(t-2) - \phi_3(y(t-2) - y(t-3)) + \varepsilon(t) \\ &= c + \delta t + (\phi_1 + \phi_2 + \phi_3 - 1)y(t-1) - (\phi_2 + \phi_3)(y(t-1) - y(t-2)) - \phi_3(y(t-2) - y(t-3)) + \varepsilon(t) \end{aligned}$$

$$\text{ここで、}\rho = \phi_1 + \phi_2 + \phi_3 - 1, \zeta_1 = -(\phi_2 + \phi_3), \zeta_2 = -\phi_3 \text{ とおくと、}\Delta y(t) = c + \delta t + \rho y(t-1) + \zeta_1 \Delta y(t-1) + \zeta_2 \Delta y(t-2) + \varepsilon(t)$$

以上の式変形を AR(p) 過程に展開すれば、

$$\Delta y(t) = c + \delta t + \rho y(t-1) + \zeta_1 \Delta y(t-1) + \zeta_2 \Delta y(t-2) + \dots + \zeta_{p-1} \Delta y(t-(p-1)) + \varepsilon(t) \quad \dots \text{(式18-5)}$$

$$\text{但し、}\rho = \phi_1 + \phi_2 + \dots + \phi_p - 1, \zeta_1 = -(\phi_2 + \phi_3 + \dots + \phi_p), \zeta_2 = -(\phi_3 + \phi_4 + \dots + \phi_p), \dots, \zeta_{p-1} = -\phi_p$$

AR(p) 過程は、特性方程式 $1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0$ が成立するとき単位根過程になるので、

AR(p) 過程の単位根検定は、帰無仮説 H_0 を $\rho = \phi_1 + \phi_2 + \dots + \phi_p - 1 = 0$ (単位根過程)、対立仮説 H_1 を $\rho < 0$ (定常過程)であると置く検定になる。

(式18-5)は目的変数 $\Delta y(t)$ の重回帰式で、 ρ は説明変数 $y(t-1)$ の偏回帰係数である。よって、単位根検定の帰無仮説の検定は、 $y(t-1)$ の偏回帰係数 $\rho = 0$ の検定を行うことになる。従って、ADF 検定は、DF 検定と同じく、(式18-3)を検定統計量とした検定になる。

ADF 検定の計算例は、(付録26)に載せる。

19. 見せかけの回帰と共和分

単位根過程の変数の関係は、見せかけの回帰か共和分の関係かの、いずれかである。

1) 見せかけの回帰(Spurious Regression)

単位根過程同志の間で回帰分析すると、本来存在しないはずの回帰関係があるというような分析結果になる。これが「見せかけの回帰」である。

「見せかけの回帰」は単根過程を対象にした現象である。

よって、時系列分析にあたっては、単位根過程であるか否かの検定(単位根検定)を実施することが大切になる。

2) 共和分(Cointegration)

n 個の時系列過程 $\mathbf{y}(t) = [y_1(t) \ y_1(t) \ \cdots \ y_n(t)]^T$ の全てが単位根過程、つまり、 $I(1)$ 過程であって、

$\mathbf{y}(t)$ の線型結合 $\beta^T \mathbf{y}(t) = \beta_1 y_1(t) + \beta_2 y_2(t) + \cdots + \beta_n y_n(t)$ が定常過程になるような係数 $\beta = [\beta_1 \ \beta_2 \ \cdots \ \beta_n]^T$ が存在するとき、

β を**共和分ベクトル**という。また、この時系列には共和分の関係がある、もしくは共和分しているといわれる。

共和分ベクトルは通常、 $\beta_1 = 1$ に設定し、 $\beta = [1 \ \beta_2 / \beta_1 \ \cdots \ \beta_n / \beta_1]^T$ のように基準化する。

3) 見せかけの回帰と共和分の関係

(式19-1)を使って、見せかけの回帰と共和分の関係を確認する。(式19-1)の $x(t)$ と $y(t)$ はいずれも単位根過程であるとする。

$$y(t) = \alpha + \beta x(t) + \varepsilon(t) \rightarrow \varepsilon(t) = y(t) - \alpha - \beta x(t) \quad \cdots \text{(式19-1)}$$

① $\varepsilon(t)$ が単位根過程になる。

このとき、 $x(t)$ と $y(t)$ の関係は見せかけの回帰の関係になる。単位根過程である $x(t)$ と $y(t)$ の線形結合である $\varepsilon(t)$ は通常、単位根過程になる。

② $\varepsilon(t)$ が定常過程になる。

このとき、 $x(t)$ と $y(t)$ の関係は共和分の関係になる。

単位根過程である $x(t)$ と $y(t)$ の線形結合が定常過程になるということで、これは共和分の関係になっていることを示す。

つまり、 $\varepsilon(t)$ の単位根検定を行えば、 $x(t)$ と $y(t)$ の関係が共和分の関係であるのか否かの判断ができる。

この手順で共和分関係の有無を検定する方法が、**Engle-Granger の共和分検定**である。この検定方法は共和分関係の数が既知である場合に限定される。

それに対して、後述の、Johansen (ヨハンセン) の検定方法は、共和分関係の数も推定することができる。

4) 見せかけの回帰を回避する方法

① 目的変数 $y(t)$ と説明変数 $x(t)$ に、時点をずらした目的変数 $y(t-1)$, $\cdots$ もしくは、説明変数 $x(t-1)$, $\cdots$ を含めて回帰分析を行う。

$$\text{(例)} \ y(t) = \alpha + \beta_1 x(t) + \beta_2 x(t-1) + \beta_3 y(t-1) + \varepsilon(t)$$

② そもそも単位根過程でなければ見せかけの回帰は起こらない。そこで、単位根過程に従う変数の差分をとって定常過程にしてから解析を行う。

$$\text{(例)} \ \Delta y(t) = \alpha + \Delta \beta x(t) + \varepsilon(t) \quad \text{但し、} \ \Delta y(t) = y(t) - y(t-1), \ \Delta x(t) = x(t) - x(t-1)$$

(付録27)に、見せかけの回帰と共和分の関係の例を載せる。

20. ベクトル誤差修正モデル(共和分関係の推定と検定)

(1)ベクトル修正誤差モデル

(付録28)に例を示すように、共和分のある単位根 VAR 過程の差分系列を VAR モデルで表現することはできない。代わりに、ベクトル誤差修正モデル(VECM : Vector Error Correction Model)を使用して表す。共和分関係のある VAR 過程を VECM で表すことができることは、**Granger 表現定理**と呼ばれている。

AR (p) 過程の差分系列(式18-5)を求めるときと同じ方法で VAR(2) 過程の差分系列を求める。差分過程モデルである VECM 表現を導出し VAR(p) 過程に拡張する。

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{c} + \Phi_1 \mathbf{y}(t-1) + \Phi_2 \mathbf{y}(t-2) + \Phi_3 \mathbf{y}(t-3) + \boldsymbol{\varepsilon}(t) \quad \boldsymbol{\varepsilon}(t) \sim \text{i.i.d. } N(\mathbf{0}, \Lambda) \\ \Delta \mathbf{y}(t) &= \mathbf{y}(t) - \mathbf{y}(t-1) \\ &= (\mathbf{c} + \Phi_1 \mathbf{y}(t-1) + \Phi_2 \mathbf{y}(t-2) + \Phi_3 \mathbf{y}(t-3) + \boldsymbol{\varepsilon}(t)) - \mathbf{y}(t-1) = \mathbf{c} + (\Phi_1 - \mathbf{I}) \mathbf{y}(t-1) + \Phi_2 \mathbf{y}(t-2) + \Phi_3 \mathbf{y}(t-3) + \boldsymbol{\varepsilon}(t) \\ &= \mathbf{c} + (\Phi_1 - \mathbf{I}) \mathbf{y}(t-1) + \Phi_2 \mathbf{y}(t-2) + \Phi_3 \mathbf{y}(t-2) - \Phi_3 \mathbf{y}(t-2) + \Phi_3 \mathbf{y}(t-3) + \boldsymbol{\varepsilon}(t) = \mathbf{c} + (\Phi_1 - \mathbf{I}) \mathbf{y}(t-1) + (\Phi_2 + \Phi_3) \mathbf{y}(t-2) - \Phi_3 (\mathbf{y}(t-2) - \mathbf{y}(t-3)) + \boldsymbol{\varepsilon}(t) \\ &= \mathbf{c} + (\Phi_1 - \mathbf{I}) \mathbf{y}(t-1) + (\Phi_2 + \Phi_3) \mathbf{y}(t-1) - (\Phi_2 + \Phi_3) \mathbf{y}(t-1) + (\Phi_2 + \Phi_3) \mathbf{y}(t-2) - \Phi_3 (\mathbf{y}(t-2) - \mathbf{y}(t-3)) + \boldsymbol{\varepsilon}(t) \\ &= \mathbf{c} + (\Phi_1 + \Phi_2 + \Phi_3 - \mathbf{I}) \mathbf{y}(t-1) - (\Phi_2 + \Phi_3) (\mathbf{y}(t-1) - \mathbf{y}(t-2)) - \Phi_3 (\mathbf{y}(t-2) - \mathbf{y}(t-3)) + \boldsymbol{\varepsilon}(t) \\ &= \mathbf{c} + (\Phi_1 + \Phi_2 + \Phi_3 - \mathbf{I}) \mathbf{y}(t-1) - (\Phi_2 + \Phi_3) \Delta \mathbf{y}(t-1) - \Phi_3 \Delta \mathbf{y}(t-2) + \boldsymbol{\varepsilon}(t) \end{aligned}$$

ここで、 $\Pi = \Phi_1 + \Phi_2 + \Phi_3 - \mathbf{I}$, $\Gamma_1 = -(\Phi_2 + \Phi_3)$, $\Gamma_2 = -\Phi_3$ とおけば、
 $\Delta \mathbf{y}(t) = \mathbf{c} + \Pi \mathbf{y}(t-1) + \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-2) + \boldsymbol{\varepsilon}(t) = \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-2) + \Pi \mathbf{y}(t-1) + \mathbf{c} + \boldsymbol{\varepsilon}(t)$

この手順を VAR(p) 過程にして、

$$\Delta \mathbf{y}(t) = \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-1) + \dots + \Gamma_{p-1} \Delta \mathbf{y}(t-(p-1)) + \Pi \mathbf{y}(t-1) + \mathbf{c} + \boldsymbol{\varepsilon}(t) \quad \dots \text{ (式20-1)}$$

ただし、 $\Pi = \Phi_1 + \Phi_2 + \dots + \Phi_p - \mathbf{I}$, $\Gamma_i = -(\Phi_{i+1} + \Phi_{i+2} + \dots + \Phi_p)$

(式20-1)で、左辺の $\Delta \mathbf{y}(t)$ は定常過程(1次和分過程、つまり、I (0) 過程)、また右辺の差分の項は全て定常過程でなくてはならない。

よって、 $\Pi = \alpha \beta^T \quad \dots \text{ (式20-2)}$ サイズ $\alpha : n \times h$, $\beta : h \times n$, $\Pi : n \times n$ とおけば、

$\Pi \mathbf{y}(t-1) = \alpha \beta^T \mathbf{y}(t-1)$ は定常過程でなくてはならない。 $\Pi = \alpha \beta^T$ は共和分関係を表している。

$$\Delta \mathbf{y}(t) = \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-1) + \dots + \Gamma_{p-1} \Delta \mathbf{y}(t-(p-1)) + \alpha \beta^T \mathbf{y}(t-1) + \mathbf{c} + \boldsymbol{\varepsilon}(t) \quad \dots \text{ (式20-3)}$$

この $\alpha \beta^T \mathbf{y}(t-1)$ の項は、(式20-3)の式のなかで差分の式を修正する項であると捉えて、**モデルの誤差修正項**(error correction term)と呼ばれている。

β は共和分ベクトルで構成されており、 $\beta^T \mathbf{y}(t-1)$ の項は定常項になる。 α はモデルの乖離を調整する速度を表していると捉える。

(2)共和分関係(β)の推定

(付録29)に、共和分ベクトル β の推定手順を載せる。 β の最尤推定量は、下記の特性方程式から求まる固有値($\lambda_1 \geq \dots \geq \lambda_n \geq 0$)に対応する固有ベクトルになる。

$$|\lambda - (\mathbf{S}_{11})^{-1} \mathbf{S}_{10} (\mathbf{S}_{00})^{-1} \mathbf{S}_{01}| = 0 \quad \dots \text{ (付式29-28)}$$

但し、 $\hat{\mathbf{u}}_0(t) : \Delta \mathbf{y}(t)$ を $\Delta \mathbf{y}(t-1), \dots, \Delta \mathbf{y}(t-(p-1))$ に回帰したときの推定残差ベクトル

$\hat{\mathbf{u}}_{01}(t) : \mathbf{y}(t-1)$ を $\Delta \mathbf{y}(t-1), \dots, \Delta \mathbf{y}(t-(p-1))$ に回帰したときの推定残差ベクトル

これらの $\hat{\mathbf{u}}_0(t)$ と $\hat{\mathbf{u}}_1(t)$ を使って、以下の標本分散共分散行列を定め、特性方程式(付式29-28)に使用する。

$$\mathbf{S}_{00} = (1/T) \sum_{t=1}^T \hat{\mathbf{u}}_0(t) (\hat{\mathbf{u}}_0(t))^T, \mathbf{S}_{01} = (1/T) \sum_{t=1}^T \hat{\mathbf{u}}_0(t) (\hat{\mathbf{u}}_1(t))^T, \mathbf{S}_{10} = (1/T) \sum_{t=1}^T \hat{\mathbf{u}}_1(t) (\hat{\mathbf{u}}_0(t))^T, \mathbf{S}_{11} = (1/T) \sum_{t=1}^T \hat{\mathbf{u}}_1(t) (\hat{\mathbf{u}}_1(t))^T$$

(3)共和分の検定

(付録29)に、共和分の検定統計量の導出手順を載せる。2種類の共和分の検定方法がある。

①トレース検定

帰無仮説 $H_2(h) : h$ 個の共和分関係しか存在しない。つまり、 $\lambda_{h+1} = 0$ 言い換えれば、 $\text{rank } \Pi \leq h$ ($h = 0, \dots, p-1$)

対立仮説 $H_1 : \text{全ての変数が定常である}$ 。つまり、 $\lambda_n > 0$ 言い換えれば、 $\text{rank } \Pi \leq p$

$$\text{検定統計量} : -2 \ln Q(H_2(h) | H_1) = -T \sum_{i=h+1}^n (1 - \hat{\lambda}_i) \quad \dots \text{ (付式29-32)}$$

この検定統計量の漸近分布が、標準ブラウン運動を使った行列(説明略)のトレースが従う分布に一致することを利用して検定を行う。

②最大固有値検定

帰無仮説 $H_2(h) : h$ 個の共和分関係しか存在しない。つまり、 $\lambda_{h+1} = 0$ 言い換えれば、 $\text{rank } \Pi \leq h$ ($h = 0, \dots, p-1$)

対立仮説 $H_2(h+1) : h+1$ 個の共和分関係が存在する。つまり、 $\lambda_{h+1} > 0$ 言い換えれば、 $\text{rank } \Pi \leq h+1$

$$\text{検定統計量} : -2 \ln Q(H_2(h) | H_2(h+1)) = -T \ln (1 - \hat{\lambda}_{h+1}) \quad \dots \text{ (付式29-33)}$$

この検定統計量の漸近分布が、標準ブラウン運動を使った行列(説明略)の最大固有値が従う分布に一致することを利用して検定を行う。

21. 条件付き不均一分散モデル (1/3) ARCH モデル

金融の分野では、標準偏差のことをボラティリティ (Volatility) と呼ぶ。金融商品では、収益率の変動する時期が集約して現れることが多く、この現象をボラティリティクラスターリング (Volatility Clustering)、または、分散不均一性 (Heteroscedasticity) という。このような現象を表すのが分散 (標準偏差の二乗) の自己相関を利用したボラティリティ変動モデルである。ARCH モデルから始まり、GARCH モデルへと発展した。ボラティリティは時系列データの予測に役立つほか、予測の不確実性 (リスク) を表しているとも言え重要な指標になっている。

(a) 条件付き不均一分散のモデル化

分散の値が条件 (過去の時点のデータ) に依存する分散を条件付きの分散という。ボラティリティ変動モデルは条件付き不均一分散モデルをベースにしている。時系列データ $y(t)$ を、時点 $t-1$ までの $y(t)$ の期待値 $\mu(t)$ と攪乱項 $u(t)$ を使って、(式21-1) のようにモデル化する。 $h(t)$ はボラティリティ (条件付き分散、時間依存) を表す。
$$y(t) = \mu(t) + u(t) = \mu(t) + h^{1/2}(t) v(t) \quad \dots \text{(式21-1)}$$

ただし、攪乱項 $u(t) = h^{1/2}(t) v(t)$ であるとする。 $h(t)$ と $v(t)$ は独立、また、 $v(t) \sim \text{i.i.d. } N(0, 1)$ つまり、 $E[v(t)] = 0, V[v(t)] = E[v^2(t)] = 1$ (式21-1) を使って、 $y(t)$ の条件付き分散を求めると、

$$\begin{aligned} V[y(t) | y(t-1), y(t-2), \dots] &= E[(y(t) - \mu(t))^2 | y(t-1), y(t-2), \dots] = E[u^2(t) | y(t-1), y(t-2), \dots] = E[h(t) v^2(t) | y(t-1), y(t-2), \dots] \\ &= E[h(t) | y(t-1), y(t-2), \dots] E[v^2(t) | y(t-1), y(t-2), \dots] = E[h(t) | y(t-1), y(t-2), \dots] V[v(t) | y(t-1), y(t-2), \dots] \\ &= E[h(t) | y(t-1), y(t-2), \dots] \cdot 1 = E[h(t) | y(t-1), y(t-2), \dots] \quad \dots \text{(式21-2)} \end{aligned}$$

(式21-2)によれば、 $y(t)$ の条件付き分散 (左辺) は、 $h(t)$ の条件付き期待値 (右辺) で表される。言い換えれば、条件付き不均一分散モデルは、 $h(t)$ のモデル化で実現できることが分かる。なお、時点 $t-1$ で $h(t)$ が既知であるようなモデルを作ることを前提とする (時点 $t-1$ までの情報で $h(t)$ は確定している、予測できると考える)。つまり、(式21-2)から、 $V[y(t) | y(t-1), y(t-2), \dots] = E[u^2(t) | y(t-1), y(t-2), \dots] = E[h(t) | y(t-1), y(t-2), \dots] = h(t) > 0 \quad \dots \text{(式21-3)}$

(b) $u(t)$ の条件付きの期待値と分散

$$\begin{aligned} E[u(t) | u(t-1), u(t-2), \dots] &= E[h^{1/2}(t) v(t) | u(t-1), u(t-2), \dots] = E[h^{1/2}(t)] E[v(t) | u(t-1), u(t-2), \dots] = h^{1/2}(t) \cdot 0 = 0 \quad \dots \text{(式21-4)} \\ V[u(t) | u(t-1), u(t-2), \dots] &= E[(u(t) - E[u(t)])^2 | u(t-1), u(t-2), \dots] = E[(u(t) - 0)^2 | u(t-1), u(t-2), \dots] = E[u^2(t) | u(t-1), u(t-2), \dots] \\ &= E[h(t) v^2(t) | u(t-1), u(t-2), \dots] = E[h(t) | u(t-1), u(t-2), \dots] E[v^2(t) | u(t-1), u(t-2), \dots] \\ &= h(t) E[(v(t) - 0)^2 | u(t-1), u(t-2), \dots] = h(t) E[(v(t) - E[v(t)])^2 | u(t-1), u(t-2), \dots] = h(t) V[v(t) | u(t-1), u(t-2), \dots] = h(t) \cdot 1 = h(t) \quad \dots \text{(式21-5)} \end{aligned}$$

(c) $u(t)$ の条件なしの期待値と分散

期待値の繰り返し公式 (付録M) を使って、(式21-4) と (式21-3) をそれぞれ代入して、
$$\begin{aligned} E[u(t)] &= E[h^{1/2}(t) v(t)] = E[E[h^{1/2}(t) v(t) | u(t-1), u(t-2), \dots]] = E[0] = 0 \quad \dots \text{(式21-6)} \\ V[u(t)] &= E[(u(t) - E[u(t)])^2] = E[(u(t) - 0)^2] = E[u^2(t)] = E[(h^{1/2}(t) v(t))^2] = E[h(t) v^2(t)] = E[E[h(t) v^2(t) | u(t-1), u(t-2), \dots]] = E[h(t)] \quad \dots \text{(式21-7)} \end{aligned}$$

(1) ARCH モデル (Autoregressive Conditional Heteroskedasticity Model: 自己回帰条件付き不均一分散モデル)

ARCH(m) モデルでは、(式21-1) の $h(t)$ を以下のようにモデル化する。今期の条件付き分散は過去の攪乱項を使って表せるとしている。計算事例を (付録30) に載せる。
$$h(t) = \omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) \quad \text{ただし、} \omega > 0, \alpha_1, \alpha_2, \dots, \alpha_m \geq 0 \quad \dots \text{(式21-7)}$$

ここで、 $w(t) = u^2(t) - h(t)$ と置けば、(式21-7) より、 $E[w(t)] = E[u^2(t) - h(t)] = E[u^2(t)] - E[h(t)] = V[u(t)] - E[h(t)] = E[h(t)] - E[h(t)] = 0 \quad \dots \text{(式21-8)}$

また、 $u^2(t) = h(t) + w(t) = \omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) + w(t) \quad \dots \text{(式21-9)}$

この式は、攪乱項の2乗 $u^2(t)$ を AR(m) 過程で表したモデルと見なすことができる。よって、(式21-7) のパラメータ $\omega, \alpha_1, \alpha_2, \dots, \alpha_m$ は重回帰式の偏回帰係数として求まる。

- ① (式6-7) から分かるように、特性方程式 $1 - \alpha_1 z^{-1} - \alpha_2 z^{-2} - \dots - \alpha_m z^{-m} = 0$ の全ての解が 1 より大きいとき、 $u^2(t)$ は定常過程になる。
また、 $u^2(t)$ が定常過程であれば、分散 $E[u^2(t)]$ は時間に依らず一定の値になる。つまり、 $E[u^2(t)] = E[u^2(t-1)] = \dots = E[u^2(t-m)] = V[u(t)] \quad \dots \text{(式21-10)}$ であるので、(式21-9) より、 $V[u(t)] = E[\omega] + \alpha_1 V[u(t-1)] + \alpha_2 V[u(t-2)] + \dots + \alpha_m V[u(t-m)] + E[w(t)] = \omega + \alpha_1 V[u(t)] + \alpha_2 V[u(t)] + \dots + \alpha_m V[u(t)] + E[w(t)]$
ここで、(式21-8) を使って、 $(1 - \alpha_1 - \alpha_2 - \dots - \alpha_m) V[u(t)] = \omega + E[w(t)] = \omega + 0 = \omega > 0$
 $V[u(t)] > 0$ であるので、 $\alpha_1 + \alpha_2 + \dots + \alpha_m < 1$ のとき、 $u^2(t)$ が定常過程になる、ということもできる。
- ② $u^2(t)$ は AR 過程 (式21-9) であるので、(式6-2) より、 $E[u^2(t)] = E[h(t) v^2(t)] = E[h(t)] E[v^2(t)] = E[h(t)] \cdot 1 = E[h(t)] = \omega / (1 - \alpha_1 - \alpha_2 - \dots - \alpha_m) \quad \dots \text{(式21-11)}$

(2) ARCH(m) モデルの1期先の95% 区間予測

ARCH(m) モデルは、 $\mu(t) = 0, w(t) = 0$ とした場合、(式21-1) より、 $y(t) = u(t) = h^{1/2}(t) v(t)$ (式21-9) より、 $u^2(t) = h(t) = \omega + \alpha_1 u^2(t-1) + \dots + \alpha_m u^2(t-m)$
 $y(t+1)$ の条件付き期待値は、(式21-4) より、 $E[y(t+1) | y(t), y(t-1), \dots] = E[u(t+1) | u(t), u(t-1), \dots] = 0$
 $y(t+1)$ の条件付き分散は、(式21-5) より、 $V[y(t+1) | y(t), y(t-1), \dots] = V[u(t+1) | u(t), u(t-1), \dots] = h(t+1) = \omega + \alpha_1 y^2(t) + \dots + \alpha_m y^2(t-m+1)$
これから、 $y(t+1)$ の条件付き分布は、正規分布 $N(0, \omega + \alpha_1 y^2(t) + \dots + \alpha_m y^2(t-m+1))$ であることが分かる。従って、別途資料「統計的推定の手順」を参照して、1期先の95% 区間予測は、 $(-1.96 (\omega + \alpha_1 y^2(t) + \dots + \alpha_m y^2(t-m+1))^{1/2}, 1.96 (\omega + \alpha_1 y^2(t) + \dots + \alpha_m y^2(t-m+1))^{1/2}) \quad \dots \text{(式21-12)}$
 $\alpha_1 + \alpha_2 + \dots + \alpha_m < 1$ (定常条件) で、 $\alpha_1, \alpha_2, \dots, \alpha_m$ の値が大きくなるほど、時点 t 以前の $y(t)$ の値の影響、つまり、ボラティリティの持続性が強くなることが分かる。

22. 条件付き不均一分散モデル (2/3) GARCH モデル

(1) GARCH モデル (Generalized ARCH Model: 一般化された自己回帰条件付き不均一分散モデル)

GARCH(r, m) モデルでは、今期の条件付き分散 (ボラティリティ) $h(t)$ は、過去の攪乱項の2乗 $u^2(t-1), \dots$ に加えて、過去の条件付き分散 (ボラティリティ) $h(t-1), \dots$ を加味して表すというモデルであり、ARCH(m) モデルを一般化したモデルである。GARCH(r, m) モデルを(式21-13)に示す。計算事例を(付録31)に載せる。

$$h(t) = \omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) + \beta_1 h(t-1) + \beta_2 h(t-2) + \dots + \beta_r h(t-r) \quad \text{但し、} \omega > 0, \alpha_1, \alpha_2, \dots, \alpha_m, \beta_1, \beta_2, \dots, \beta_r \geq 0 \quad \dots (式21-13)$$

$$\begin{aligned} w(t) &= u^2(t) - h(t) \text{ と置けば、} h(t) = u^2(t) - w(t) \quad \text{従って、} \\ u^2(t) &= h(t) + w(t) = \omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) + \beta_1 h(t-1) + \beta_2 h(t-2) + \dots + \beta_r h(t-r) + w(t) \\ &= \omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) + \beta_1 (u^2(t-1) - w(t-1)) + \beta_2 (u^2(t-2) - w(t-2)) + \dots + \beta_r (u^2(t-r) - w(t-r)) + w(t) \\ &= \omega + (\alpha_1 + \beta_1) u^2(t-1) + (\alpha_2 + \beta_2) u^2(t-2) + \dots + \alpha_m u^2(t-m) + \beta_r u^2(t-r) - \beta_1 w(t-1) - \beta_2 w(t-2) - \dots - \beta_r w(t-r) + w(t) \quad \dots (式21-14) \end{aligned}$$

ここで、 $p = \max[r, m]$ とおくと、時点 $t-p$ は最も過去の時点を指すことになるので、時点 $t-m$ と時点 $t-r$ を時点 $t-p$ で置き換えることができる。つまり、

$$\begin{aligned} u^2(t) &= \omega + (\alpha_1 + \beta_1) u^2(t-1) + (\alpha_2 + \beta_2) u^2(t-2) + \dots + \alpha_p u^2(t-p) + \beta_p u^2(t-p) - \beta_1 w(t-1) - \beta_2 w(t-2) - \dots - \beta_r w(t-r) + w(t) \\ &= \omega + (\alpha_1 + \beta_1) u^2(t-1) + (\alpha_2 + \beta_2) u^2(t-2) + \dots + (\alpha_p + \beta_p) u^2(t-p) + w(t) - \beta_1 w(t-1) - \beta_2 w(t-2) - \dots - \beta_r w(t-r) \quad \dots (式21-15) \end{aligned}$$

この式は、攪乱項の2乗 $u(t)^2$ を ARMA(p, r) 過程で表したモデルであると見なすことができる ($\beta_1 = \beta_2 = \dots = \beta_r = 0$ のとき、ARCH(m) モデル(式21-9)になる)。

① GARCH モデルの利点

以下に示すように、GARCHモデルは ARCH モデルに比べて低次のモデルで表すことができる。

GARCH(1, 1) モデルで、ARCH(∞) モデルを表すことができる。

(式21-13)より、GARCH(1, 1) モデル $h(t) = \omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1)$ この式に、1期前の式 $h(t-1) = \omega + \alpha_1 u^2(t-2) + \beta_1 h(t-2)$ を代入する。

$$h(t) = \omega + \alpha_1 u^2(t-1) + \beta_1 (\omega + \alpha_1 u^2(t-2) + \beta_1 h(t-2)) = (1 + \beta_1) \omega + \alpha_1 u^2(t-1) + \alpha_1 \beta_1 u^2(t-2) + \beta_1^2 h(t-2)$$

この操作を繰り返して、 $h(t) = (1 + \beta_1 + \beta_1^2 + \beta_1^3 + \dots) \omega + \alpha_1 u^2(t-1) + \alpha_1 \beta_1 u^2(t-2) + \alpha_1 \beta_1^2 u^2(t-3) + \alpha_1 \beta_1^3 u^2(t-4) + \dots$

この式は、ARCH(∞) の式を表しており、ARCH(∞) のモデルを、GARCH(1, 1) モデルの3つのパラメータ $\omega, \alpha_1, \beta_1$ を使って表せることを示している。

② GARCH モデルの定常性

GARCH(m, r) の定常性は、(式21-15)を使って、ARMA(p, r) 過程の AR 過程部分の定常性確認すればいい。

よって、特性方程式 $1 - (\alpha_1 + \beta_1) z^{-1} - (\alpha_1 + \beta_1) z^{-2} - \dots - (\alpha_1 + \beta_1) z^{-m} = 0$ の全ての解が 1 より大きいとき、 $u^2(t)$ は定常過程になる。

また、(式21-14)より、また、(式21-8)と(式21-10)を使って、

$$\begin{aligned} E[u^2(t)] &= V[u(t)] = E[\omega + \alpha_1 u^2(t-1) + \alpha_2 u^2(t-2) + \dots + \alpha_m u^2(t-m) + \beta_1 (u^2(t-1) - w(t-1)) + \beta_2 (u^2(t-2) - w(t-2)) + \dots + \beta_r (u^2(t-r) - w(t-r)) + w(t)] \\ &= \omega + \alpha_1 E[u^2(t-1)] + \dots + \alpha_m E[u^2(t-m)] + \beta_1 (E[u^2(t-1)] - E[w(t-1)]) + \dots + \beta_r (E[u^2(t-r)] - E[w(t-r)]) + E[w(t)] \\ &= \omega + \alpha_1 V[u(t-1)] + \dots + \alpha_m V[u(t-m)] + \beta_1 (V[u(t-1)] - 0) + \dots + \beta_r (V[u(t-r)] - 0) + 0 \\ &= \omega + \alpha_1 V[u(t)] + \alpha_2 V[u(t)] + \dots + \alpha_m V[u(t)] + \beta_1 V[u(t)] + \beta_2 V[u(t)] + \dots + \beta_r V[u(t)] \end{aligned}$$

従って、 $(1 - \alpha_1 - \alpha_2 - \dots - \alpha_m - \beta_1 - \beta_2 - \dots - \beta_r) V[u(t)] = \omega > 0$
 $V[u(t)] > 0$ であるので、 $\alpha_1 + \alpha_2 + \dots + \alpha_m + \beta_1 + \beta_2 + \dots + \beta_r < 1$ のとき、 $u^2(t)$ は定常過程になる、ということもできる。

(2) GARCH モデルのパラメータ推定

よく使用されている GARCH(1, 1) モデルで考える。

$$h(t) = \omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1) \quad y(t) = \mu(t) + u(t) = \mu(t) + h^{1/2}(t) v(t)$$

時点 t では、時点 $t-1$ でのボラティリティ $h(t)$ と 攪乱項 $u(t)$ の値が既知であるとする。つまり、 $\hat{h}(t) = \omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1)$ であると見なす。

(付録G)を参考にして、 $v(t) \sim \text{i.i.d. } N(0, 1)$ であれば、確率密度関数は、 $f(y(t) | y(t-1), \dots, \hat{h}(t)) = (1 / (2 \pi \hat{h}(t)))^{1/2} \exp(- (y(t) - \mu(t))^2 / (2 \hat{h}(t))) \quad \dots (式21-16)$

よって、対数尤度は、 $\ln L = - (T / 2) \ln(2 \pi) - (1 / 2) \sum_{t=1}^T \ln \hat{h}(t) - (1 / 2) \sum_{t=1}^T (u^2(t) / \hat{h}(t)) \quad \dots (式21-17)$

従って、対数尤度を未知のパラメータで偏微分して、連立方程式の解として、パラメータ ω, α, β を求めることができる。以下、 $\sum_{t=1}^T$ は、 Σ で表す。

$$\begin{aligned} \partial \ln L / \partial \omega &= - (1 / 2) \Sigma 1 / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1)) + (1 / 2) \Sigma u^2(t) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 \\ &= (1 / 2) \Sigma (u^2(t) - \omega - \alpha_1 u^2(t-1) - \beta_1 h(t-1)) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 = 0 \quad \dots (式21-17) \end{aligned}$$

$$\begin{aligned} \partial \ln L / \partial \alpha &= - (1 / 2) \Sigma u^2(t-1) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1)) + (1 / 2) \Sigma u^2(t) u^2(t-1) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 \\ &= - (1 / 2) \Sigma u^2(t-1) (u^2(t) - \omega - \alpha_1 u^2(t-1) - \beta_1 h(t-1)) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 = 0 \quad \dots (式21-18) \end{aligned}$$

$$\begin{aligned} \partial \ln L / \partial \beta &= - (1 / 2) \Sigma h(t-1) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1)) + (1 / 2) \Sigma u^2(t) h(t-1) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 \\ &= - (1 / 2) \Sigma h^2(t-1) (u^2(t) - \omega - \alpha_1 u^2(t-1) - \beta_1 h(t-1)) / (\omega + \alpha_1 u^2(t-1) + \beta_1 h(t-1))^2 = 0 \quad \dots (式21-19) \end{aligned}$$

23. 条件付き不均一分散モデル (3/3) いろいろなGARCH モデルの概要

(1) レバレッジ効果 (Leverage Effect)

金融市場では、負のショックがより大きな影響力を持つと言われている。この影響力のことをレバレッジ効果と呼ぶ。このレバレッジ効果を表すためのモデルを以下に示す。

1) GJR モデル (Glosten, Jagannathan and Runkle model)

$$y(t) = \mu(t) + u(t), \quad u(t) = h^{1/2}(t) v(t), \quad v(t) \sim \text{i.i.d. } N(0, 1)$$

$$h(t) = \omega + \beta h(t-1) + \alpha u^2(t-1) + \gamma u^2(t-1) I(t-1) \quad \text{ただし、} u(t-1) < 0 \text{ のとき } I(t-1) = 1, \quad u(t-1) \geq 0 \text{ のとき } I(t-1) = 0$$

$I(t-1)$ を導入することでモデルに非対称性が生まれ、 $h(t)$ に対する正のショック $u^2(t)$ の影響は α 、 $h(t)$ に対する負のショック $u^2(t)$ の影響は $\alpha + \gamma$ になる。

つまり、 $\gamma > 0$ のとき、負のショックは正のショックより、条件付き分散 $h(t)$ に対して大きな影響を持つ。

2) EGARCH モデル (Exponential GARCH model)

$$y(t) = \mu(t) + u(t), \quad u(t) = h^{1/2}(t) v(t), \quad v(t) \sim \text{i.i.d. } N(0, 1)$$

$$\ln h(t) = \omega + \beta \ln h(t-1) + \gamma v(t-1) + \delta (|v(t-1)| - E[v(t-1)])$$

条件付き分散 $h(t)$ に対数を使うことで、暗に $h(t) > 0$ を求めている。 $\gamma < 0$ のとき、負のショックは正のショックより、条件付き分散 $h(t)$ に対して大きな影響を持つ。

3) GARCH-M model (GAARCH in the mean model)

$$y(t) = \mathbf{x}^T \boldsymbol{\beta} + \delta h(t) + u(t), \quad u(t) = h^{1/2}(t) v(t), \quad v(t) \sim \text{i.i.d. } N(0, 1)$$

$$h(t) = \omega + \beta h(t-1) + \alpha u^2(t-1)$$

$\delta > 0$ のとき、条件付き分散 $h(t)$ は $y(t)$ の条件付き期待値と正の相関を持つ。条件付き分散が高いと条件付き期待値も高くなるというハイリスクハイリターンモデルになる。

(2) 多変量 GARCH モデル

$$\mathbf{y}(t) = \boldsymbol{\mu}(t) + \mathbf{u}(t), \quad \mathbf{u}(t) = \mathbf{H}^{1/2}(t) \mathbf{v}(t), \quad \mathbf{v}(t) \sim \text{i.i.d. } N(\mathbf{0}, \mathbf{I}_n)$$

$$\mathbf{H}(t) = E[\mathbf{u}^T(t) \mathbf{u}(t) | \boldsymbol{\Omega}(t-1)] = E[(\mathbf{y}(t) - \boldsymbol{\mu}(t))(\mathbf{y}(t) - \boldsymbol{\mu}(t))^T | \boldsymbol{\Omega}(t-1)] \quad \text{ここで、} \mathbf{H}(t) \text{ は条件付き分散共分散行列で対称行列}$$

1) VEC モデル (vec 作用素を使うモデル) ... vech 作用素に関しては、(付録N)を参照

$$\mathbf{H}(t) \text{ に関して、} \text{vech}(\mathbf{H}(t)) = \mathbf{C} + \mathbf{B} \text{vech}(\mathbf{H}(t-1)) + \mathbf{A} \text{vech}(\mathbf{u}(t-1) \mathbf{u}^T(t-1))$$

このモデルの問題点は、モデルの次数が大きくなるとパラメータの数が急速に多くなることである(付録32)。

2) DVEC モデル (Diagonal VEC model)

$$\mathbf{H}(t)_{ij} = \mathbf{C}_{ij} + \mathbf{B}_{ij} \mathbf{H}(t-1) + \mathbf{A}_{ij} u_i(t-1) u_j(t-1) \quad \text{ここでは、例えば、} \mathbf{H}_{ij}(t) \text{ は行列 } \mathbf{H}(t) \text{ の } ij \text{ 要素を表す。} u_i(t) \text{ はベクトル } \mathbf{u}(t) \text{ の第 } i \text{ 要素を表す。}$$

1) のパラメータの数が急速に多くなるという VEC モデルの課題を克服するため、VEC モデルの行列 $\mathbf{A}$ と $\mathbf{B}$ を対角行列に限定したモデルである。

しかし、条件付き分散共分散行列 $\mathbf{H}(t)$ の要素間の相互依存関係がなくなること、つまり、多変量であることの意味が薄れることに課題がある。

3) BEKK モデル (Baba, Engle, Kraft and Kroner model)

$$\mathbf{H}(t) = \mathbf{C} + \mathbf{B} \mathbf{H}(t-1) \mathbf{B}^T + \mathbf{A} \mathbf{u}(t-1) \mathbf{u}^T(t-1) \mathbf{A}^T$$

VEC モデルより少ないパラメータ数で条件付き分散の相互依存関係を表現できるモデルで、よく使用されているモデルだとのことである。

また、BEKK モデルで、行列 $\mathbf{A}$ と $\mathbf{B}$ を対角行列に限定したのが、DBEKK モデル (Diagonal BEKK model) である。

4) CCC モデル (Constant Conditional Correlation model)

$u(t)$ の条件付き相関 $\mathbf{R}$ が時間を通して一定であると仮定し、条件付き分散と共分散だけが動的依存関係を持つとしたモデルで、よく使用されているモデルだとのことである。

$$\mathbf{D}(t) = \text{diag}(\mathbf{H}_{11}(t), \mathbf{H}_{22}(t), \dots, \mathbf{H}_{nn}(t))^{1/2}, \quad \mathbf{H}_{ij}(t) = \mathbf{R}_{ij} (\mathbf{H}_{ii}(t) \mathbf{H}_{jj}(t))^{1/2} \quad \text{で定義される行列 } \mathbf{D}(t) \text{ と } \mathbf{R} \text{ を使用して、} \mathbf{H}(t) = \mathbf{D}(t) \mathbf{R} \mathbf{D}(t)$$

(4) 相関変動モデル (Leverage Effect)

条件付き相関が期間を通して一定であると仮定し、条件付き分散共分散行列 $\mathbf{H}(t)$ をモデル化したのが前述の CCC モデルであるが、それに対して、条件付き相関を直接モデル化するのが、相関変動モデルである。

1) DCC モデル (Dynamic Conditional Correlation model)

$$\mathbf{H}(t) = \mathbf{D}(t) \mathbf{R}(t) \mathbf{D}(t)$$

CCC モデルで、CCC モデルの条件付き相関係数 $\mathbf{R}$ を時間依存する形に拡張したモデルである。 $\mathbf{R}(t)$ のモデル化方法によって、様々な DCC モデルが考えられる。

2) DCD モデル (Dynamic Conditional Dependence model)

DCC モデルは、条件付き相関係数をモデル化したものであるが、相関係数は線形依存関係を表すものであり、非線形な依存関係を捉えることはできない。

つまり、DCC モデルは、対称な依存関係しか表すことができない、といった問題点を抱えている。

そこで、より複雑な依存関係、例えば、非対称な依存構造を記述することができる分布関数、コピュラ (Copula) (付録33) が導入される。

コピュラは、通常2変量のモデルが対象になるが、このコピュラのパラメータを変動させたモデルが、DCD モデルである。

24. 閾値モデルと平滑推移モデル

観測できる変数によって状態が変化するモデルを取り上げる。

(1) **閾値モデル**(Threshold model, TR モデル)

ある変数(状態) $s(t)$ が閾値 c を越えているどうか、離散的变化によって、時系列データ $y(t)$ が従うモデルが変化するというモデルである。閾値を境にモデルが大きく変化する。AR(1) モデルを対象に、2状態の閾値モデル TAR(1) を(式24-1)に示す。

$$\left. \begin{aligned} s(t) < c \text{ のとき、} & y(t) = \phi_{01} + \phi_{11} y(t-1) + \sigma_1 \varepsilon(t), \quad \varepsilon(t) \sim \text{i.i.d.}(0, 1) \\ s(t) \geq c \text{ のとき、} & y(t) = \phi_{02} + \phi_{12} y(t-1) + \sigma_2 \varepsilon(t) \end{aligned} \right\} \quad \cdots \text{ (式24-1)}$$

(2) **平滑推移モデル**(Smooth - Transition model, ST モデル)

推移関数(Transition Function) $G(s(t); \gamma, c)$ を導入して、閾値モデルの状態変化が徐々に起こるように閾値モデルを拡張したのが、平滑推移モデルである。

例として挙げる(式24-2)では、平滑推移モデルは2つのモデル $\phi_{01} + \phi_{11} y(t-1)$ と $\phi_{02} + \phi_{12} y(t-1)$ を $1 - G(s(t))$ と $G(s(t))$ との比率で混合したモデルになっている。

$$y(t) = (\phi_{01} + \phi_{11} y(t-1))(1 - G(s(t))) + (\phi_{02} + \phi_{12} y(t-1))G(s(t)) + \sigma \varepsilon(t), \quad \varepsilon(t) \sim \text{i.i.d.}(0, 1) \quad \cdots \text{ (式24-2)}$$

以下、AR(1) モデルを対象に、2種類の平滑推移モデル STAR(1) を示す。

1) **LSTAR モデル**(Logistic STAR model)

①(式24-3)に示すように、推移関数として、1次の**ロジスティック型の関数**を使用する。

$$G(s(t); \gamma, c) = 1 / (1 + \exp(-\gamma(s(t) - c))) \quad \cdots \text{ (式24-3)}$$

(図24-1)に、この推移関数を示す。図から分かるように、対称型の関数で、 γ が推移の速さを決め、 c が推移の中心を決める。

なお、 $\gamma = 0$ では、推移が存在しないことが分かる。また、 $\gamma = \infty$ では、推移関数が0から1へ一気に変化することが予想される。

②(式24-4)に示すように、推移関数として、2次のロジスティック型の関数を使用する。

$$G(s(t); \gamma, c_1, c_2) = 1 / (1 + \exp(-\gamma(s(t) - c_1)(s(t) - c_2))) \quad \cdots \text{ (式24-4)}$$

(図24-2)に、この推移関数を示す。図から分かるように、非対称型の関数で、 γ が推移の速さを決め、 c_1 と c_2 が推移の中心を決める。

なお、 $\gamma = 0$ では、推移が存在しないことが分かる。

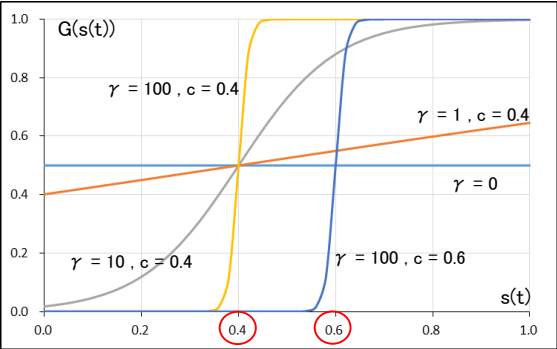
2) **ESTAR モデル**(Exponential model)

(式24-5)に示すように、推移関数として、**指数型の関数**を使用する。

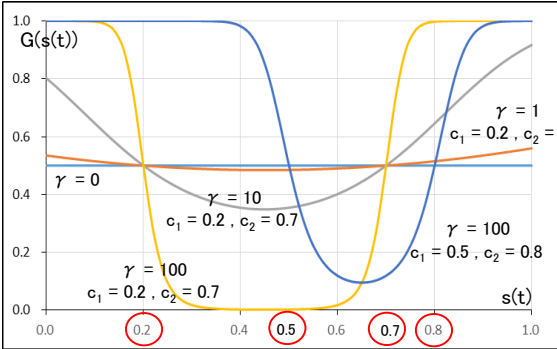
$$G(s(t); \gamma, c) = 1 - \exp(-\gamma(s(t) - c)^2) \quad \cdots \text{ (式24-5)}$$

(図24-3)に、この推移関数を示す。図から分かるように、 γ が推移の速さを決め、 c が推移の中心を決める。

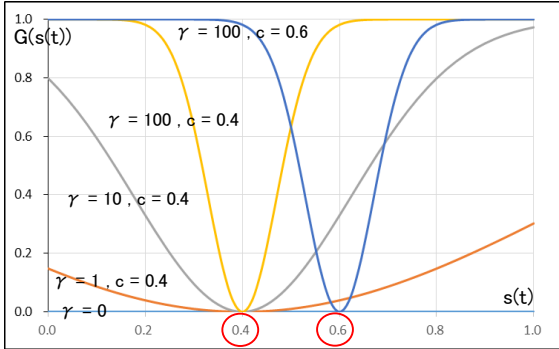
なお、 $\gamma = 0$ では、推移が存在しないことが分かる。



(図24-1) 1次のロジスティック型の推移関数



(図24-2) 2次のロジスティック型の推移関数



(図24-3) 指数型の推移関数

AR(1) モデルを対象とした、閾値モデルと平滑推移モデルの計算例を(付録34)に乗せておく。

25. マルコフ転換モデル（フィルター化確率）

時系列データ $y(t)$ の従う過程が観測できない変数の状態（レジームともいう）に応じて変化するとするモデルが、**マルコフ転換モデル**（MS モデル Markov Switching model）である。レジーム転換モデル（RS モデル Regime Switching model）とも言う。経済・ファイナンス分野の時系列データは、景気、投資家心理など観測できない変数に影響を受けるものが多く、マルコフ転換モデルを使用して、例えば、株式市場がベア市場（弱気）なのか、ブル市場（強気）なのかの特徴付けに使用されたりする。

状態（Regime）を $s(t)$ で表し、この状態の数が2つのときのマルコフ転換モデルと確率密度関数は、例えば、MSAR(1)モデル（Markov Switching Auto Regressive model）を使う場合、

$$\begin{aligned} s(t) = 1 \text{ のとき、} & y(t) = \phi_{01} + \phi_{11} y(t-1) + \sigma_1 \varepsilon(t) \\ s(t) = 2 \text{ のとき、} & y(t) = \phi_{02} + \phi_{12} y(t-1) + \sigma_2 \varepsilon(t) \end{aligned}$$

ただし、例えば、 $\varepsilon(t) \sim \text{i.i.d. } N(0, 1)$ 正規ホワイトノイズ

また、観測値 $y(t)$ の確率密度関数は、 $s(t) = 1$ のとき $f_1(y(t))$ 、 $s(t) = 2$ のとき $f_2(y(t))$ であるとする。

}

...（式25-1）

ここで、観測できない状態 $s(t)$ の従う確率過程を、（付録0）に示すマルコフ連鎖（Markov Chain）であるとするのが、マルコフ転換モデルである。計算例を、（付録35）に示す。

■ **フィルター化確率**（Filtered Probability）の計算手順（状態数が2つのときを例に）

マルコフ転換モデルのパラメータ推定とか、後述する平滑化確率の推定に必要な、フィルター化確率を求める手順を示す。

フィルター化確率は、時点 t までの時系列データ $\Omega(t) = \{y(t), y(t-1), \dots, y(1)\}$ が観測された上での $s(t)$ の状態の条件付き確率であり、 $s(t) = 1$ であるフィルター化確率が $P(s(t) = 1 | \Omega(t))$ また、 $s(t) = 2$ であるフィルター化確率が $P(s(t) = 2 | \Omega(t))$ である。

1) 確率 $P(s(t) = 1 | \Omega(t-1))$ と確率 $P(s(t) = 2 | \Omega(t-1))$ の $t = 1$ のときの値（初期値）は、 $\Omega(0)$ となるので、マルコフ連鎖の定常確率（付式0-6）と（付式0-7）で与える。

2) $y(t)$ の確率分布（条件付き確率密度関数）を求める。

$s(t) = 1$ のとき、（式25-1）と（式6-2）より、 $\mu_1 = E[y(t)] = \phi_{01} / (1 - \phi_{11})$ また、（式25-1）と（式6-3）、（式6-5）より、 $\sigma_1 = V[y(t)] = \sigma_1^2 / (1 - \phi_{11}^2)$

従って、 $f_1(y(t) | \Omega(t-1)) = (1 / (2 \pi \sigma_1^2)^{1/2}) \exp(-(y(t) - \mu_1)^2 / (2 \sigma_1^2))$...（式25-2）

同様に、 $s(t) = 2$ のとき、 $\mu_2 = E[y(t)] = \phi_{02} / (1 - \phi_{12})$ また、 $\sigma_2 = V[y(t)] = \sigma_2^2 / (1 - \phi_{12}^2)$

従って、 $f_2(y(t) | \Omega(t-1)) = (1 / (2 \pi \sigma_2^2)^{1/2}) \exp(-(y(t) - \mu_2)^2 / (2 \sigma_2^2))$...（式25-3）

3) $f_1(y(t) | \Omega(t-1))$ と $f_2(y(t) | \Omega(t-1))$ の**混合分布**（Mixture Distribution）を求める。

$f(y(t) | \Omega(t-1)) = P(s(t) = 1 | \Omega(t-1)) f_1(y(t) | \Omega(t-1)) + P(s(t) = 2 | \Omega(t-1)) f_2(y(t) | \Omega(t-1))$...（式25-4）

4) **フィルター化確率** 確率 $P(s(t) = 1 | \Omega(t))$ と確率 $P(s(t) = 2 | \Omega(t))$ を求める。

$P(s(t) = 1 | \Omega(t-1))$ を $P(s(t) = 1 | \Omega(t))$ に更新する。同様に、 $P(s(t) = 2 | \Omega(t-1))$ を $P(s(t) = 2 | \Omega(t))$ に更新する。

（付録36）より、 $s(t) = 1$ のとき、 $P(s(t) = 1 | \Omega(t)) = P(s(t) = 1 | \Omega(t-1)) f_1(y(t) | \Omega(t-1)) / f(y(t) | \Omega(t-1))$...（付式36-2）

$s(t) = 2$ のとき、 $P(s(t) = 2 | \Omega(t)) = P(s(t) = 2 | \Omega(t-1)) f_2(y(t) | \Omega(t-1)) / f(y(t) | \Omega(t-1))$...（付式36-3）

5) 確率 $P(s(t+1) = 1 | \Omega(t))$ と確率 $P(s(t+1) = 2 | \Omega(t))$ を求める。

時点 t までの情報に基づき、推移確率を使用して、時点 $t+1$ の条件付き確率である $P(s(t+1) = 1 | \Omega(t))$ と $P(s(t+1) = 2 | \Omega(t))$ を求める。

下記で、例えば、 $P(s(t+1) = 1 | s(t) = 1)$ は、状態 $s(t) = 1$ から、状態 $s(t+1) = 1$ に遷移する確率を表す。

$P(s(t+1) = 1 | \Omega(t)) = P(s(t) = 1 | \Omega(t)) P(s(t+1) = 1 | s(t) = 1) + P(s(t) = 2 | \Omega(t)) P(s(t+1) = 1 | s(t) = 2)$...（式25-5）

時点 $t+1$ で $s(t+1) = 1$ になる確率は、時点 t で $s(t) = 1$ が時点 $t+1$ で $s(t+1) = 1$ に遷移する確率と $s(t) = 2$ が $s(t+1) = 1$ に遷移する確率の合計である。

$P(s(t+1) = 2 | \Omega(t)) = P(s(t) = 1 | \Omega(t)) P(s(t+1) = 2 | s(t) = 1) + P(s(t) = 2 | \Omega(t)) P(s(t+1) = 2 | s(t) = 2)$...（式25-6）

時点 $t+1$ で $s(t+1) = 2$ になる確率は、時点 t で $s(t) = 1$ が時点 $t+1$ で $s(t+1) = 2$ に遷移する確率と $s(t) = 2$ が $s(t+1) = 2$ に遷移する確率の合計である。

6) 時点 $t+1$ にして、2)以降を繰り返す。

$P(s(t+1) = 1 | \Omega(t))$ を $P(s(t) = 1 | \Omega(t-1))$ として使う。

$P(s(t+1) = 2 | \Omega(t))$ を $P(s(t) = 2 | \Omega(t-1))$ として使う。

以上の計算手順を使って、（付録35）のマルコフ転換モデルのフィルター確率を求めた結果を（付録36）に載せる。

26. マルコフ転換モデル（平滑化確率）

状態 (Regime) を $s(t)$ で表し、この状態の数が2つのときのマルコフ転換モデルと確率密度関数は、例えば、MSAR(1)モデル (Markov Switching Auto Regressive model) を使う場合、
 $s(t) = 1$ のとき、 $y(t) = \phi_{01} + \phi_{11} y(t-1) + \sigma_{1'} \varepsilon(t)$
 $s(t) = 2$ のとき、 $y(t) = \phi_{02} + \phi_{12} y(t-1) + \sigma_{2'} \varepsilon(t)$ ただし、例えば、 $\varepsilon(t) \sim \text{i.i.d. } N(0, 1)$ 正規ホワイトノイズ } ... (式25-1)
また、確率密度関数は、 $s(t) = 1$ のとき $f_1(y(t))$, $s(t) = 2$ のとき $f_2(y(t))$ であるとする。

■ **平滑化確率 (Smoothed Probability)** の計算手順 (状態数が2つのときを例に)
全ての時系列データを使って、マルコフ転換モデルの観測できない状態 $s(t)$ の確率を、平滑化確率として各時点ごとに評価することができる。平滑化確率を求める手順を示す。
なお、時点 t までの時系列データの集合を $\Omega(t) = \{y(t), y(t-1), \dots, y(1)\}$ で表す。

平滑化確率は、 $t = T$ から初めて、 $t = T-1, \dots, t = 1$ と時点を遡って計算していく。
1) $t = T$ のときの、平滑化確率 (初期値) は、時点 T のときのフィルター化確率 $P(s(t) = 1 \mid \Omega(T))$ と $P(s(t) = 2 \mid \Omega(T))$ である。

2) 時点 t の平滑化確率を計算する (付録37)。
① $s(t) = 1$ の平滑化確率
 $P(s(t) = 1 \mid \Omega(T)) = P(s(t+1) = 1 \mid \Omega(T)) P(s(t) = 1 \mid \Omega(t)) p_{11} / P(s(t+1) = 1 \mid \Omega(t)) + P(s(t+1) = 2 \mid \Omega(T)) P(s(t) = 1 \mid \Omega(t)) p_{12} / P(s(t+1) = 2 \mid \Omega(t)) \dots$ (付37-7)
② $s(t) = 2$ の平滑化確率の導出 (付式37-2)に、(式37-6)を代入する。
 $P(s(t) = 2 \mid \Omega(T)) = P(s(t+1) = 1 \mid \Omega(T)) P(s(t) = 2 \mid \Omega(t)) p_{21} / P(s(t+1) = 1 \mid \Omega(t)) + P(s(t+1) = 2 \mid \Omega(T)) P(s(t) = 2 \mid \Omega(t)) p_{22} / P(s(t+1) = 2 \mid \Omega(t)) \dots$ (付37-8)

3) 時点を遡って、時点 $t-1$ の平滑化確率を計算する。つまり、2)の計算を繰り返す。

以上の計算手順を使って、(付録35)のマルコフ転換モデルの平滑化確率を求めた結果を(付録37)に載せる。
また、いろいろなマルコフ転換モデルの平滑化確率の計算例を(付録38)に載せる。

補足 : マルコフ転換モデルのパラメータの統計的推定方法 (概要)
マルコフ転換モデルは他のモデルと異なり 観測できない状態 $s(t)$ を持つモデルであるので、モデルパラメータの推定には工夫が必要になる。

1) **EM アルゴリズム** (Expectation - Maximization Algorithm) を使った推定
マルコフ転換モデルのパラメータは、モデルの従う確率分布の形が分かれば最尤法を使用して求めることができる。
つまり、フィルター化確率を使って混合分布を求め、この混合分布の対数尤度を最大化するパラメータを最尤推定値とすることができる。
しかし、パラメータ数が増えると通常の最尤法の計算は難しくなることが容易に推測できる。そこで、使われるのが EM アルゴリズムである (詳細は文献参照)。
E ステップ : 観測されない状態 $s(t)$ を潜在変数として事後確率の期待値を求める。
M ステップ : その上で、尤度関数を最大化するモデルのパラメータを繰り返し計算で求め、最尤推定値とする。

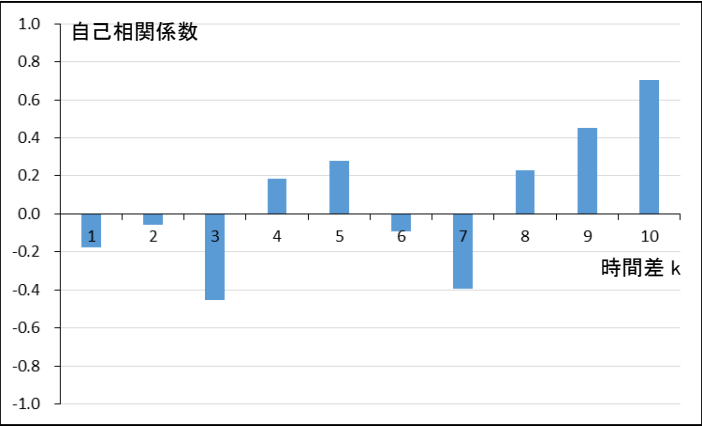
2) **ベイズ推定法** (Bayesian Estimation) を使った推定
マルコフ転換モデルが従う確率分布の形が分からないときには尤度関数の形が定まらないため、最尤法を直接使用することができない。
このような場合、つまり、モデルの尤度関数を直接計算できない場合でも、**MCMC 法** (Markov Chain Monte Carlo method 別途資料「ベイズ統計の概要」を参照)によって
例えばメトロポリス法を使って確率分布に沿ったマルコフ連鎖の点列を生成し、モデルが従う確率分布 (事後分布) の形をあぶり出すことができる。
ベイズ推定法ではこの確率分布をあぶり出したときに使用したモデルパラメータを推定値とする。

付録1. 自己相関係数とコレログラム(計算例)

(付表1-1)に、15個の時系列データを対象に、時点の差 $k = 0$ と $k = 1$ のときの自己共分散の算出手順を示す。Excel では、自己共分散の計算に、COVAR 関数が用意されている。
(付表1-2)は、COVAR 関数を使用して、 $k = 0 \sim 10$ のときの自己共分散を求め、合わせて自己相関係数も求めた結果も示す。
(付図1-1)は求めた自己相関係数のコレログラムである。

(付表1-1) 自己共分散の算出手順($k = 0$ と $k = 1$ のとき)

時点1 ～ 14 のデータ				時点 2 ～ 15 のデータ			
時点	時系列データ	偏差	偏差の二乗 (偏差平方)	時点	時系列データ	偏差	2つの時点の偏差の積
1	313.6	313.6 - 297.8	$(313.6 - 297.8)^2$	2	275.4	275.4 - 298.3	$(313.6 - 297.8) * (275.4 - 298.3)$
2	275.4	275.4 - 297.8	$(275.4 - 297.8)^2$	3	289.6	289.6 - 298.3	$(275.4 - 297.8) * (289.6 - 298.3)$
3	289.6	289.6 - 297.8	$(289.6 - 297.8)^2$	4	281.9	281.9 - 298.3	$(289.6 - 297.8) * (281.9 - 298.3)$
4	281.9	281.9 - 298.3	$(281.9 - 298.3)^2$	5	321.8	321.8 - 298.3	$(281.9 - 298.3) * (321.8 - 298.3)$
5	321.8	321.8 - 297.8	$(321.8 - 297.8)^2$	6	294.2	294.2 - 298.3	$(321.8 - 297.8) * (294.2 - 298.3)$
6	294.2	294.2 - 297.8	$(294.2 - 297.8)^2$	7	302.5	302.5 - 298.3	$(294.2 - 297.8) * (302.5 - 298.3)$
7	302.5	302.5 - 297.8	$(302.5 - 297.8)^2$	8	294.8	294.8 - 298.3	$(302.5 - 297.8) * (294.8 - 298.3)$
8	294.8	294.8 - 297.8	$(294.8 - 297.8)^2$	9	300.2	300.2 - 298.3	$(294.8 - 297.8) * (300.2 - 298.3)$
9	300.2	300.2 - 297.8	$(300.2 - 297.8)^2$	10	316.2	316.2 - 298.3	$(300.2 - 297.8) * (316.2 - 298.3)$
10	316.2	316.2 - 297.8	$(316.2 - 297.8)^2$	11	290.7	290.7 - 298.3	$(316.2 - 297.8) * (290.7 - 298.3)$
11	290.7	290.7 - 297.8	$(290.7 - 297.8)^2$	12	278.6	278.6 - 298.3	$(290.7 - 297.8) * (278.6 - 298.3)$
12	278.6	278.6 - 297.8	$(278.6 - 297.8)^2$	13	307.2	307.2 - 298.3	$(278.6 - 297.8) * (307.2 - 298.3)$
13	307.2	307.2 - 297.8	$(307.2 - 297.8)^2$	14	303.1	303.1 - 298.3	$(307.2 - 297.8) * (303.1 - 298.3)$
14	303.1	303.1 - 297.8	$(303.1 - 297.8)^2$	15	320.6	320.6 - 298.3	$(303.1 - 297.8) * (320.6 - 298.3)$
15	320.6	320.6 - 297.8	$(320.6 - 297.8)^2$				
平均	297.8		分散 $\gamma_0 =$ 偏差平方和/15 = 203.6	平均	298.3		自己共分散 $\gamma_1 =$ 偏差の積の総和 / (15 - 1) = -36.2
分散	183.7			分散	202.7		



(付図1-1) コレログラム

(付表1-2) 自己共分散と自己相関係数

時間差 (k)	k 次の自己 共分散 γ_k	k 次の自己 相関係数 ρ_k
0	203.6	1
1	-36.2	-0.178
2	-11.3	-0.056
3	-92.5	-0.454
4	38.0	0.186
5	56.9	0.279
6	-18.4	-0.090
7	80.8	-0.397
8	46.9	0.230
9	92.3	0.453
10	143.8	0.706

付録2. MA(2) 過程の例

MA(2) 過程は、(式5-1)より、 $q=2$, $y(t) = \mu + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2)$ で表される。ここで、 $\varepsilon(t)$ は(図3-2)に示す正規ホワイトノイズ W.N. (0, 1) であるとする。分散は(式5-3)を使って求め、自己相関係数は(式1-5)を使用して求める。

① $\mu = 0, \theta_1 = 0.6, \theta_2 = -0.3$ のとき

分散 $V[y(t)] = (1 + \theta_1^2 + \theta_2^2)\sigma^2 = (1 + 0.6^2 + (-0.3)^2) \cdot 1 = 1.45$

自己相関係数 $\rho_1 = (\theta_1 + \theta_2 \theta_1) / (1 + \theta_1^2 + \theta_2^2) = (0.6 + (-0.3) \cdot 0.6) / (1 + 0.6^2 + (-0.3)^2) = 0.290$

$\rho_2 = \theta_2 / (1 + \theta_1^2 + \theta_2^2) = (-0.3) / (1 + 0.6^2 + (-0.3)^2) = -0.207$

$\rho_3, \rho_4, \rho_5, \dots = 0$

この場合の MA(2) 過程を(付図2-1)に、上式に基づくコレログラムを(付図2-2)に示す。なお、(付図2-1)の時系列データを使用して求めたコレログラムは(付図2-3)に示す。(付図2-3)でも、時間差 3 以上の自己相関係数がほぼ 0 であることが分かる。

② $\mu = 0, \theta_1 = 0.8, \theta_2 = 0.4$ のとき

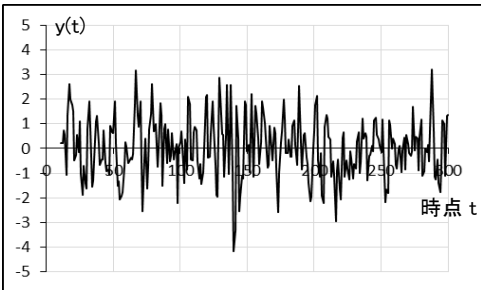
分散 $V[y(t)] = (1 + \theta_1^2 + \theta_2^2)\sigma^2 = (1 + 0.8^2 + 0.4^2) \cdot 1 = 1.80$

自己相関係数 $\rho_1 = (\theta_1 + \theta_2 \theta_1) / (1 + \theta_1^2 + \theta_2^2) = (0.8 + 0.4 \cdot 0.8) / (1 + 0.8^2 + 0.4^2) = 0.622$

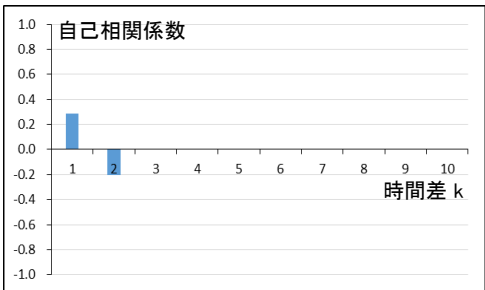
$\rho_2 = \theta_2 / (1 + \theta_1^2 + \theta_2^2) = 0.4 / (1 + 0.8^2 + 0.4^2) = 0.222$

$\rho_3, \rho_4, \rho_5, \dots = 0$

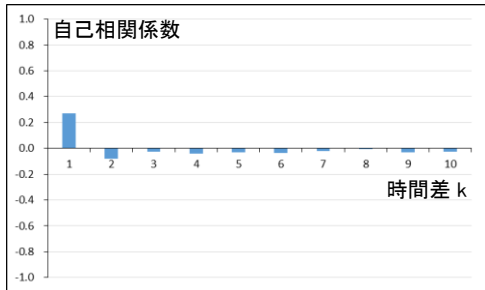
この場合の MA(2) 過程を(付図2-4)に、上式に基づくコレログラムを(付図2-5)に示す。なお、(付図2-4)の時系列データを使用して求めたコレログラムは(付図2-6)に示す。(付図2-6)でも、時間差 3 以上の自己相関係数がほぼ 0 であることが分かる。



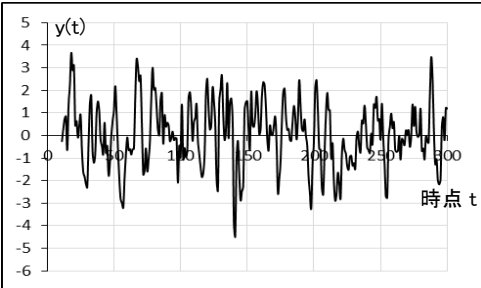
(付図2-1) MA(2) 過程 ($\mu = 0, \theta_1 = 0.6, \theta_2 = -0.3$)



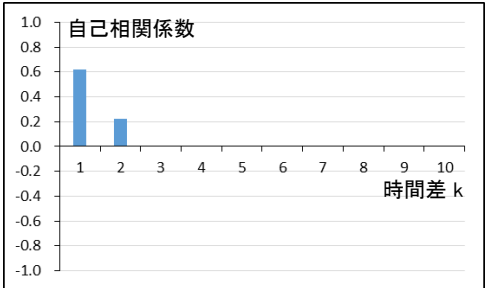
(付図2-2)
(付図2-1)のモデル式に基づくコレログラム



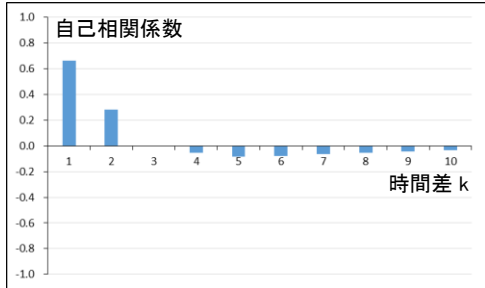
(付図2-3)
(付図2-1)のデータに基づくコレログラム



(付図2-4) MA(2) 過程 ($\mu = 0, \theta_1 = 0.8, \theta_2 = 0.4$)



(付図2-5)
(付図2-4)のモデル式に基づくコレログラム



(付図2-6)
(付図2-4)のデータに基づくコレログラム

付録3. AR(2) 過程の例 (1/2)

AR(2) 過程は、(式6-1)より、

$$p = 2, y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \varepsilon(t) \text{ で表される。}$$

ここで、 $\varepsilon(t)$ は(図3-2)に示す正規ホワイトノイズ W.N. (0, 1) であるとする。

① $c = 2, \phi_1 = 0.2, \phi_2 = 0.4$ のとき (なお、 $\phi_3 = \phi_4 = \dots = 0$ である)

・自己相関係数は、 $k = 1$ のとき、(式6-5)より、 $\rho_1 = \phi_1 \rho_0 + \phi_2 \rho_{-1}$ ここで、 $\rho_0 = \gamma_0 / \gamma_0 = 1, \rho_{-1} = \rho_1$ であるので、 $\rho_1 = \phi_1 + \phi_2 \rho_1$

従って、 $\rho_1 = \phi_1 / (1 - \phi_2)$ よって、 $\rho_1 = \phi_1 / (1 - \phi_2) = 0.2 / (1 - 0.4) = 0.333$

以降、 $k = 2$ のとき、 $\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 = 0.2 * 0.333 + 0.4 * 1 = 0.467$

$k = 3$ のとき、 $\phi_3 = 0$ であるから、 $\rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1 = 0.2 * 0.467 + 0.4 * 0.333 = 0.227$

・分散は(式6-3)より、 $\gamma_0 = \sigma^2 / (1 - \phi_1 \rho_1 - \phi_2 \rho_2) = 1 / (1 - 0.2 * 0.333 - 0.4 * 0.467) = 1.34$

この場合の AR(2)過程を(付図3-1)に、上式に基づくコレログラムを(付図3-2)に示す。

なお、(付図3-1)の時系列データを使用して求めたコレログラムは(付図3-3)に示す。当然のことであるが、2つのコレログラムの傾向は似ている。

・AR(2) 過程の特性方程式は(式2-6)より、 $\lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 - 0.2 \lambda - 0.4 = 0$

この方程式を解いて、 $\lambda_1 = 0.740, \lambda_2 = -0.540$ よって、 $|\lambda_1| = |0.740| < 1, |\lambda_2| = |-0.540| < 1$ であるので、この AR(2) 過程は定常である。

② $c = 2, \phi_1 = 0.2, \phi_2 = -0.4$ のとき (なお、 $\phi_3 = \phi_4 = \dots = 0$ である)

・自己相関係数は、 $k = 1$ のとき、 $\rho_1 = \phi_1 / (1 - \phi_2) = 0.2 / (1 - (-0.4)) = 0.143$

以降、 $k = 2$ のとき、 $\rho_2 = \phi_1 \rho_1 + \phi_2 \rho_0 = 0.2 * 0.143 + (-0.4) * 1 = -0.371$

$k = 3$ のとき、 $\phi_3 = 0$ であるから、 $\rho_3 = \phi_1 \rho_2 + \phi_2 \rho_1 = 0.2 * (-0.371) + (-0.4) * 0.143 = -0.131$

・分散は(式6-3)より、 $\gamma_0 = \sigma^2 / (1 - \phi_1 \rho_1 - \phi_2 \rho_2) = 1 / (1 - 0.2 * 0.143 - (-0.4) * (-0.371)) = 1.251$

この場合の AR(2)過程を(付図3-4)に、上式に基づくコレログラムを(付図3-5)に示す。

なお、(付図3-4)の時系列データを使用して求めたコレログラムは(付図3-6)に示す。当然のことであるが、2つのコレログラムの傾向は似ている。

・AR(2) 過程の特性方程式は(式6-6)より、 $\lambda^2 - 0.2 \lambda + 0.4 = 0$ この方程式を解いて、 $\lambda_1, \lambda_2 = (0.2 \pm j 1.249) / 2$ (j は虚数単位、 $j = (-1)^{1/2}$)

よって、 $|\lambda_1| = |\lambda_2| = ((0.2/2)^2 + (1.249/2)^2)^{1/2} = 0.632 < 1$ であるので、この AR(2) 過程は定常である。

③ $c = 2, \phi_1 = 0.2, \phi_2 = 0.9$ のとき

特性方程式は(式6-6)より、 $\lambda^2 - 0.2 \lambda - 0.9 = 0$ この方程式を解いて、 $\lambda_1 = 1.054 > 1, \lambda_2 = 0.740 < 1$ 全ての根が 1 より小さくないので非定常である。

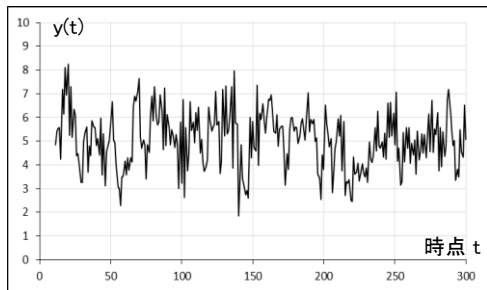
(付図3-7)より、発散している過程(非定常な過程)であることを確認できる。

④ $c = 2, \phi_1 = 0.6, \phi_2 = 0.4$ のとき

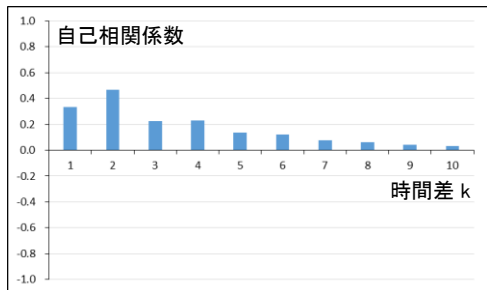
特性方程式は(式6-6)より、 $\lambda^2 - 0.6 \lambda - 0.4 = 0$ この方程式を解いて、 $\lambda_1 = 1, \lambda_2 = -0.4 < 1$ つまり、根の 1 つが 1 であり、この過程は単位根過程と

呼ばれている(後述)。時系列データを(付図3-8)に示す。

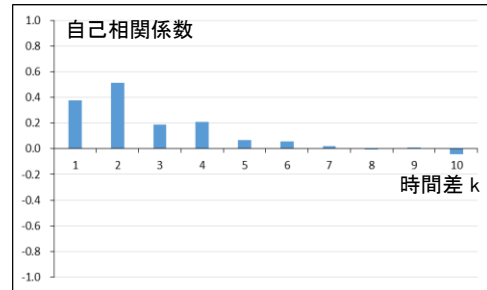
付録3. AR(2) 過程の例 (2/2)



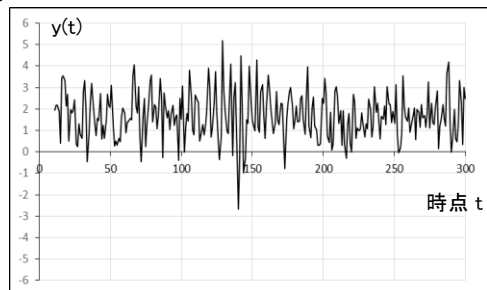
(付図3-1) AR(2) 過程 ($c = 2, \phi_1 = 0.2, \phi_2 = 0.4$)



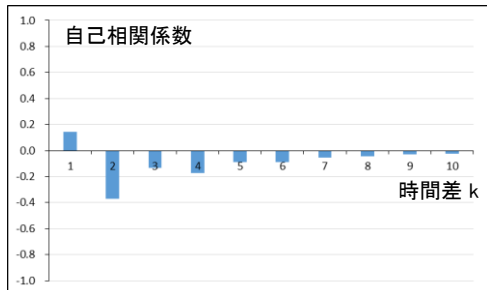
(付図3-2) (付図3-1) のコログラム
... 計算式に基づく



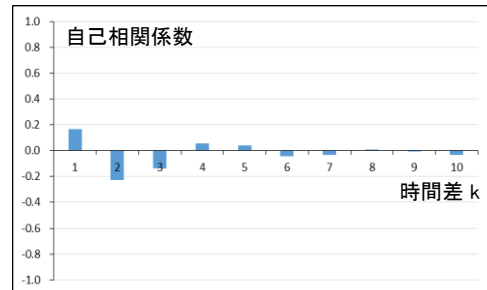
(付図3-3) (付図3-1) のコログラム
... (図3-1) のデータを使用



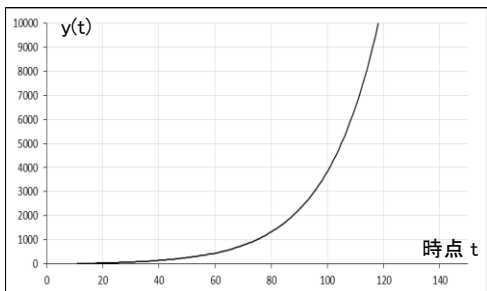
(付図3-4) AR(2) 過程 ($c = 2, \phi_1 = 0.2, \phi_2 = -0.4$)



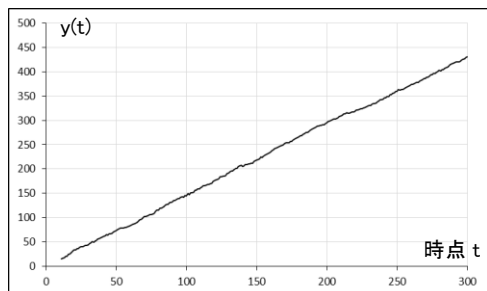
(付図3-5) (付図3-4) のコログラム
... 計算式に基づく



(付図3-6) (付図3-4) のコログラム
... (図3-4) のデータを使用



(付図3-7) AR(2) 過程 ($c = 2, \phi_1 = 0.2, \phi_2 = 0.9$)
非定常過程



(付図3-8) AR(2) 過程 ($c = 2, \phi_1 = 0.6, \phi_2 = 0.4$)
非定常過程 (単位根過程)

付録4. ARMA(2, 2) 過程の例

ARMA(2, 2) 過程は、(式7-1)より、 $p = 2$, $q = 2$, $y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \varepsilon(t) + \theta_1 \varepsilon(t-1) + \theta_2 \varepsilon(t-2)$ で表される。
ここで、 $\varepsilon(t)$ は(付3-2)に示す正規ホワイトノイズ W.N. (0, 1) であるとする。

① $c = 1$, $\phi_1 = -0.2$, $\phi_2 = -0.5$, $\theta_1 = 0.5$, $\theta_2 = 0.8$ のとき

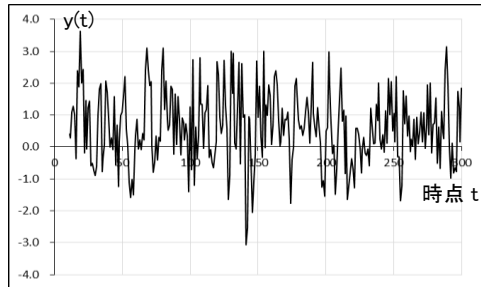
- ・定常性は、ARMA(2,2) 過程のAR(2) 過程の特性方程式で確認できる。この過程を(付図4-1)に示す。
また、コレログラムを(付図4-2)に示す。コレログラムは、時間差が大きくなるにつれ、減衰していく様子が伺える。
 $\lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 + 0.2\lambda + 0.5 = 0$ より、 $\lambda_1, \lambda_2 = -0.1 \pm 0.7j \rightarrow |\lambda| = 0.707 < 1$ 従って、この過程は定常過程である。
- ・分散 γ_0 と自己共分散 $\gamma_1, \gamma_2, \gamma_3, \dots$ は、Excel の COVAR 関数を使用して、時系列データ 280 個から求めた。
- ・また、自己相関係数は、 $\rho_k = \gamma_k / \gamma_0$ にて求めた。

② $c = 2$, $\phi_1 = 0.8$, $\phi_2 = 0.3$, $\theta_1 = 0.5$, $\theta_2 = 0.8$ のとき

- ・定常性は、ARMA(2,2) 過程のAR(2) 過程の特性方程式で確認できる。この過程を(付図4-3)に示す。
 $\lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 - 0.8\lambda - 0.3 = 0$ より、 $\lambda_1 = 1.078 > 1$, $\lambda_2 = -0.278 < 1$
従って、この過程は発散する過程(非定常過程)である。

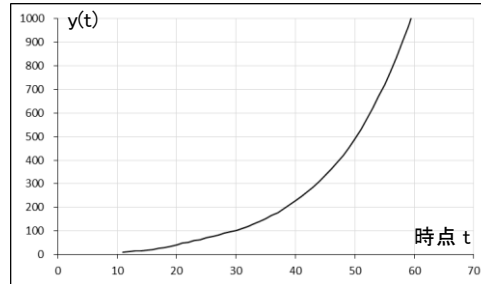
③ $c = 2$, $\phi_1 = 0.6$, $\phi_2 = 0.4$, $\theta_1 = 0.5$, $\theta_2 = 0.8$ のとき

- ・定常性は、ARMA(2,2) 過程のAR(2) 過程の特性方程式で確認できる。この過程を(付図4-4)に示す。
 $\lambda^2 - \phi_1 \lambda - \phi_2 = \lambda^2 - 0.6\lambda - 0.4 = 0$ より、 $\lambda_1 = 1$, $\lambda_2 = -0.278 < 1$
特性根のいずれかが 1 のとき、このような過程を「単位根過程」(後述)という。



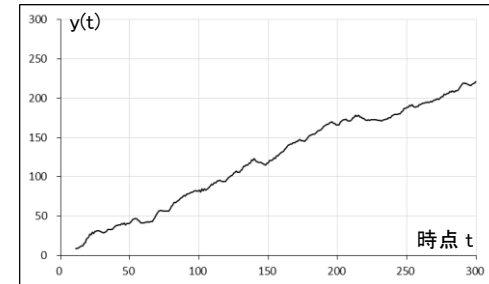
(付図4-1) ARMA(2,2)過程

($c = 1$, $\phi_1 = -0.2$, $\phi_2 = -0.5$, $\theta_1 = 0.5$, $\theta_2 = 0.8$)



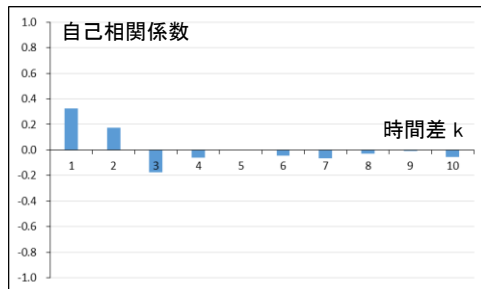
(付図4-3) ARMA(2,2)過程

($c = 1$, $\phi_1 = 0.8$, $\phi_2 = 0.3$, $\theta_1 = 0.5$, $\theta_2 = 0.8$)



(付図4-4) ARMA(2,2)過程

($c = 1$, $\phi_1 = 0.6$, $\phi_2 = 0.4$, $\theta_1 = 0.5$, $\theta_2 = 0.8$)



(付図4-2) (付図4-1) のコレログラム

付録5. AR(p) 過程の定常条件(対角行列を使って確認する方法)

(式6-1)で表される AR(p) 過程 $y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t)$ を行列とベクトルを使って表す。

$$\xi(t) = \begin{bmatrix} y(t) \\ y(t-1) \\ \vdots \\ y(t-(p-1)) \end{bmatrix}, \quad F = \begin{bmatrix} \phi_1 & \phi_2 & \dots & \phi_{p-1} & \phi_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad \varepsilon(t) = \begin{bmatrix} \varepsilon(t) \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} c \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

とおくと、AR 過程は下記のように行列を使って表すことができる。

$$\begin{aligned} \xi(t) &= c + F \xi(t-1) + \varepsilon(t) = F \xi(t-1) + \varepsilon(t) + c \\ \xi(t) &= F(F \xi(t-2) + \varepsilon(t-1) + c) + \varepsilon(t) + c = F^2 \xi(t-2) + F \varepsilon(t-1) + \varepsilon(t) + c + F c = F^2(F \xi(t-3) + \varepsilon(t-2) + c) + F \varepsilon(t-1) + \varepsilon(t) + (I + F) c \\ &= F^3 \xi(t-3) + F^2 \varepsilon(t-2) + F \varepsilon(t-1) + \varepsilon(t) + (I + F + F^2) c = \dots \\ &= F^t \xi(0) + F^{t-1} \varepsilon(1) + F^{t-2} \varepsilon(2) + \dots + F \varepsilon(t-1) + \varepsilon(t) + (I + F + F^2 + \dots + F^{t-1}) c \quad \dots \text{ (付式5-1)} \end{aligned}$$

行列 F を対角化する(別途資料「制御工学(現代制御)の基礎知識」を参照)。
 $F = T^{-1} A T$ 但し、行列 A は、対角要素に行列 F の固有値 $\lambda_1, \lambda_2, \dots, \lambda_p$ を持つ対称行列である。

行列 F の特性方程式(下記)を固有値の多項式として表す。行列式の展開には余因子行列を使う(別途資料「制御工学(現代制御)の基礎知識」を参照)。

$$\begin{aligned} |F - \lambda I| &= \begin{vmatrix} \phi_1 - \lambda & \phi_2 & \dots & \phi_{p-1} & \phi_p \\ 1 & -\lambda & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & -\lambda \end{vmatrix} = (-1)^{1+1}(\phi_1 - \lambda)(-\lambda)^{p-1} + (-1)^{1+2}\phi_2(-\lambda)^{p-2} + \dots + (-1)^{1+p-1}\phi_{p-1}(-\lambda) + (-1)^{1+p}\phi_p \\ &= (-1)^p \lambda^p + (-1)^{p-1}\phi_1 \lambda^{p-1} + (-1)^{1+2}\phi_2(-\lambda)^{p-2} + \dots + (-1)^{1+p-1}\phi_{p-1}(-\lambda) + (-1)^{1+p}\phi_p \\ &= (-1)^p (\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_{p-1} \lambda - \phi_p) = 0 \end{aligned}$$

従って、行列 F の特性方程式は次式で表される。 $\lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_{p-1} \lambda - \phi_p = 0 \quad \dots \text{ (付式5-2)}$

行列 F の固有値は(付式5-2)を解いて求まる。

ところで、 $F^2 = (T^{-1} A T)(T^{-1} A T) = T^{-1} A (T T^{-1}) A T = T^{-1} A I A T = T^{-1} A A T = T^{-1} A^2 T$ であるから、 $F^t = T^{-1} A^t T$ である。

$$\text{また、} A^2 = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \lambda_p \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \lambda_p \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & 0 & \dots & 0 & 0 \\ 0 & \lambda_2^2 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \lambda_p^2 \end{bmatrix}, \quad \dots, \quad A^t = \begin{bmatrix} \lambda_1^t & 0 & \dots & 0 & 0 \\ 0 & \lambda_2^t & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \lambda_p^t \end{bmatrix}$$

従って、全ての $|\lambda_i| < 1$ であれば、 $t \rightarrow \infty$ で、 $A^t \rightarrow 0$ よって、 $F^t \rightarrow 0$ となる。
つまり、特性方程式(付式5-2)の根(固有値)が全て 1 より小さければ、(付式5-1)で表される時系列データ $\xi(t)$ は収束する。
これは、全ての $|\lambda_i| < 1$ のとき、時系列データ $y(t), y(t-1), \dots$ が発散しないこと、定常の確率過程であることを示している。

付録6. AR(p) 過程の定常条件(z 変換を使って確認する方法)

定常条件を確認する上で(式6-1)の定数 c は関係ない。

そこで、c = 0 として(式6-1)を、 $y(t) - \phi_1 y(t-1) - \phi_2 y(t-2) - \dots - \phi_p y(t-p) = \varepsilon(t)$ と変形し、z 変換する(付録D-1, D-2)。

$$\begin{aligned} Z[y(t) - \phi_1 y(t-1) - \phi_2 y(t-2) - \dots - \phi_p y(t-p)] &= Z[\varepsilon(t)] \\ Z[y(t)] - \phi_1 Z[y(t-1)] - \phi_2 Z[y(t-2)] - \dots - \phi_p Z[y(t-p)] &= Z[\varepsilon(t)] \\ Z[y(t)] - \phi_1 z^{-1} Z[y(t)] - \phi_2 z^{-2} Z[y(t)] - \dots - \phi_p z^{-p} Z[y(t)] &= Z[\varepsilon(t)] \\ Z[y(t)](1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p}) &= Z[\varepsilon(t)] \\ \text{よって、} Z[y(t)] &= (1 / (1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p})) Z[\varepsilon(t)] \quad \dots \text{ (付式6-1)} \end{aligned}$$

この右辺の分母を特性多項式といい、

$$1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0 \quad \dots \text{ (付式6-2)}$$

この特性方程式の根を $z^{-1} = \lambda_1, \lambda_2, \dots, \lambda_p$ とおくと、特性方程式は、

$$(1 - z^{-1}/\lambda_1)(1 - z^{-1}/\lambda_2) \dots (1 - z^{-1}/\lambda_p) = 0$$

(付式6-1)の $1 / (1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p})$ を部分分数に分解すると、

$$c_1 / (1 - z^{-1}/\lambda_1) + c_2 / (1 - z^{-1}/\lambda_2) + \dots + c_p / (1 - z^{-1}/\lambda_p) = 0 \quad \text{但し、} c_1, c_2, \dots, c_p \text{ は定数}$$

これを(付式6-1)に代入すると、

$$Z[y(t)] = (c_1 / (1 - z^{-1}/\lambda_1) + c_2 / (1 - z^{-1}/\lambda_2) + \dots + c_p / (1 - z^{-1}/\lambda_p)) Z[\varepsilon(t)] \quad \dots \text{ (付式6-3)}$$

ここで、 $c_p / (1 - z^{-1}/\lambda_p)$ の項に着目し、マウローリン展開(別途資料「いろいろな確率分布」を参照)すると、

$$\begin{aligned} c_p / (1 - z^{-1}/\lambda_p) &= c_p (1 + (z^{-1}/\lambda_p) + (z^{-1}/\lambda_p)^2 + \dots) \\ c_p / (1 - z^{-1}/\lambda_p) Z[\varepsilon(t)] &= c_p (1 + (z^{-1}/\lambda_p) + (z^{-1}/\lambda_p)^2 + \dots) Z[\varepsilon(t)] = c_p (\varepsilon(t) + (1/\lambda_p) \varepsilon(t-1) + (1/\lambda_p)^2 \varepsilon(t-2) + \dots) \end{aligned}$$

この式の分散を、(付録B)の公式を使って求めると、

$$\begin{aligned} V[c_p (\varepsilon(t) + (1/\lambda_p) \varepsilon(t-1) + (1/\lambda_p)^2 \varepsilon(t-2) + \dots)] \\ = c_p^2 (V[\varepsilon(t)] + V[1/\lambda_p \varepsilon(t-1)] + V[(1/\lambda_p)^2 \varepsilon(t-2)] + \dots) \\ = c_p^2 (\sigma^2 + (1/\lambda_p)^2 \sigma^2 + (1/\lambda_p)^4 \sigma^2 + \dots) \\ = c_p^2 \sigma^2 (1 + (1/\lambda_p)^2 + (1/\lambda_p)^4 + \dots) \quad \dots \text{ (付式6-4)} \end{aligned}$$

$1 + (1/\lambda_p)^2 + (1/\lambda_p)^4 + \dots$ は、初項 1、公比 $(1/\lambda_p)^2$ の無限等比数列の和である。

この無限等比数列の和が収束するのは公比が 1 未満のときである(付録A)。これは、 $(1/\lambda_p)^2 < 1$ のときである。言い換えれば、 $|\lambda_p| > 1$ のときである。

従って、 $|\lambda_p| > 1$ のとき、(付式6-4)の右辺の分散は収束する。

これは、全ての $|\lambda_i| > 1$ のとき、 $Z[y(t)]$ が、つまり、時系列データ $y(t), y(t-1), \dots$ が発散しないこと、定常の自系列過程であることを示している。

(注1) ここでの考え方は、別途資料「制御工学(古典制御)の基礎知識」の「動的システムの応答と安定性」と同じである。

(注2) 本文から、

$$\text{特性方程式は、} \lambda^p - \phi_1 \lambda^{p-1} - \phi_2 \lambda^{p-2} - \dots - \phi_{p-1} \lambda - \phi_p = 0 \quad \dots \text{ (式6-6)}$$

$$1 - \phi_1 z^{-1} - \phi_2 z^{-2} - \dots - \phi_p z^{-p} = 0 \quad \dots \text{ (式6-7)}$$

従って、定常過程の条件は特性方程式の固有値に関する条件が、

$$\text{(式6-6)では、全ての根 } \lambda \text{ に関して、} |\lambda| < 1$$

$$\text{(式6-7)では、全ての } z^{-1} \text{ 根 に関して、} |z^{-1}| < 1, \text{ つまり、} |z| > 1 \text{ となる。}$$

付録7. 時系列過程の反転事例

(1) AR(2) 過程を MA(∞) 過程で表す事例(反転事例)

$p = 2$ のとき、AR(2) 過程 $y(t) = 0.9 y(t-1) - 0.2 y(t-2) + \varepsilon(t)$ を MA(∞) 過程で表す。

特性方程式は(式8-1)より、 $1 - 0.9 z^{-1} - (-0.2) z^{-2} = 0 \rightarrow 1 - 0.9 z^{-1} + 0.2 z^{-2} = 0 \rightarrow (1 - 0.4 z^{-1})(1 - 0.5 z^{-1}) = 0$

この特性方程式の根は $z^{-1} = 1 / 0.4 = 2.5 > 1$ と $z^{-1} = 1 / 0.5 = 0.2 > 1$ といずれも 1 より大きい。

つまり、この AR(2) 過程は定常過程である。 よって、反転可能である。

反転した MA(∞) 過程を求める。

AR(2) 過程を z 変換すると、 $Z[y(t)] = (0.9 z^{-1} - 0.2 z^{-2}) Z[y(t)] + Z[\varepsilon(t)]$

$Z[y(t)] = (1 / (1 - 0.9 z^{-1} + 0.2 z^{-2})) Z[\varepsilon(t)] = (1 / ((1 - 0.4 z^{-1})(1 - 0.5 z^{-1}))) Z[\varepsilon(t)] = ((1 / (1 - 0.4 z^{-1}))(1 / (1 - 0.5 z^{-1}))) Z[\varepsilon(t)]$

ここで、 $1 / (1 - 0.4 z^{-1})$ は、初項 1、公比 $0.4 z^{-1}$ の無限等比数列の和である。

同様に、 $1 / (1 - 0.5 z^{-1})$ は、初項 1、公比 $0.5 z^{-1}$ の無限等比数列の和である。

$$\begin{aligned} \text{よって、} Z[y(t)] &= \left(\sum_{i=0}^{\infty} (0.4 z^{-1})^i \right) \left(\sum_{j=0}^{\infty} (0.5 z^{-1})^j \right) Z[\varepsilon(t)] \\ &= (1 + (0.4 z^{-1}) + (0.4 z^{-1})^2 + (0.4 z^{-1})^3 + \dots) (1 + (0.5 z^{-1}) + (0.5 z^{-1})^2 + (0.5 z^{-1})^3 + \dots) Z[\varepsilon(t)] \\ &= (1 + (0.4 + 0.5) z^{-1} + (0.4 \cdot 0.4 + 0.4 \cdot 0.5 + 0.5 \cdot 0.5) z^{-2} + (0.4 \cdot 0.4 \cdot 0.4 + 0.4 \cdot 0.4 \cdot 0.5 + 0.4 \cdot 0.5 \cdot 0.5 + 0.5 \cdot 0.5 \cdot 0.5) z^{-3} + \dots) Z[\varepsilon(t)] \\ &= (1 + 0.9 z^{-1} + 0.61 z^{-2} + 0.369 z^{-3} + \dots) Z[\varepsilon(t)] \end{aligned}$$

従って、AR(2) 過程 $y(t) = 0.9 y(t-1) - 0.2 y(t-2) + \varepsilon(t)$ は、

MA(∞) 過程 $y(t) = \varepsilon(t) + 0.9 \varepsilon(t-1) + 0.61 \varepsilon(t-2) + 0.369 \varepsilon(t-3) + \dots$ で表すことができる。

(2) MA(2) 過程を AR(∞) 過程で表す事例(反転事例)

$q = 2$ のとき、MA(2) 過程 $y(t) = \varepsilon(t) - 0.9 \varepsilon(t-1) + 0.2 \varepsilon(t-2)$ を AR(∞) 過程で表す。

特性方程式は(式8-2)より、 $1 + (-0.9) z^{-1} + 0.2 z^{-2} = 0 \rightarrow 1 - 0.9 z^{-1} + 0.2 z^{-2} = 0 \rightarrow (1 - 0.4 z^{-1})(1 - 0.5 z^{-1}) = 0$

この特性方程式の根は $z^{-1} = 1 / 0.4 = 2.5 > 1$ と $z^{-1} = 1 / 0.5 = 0.2 > 1$ といずれも 1 より大きい。 よって反転可能である。

反転した AR(∞) 過程を求める。

MA(2) 過程を z 変換すると、 $Z[y(t)] = (1 - 0.9 z^{-1} + 0.2 z^{-2}) Z[\varepsilon(t)]$

$Z[\varepsilon(t)] = (1 / (1 - 0.9 z^{-1} + 0.2 z^{-2})) Z[y(t)] = (1 / ((1 - 0.4 z^{-1})(1 - 0.5 z^{-1}))) Z[y(t)] = ((1 / (1 - 0.4 z^{-1}))(1 / (1 - 0.5 z^{-1}))) Z[y(t)]$

ここで、 $1 / (1 - 0.4 z^{-1})$ は、初項 1、公比 $0.4 z^{-1}$ の無限等比数列の和である。

同様に、 $1 / (1 - 0.5 z^{-1})$ は、初項 1、公比 $0.5 z^{-1}$ の無限等比数列の和である。

$$\begin{aligned} \text{よって、} Z[\varepsilon(t)] &= \left(\sum_{i=0}^{\infty} (0.4 z^{-1})^i \right) \left(\sum_{j=0}^{\infty} (0.5 z^{-1})^j \right) Z[y(t)] \\ &= (1 + (0.4 z^{-1}) + (0.4 z^{-1})^2 + (0.4 z^{-1})^3 + \dots) (1 + (0.5 z^{-1}) + (0.5 z^{-1})^2 + (0.5 z^{-1})^3 + \dots) Z[y(t)] \\ &= (1 + (0.4 + 0.5) z^{-1} + (0.4 \cdot 0.4 + 0.4 \cdot 0.5 + 0.5 \cdot 0.5) z^{-2} + (0.4 \cdot 0.4 \cdot 0.4 + 0.4 \cdot 0.4 \cdot 0.5 + 0.4 \cdot 0.5 \cdot 0.5 + 0.5 \cdot 0.5 \cdot 0.5) z^{-3} + \dots) Z[y(t)] \\ &= (1 + 0.9 z^{-1} + 0.61 z^{-2} + 0.369 z^{-3} + \dots) Z[y(t)] \end{aligned}$$

これを時系列データに戻せば、 $\varepsilon(t) = y(t) + 0.9 y(t-1) + 0.61 y(t-2) + 0.369 y(t-3) + \dots$ より、

$y(t) = \varepsilon(t) - 0.9 y(t-1) - 0.61 y(t-2) - 0.369 y(t-3) + \dots$ となる。

従って、MA(2) 過程 $y(t) = \varepsilon(t) - 0.9 \varepsilon(t-1) + 0.2 \varepsilon(t-2)$ は、

AR(∞) 過程 $y(t) = \varepsilon(t) - 0.9 y(t-1) - 0.61 y(t-2) - 0.369 y(t-3) + \dots$ で表すことができる。

付録8. ボックス・ピアース(Box – Pierce)検定統計量の導出

k を時間差として、 $E[y(t)] = E[y(t - k)] = \bar{y} = 0$, $V[y(t)] = V[y(t - k)] = \sigma^2$ であること、 $y(t)$ が定常性過程であることを前提に、ボックス・ピアース検定統計量を導出する。

自己相関係数 $\rho_k(t)$ を $y(t) y(t - k)$ を使用して表す。

$$\begin{aligned}\gamma_k(t) &= \text{Cov}[y(t), y(t - k)] = (1/n) \sum_{k=1}^n (y(t) - \bar{y})(y(t - k) - \bar{y}) = (1/n) \sum_{k=1}^n (y(t) - 0)(y(t - k) - 0) = (1/n) \sum_{k=1}^n y(t) y(t - k) \\ \gamma_0(t) &= V[y(t)] = (1/n) \sum_{k=1}^n (y(t) - \bar{y})^2 = (1/n) \sum_{k=1}^n (y(t) - 0)^2 = (1/n) \sum_{k=1}^n y^2(t) = \sigma^2 \quad \text{より} \quad \sum_{k=1}^n y^2(t) = n \sigma^2 \\ \gamma_0(t - k) &= V[y(t - k)] = (1/n) \sum_{k=1}^n (y(t - k) - \bar{y})^2 = (1/n) \sum_{k=1}^n (y(t - k) - 0)^2 = \sigma^2 \quad \text{より} \quad \sum_{k=1}^n y^2(t - k) = n \sigma^2\end{aligned}$$

従って、 $\rho_k(t) = \gamma_k(t) / (\gamma_0(t) \gamma_0(t - k))^{1/2} = (1/n) \sum_{k=1}^n y(t) y(t - k) / (\sigma^2 \sigma^2)^{1/2} = \sum_{k=1}^n y(t) y(t - k) / (n \sigma^2) \quad \cdots \text{ (付式8-1)}$

ここで、 $y(t) y(t - k)$ の期待値と分散を求めると、(付録B)の分散と共分散の公式を使って、

$$\begin{aligned}E[y(t) y(t - k)] &= E[y(t)] E[y(t - k)] = 0 \cdot 0 = 0 \\ V[y(t) y(t - k)] &= V[y(t)] V[y(t - k)] + E[y(t)]^2 V[y(t - k)] + V[y(t)] E[y(t - k)]^2 = \sigma^2 \cdot \sigma^2 + 0^2 \cdot \sigma^2 + \sigma^2 \cdot 0 = \sigma^4\end{aligned}$$

一方、n が十分に大きいとき、中心極限定理(別途資料「いろいろな確率分布」を参照)により、

$y(t) y(t - k)$ の平均 $(1/n) \sum_{k=1}^n y(t) y(t - k)$ は、平均が $E[y(t) y(t - k)] = 0$ で、分散が $(1/n) V[y(t) y(t - k)] = \sigma^4 / n$ である正規分布に近づく。

つまり、 $(1/n) \sum_{k=1}^n y(t) y(t - k) \sim N(0, \sigma^4 / n) \quad \cdots \text{ (付式8-2)}$ が成立する。

従って、(付式8-2)より、

$$\begin{aligned}E[\rho_k(t)] &= E[(1/n) \sum_{k=1}^n y(t) y(t - k) / (n \sigma^2)] = (1/\sigma^2) E[(1/n) \sum_{k=1}^n y(t) y(t - k)] = (1/\sigma^2) \cdot 0 = 0 \\ V[\rho_k(t)] &= V[(1/n) \sum_{k=1}^n y(t) y(t - k) / (n \sigma^2)] = (1/\sigma^4) V[(1/n) \sum_{k=1}^n y(t) y(t - k)] = (1/\sigma^4) (\sigma^4 / n) = 1/n\end{aligned}$$

つまり、 $\rho_k(t) \sim N(0, 1/n) \quad \cdots \text{ (付式8-3)}$ が成立する。

(付式8-3)を $z = (\rho_k(t) - E[\rho_k(t)]) / (V[\rho_k(t)])^{1/2} = (\rho_k(t) - 0) / (1/n)^{1/2} = n^{1/2} \rho_k(t)$ という変数変換で標準化すると、

この新しい確率変数 z は標準正規分布に従う(別途資料「いろいろな確率分布」を参照)。

つまり、 $n^{1/2} \rho_k(t) \sim N(0, 1) \quad \cdots \text{ (付式8-4)}$ が成立する。

統計量 $R(m)$ を下式で定義する。ここで、m は時間差の上限である。つまり、時間差 $k = 1, \cdots, m$ 全体の統計量を考える。

$$R(m) = n \sum_{k=1}^m \rho_k^2(t) = \sum_{k=1}^m (n^{1/2} \rho_k(t))^2 \quad \cdots \text{ (付式8-5)}$$

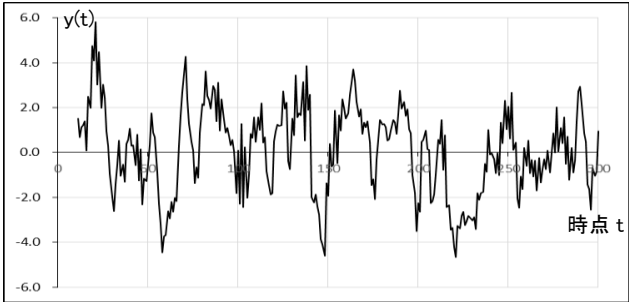
(付式8-4)より、 $n^{1/2} \rho_k(t)$ は標準正規分布に従い、 $R(m)$ は $(n^{1/2} \rho_k(t))$ の m 個の平方和であるので自由度 m のカイニ乗分布に従うことが分かる(付録N)。

$R(m)$ は、**ボックス・ピアース検定統計量** と呼ばれ、時系列分析において、自己相関係数をまとめて検定するための検定統計量である。

なお、ボックス・ピアース検定統計量は、時間差ごとの検定統計量の重みを変えた、**jung・ボックス検定統計量** $Q(m)$ として使用されている(式4-5)。

付録9. 自己相関の検定(計算例)

(付図9-1)は、ARMA(2,2) 過程で、 $c = 0$, $\phi_1 = 0.3$, $\phi_2 = 0.4$, $\theta_1 = 0.3$, $\theta_2 = 0.4$, とし、(図3-2)の正規ホワイトノイズを使用した時系列データである。この時系列データの自己相関係数の検定を行う。

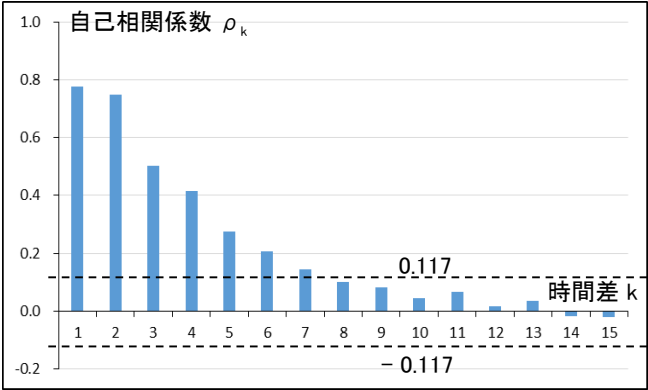


(付図9-1) ARMA(2,2) 過程

時間差を k として、自己共分散 γ_k と自己相関係数 ρ_k の計算結果を(付表9-1)に示す。また、自己相関係数 ρ_k のコレログラムを(付図9-2)に示す。

(付表9-1) 検定結果 (注) $\chi^2(p)$ 上側 5% は、Excel 関数を使って、CHISQ.INV.RT(Q(p), p) で求めることができる。

	時間差 k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
①個々の自己相関検定	γ_k	2.932	2.831	1.8945	1.562	1.033	0.778	0.539	0.374	0.302	0.169	0.243	0.060	0.134	-0.064	-0.077
	ρ_k	0.777	0.750	0.502	0.414	0.274	0.206	0.143	0.099	0.080	0.045	0.064	0.016	0.036	-0.017	-0.020
	検定結果	棄却	棄却	棄却	棄却	棄却	棄却	棄却	採択	採択	採択	採択	採択	採択	採択	採択
②まとめた自己相関検定	$Q(m)$	170.8	330.6	402.4	451.4	472.9	485.2	491.0	439.9	495.8	496.4	497.6	497.7	498.0	498.1	498.2
	$\chi^2(m)$ 5%	3.841	5.991	7.815	9.488	11.070	12.592	14.067	15.507	16.919	18.307	19.675	21.026	22.362	23.685	24.996
	p 値	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	検定結果	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却	棄却



(付図9-2) 自己相関係数のコレログラム

- (1) 時間差ごとの個々の自己相関の検定
この検定結果を(付表9-1)の①に示す。なお、 $\gamma_0 = 3.774$ である。
帰無仮説 $\rho_k = 0$ が棄却されるのは、 $|\hat{\rho}_k| > |1.96 / T^{1/2}| \dots$ (式4-4) が成立するときであるが、サンプル数 $T = 280$ であるので、 $|1.96 / T^{1/2}| = |1.96 / 200^{1/2}| = 0.117$ となり、
帰無仮説 $\rho_1 = 0, \rho_2 = 0, \dots, \rho_7 = 0$ が棄却される。つまり、これらの相関係数は有意であると言える。
(付図9-2)に棄却/採択の判断になる 0.117 のラインを記入しておく。ラインを超えると棄却されることを示す。
- (2) ある時間差までの自己相関をまとめた検定 (Ljung - Box 検定)
この検定結果を(付表9-1)の②に示す。
帰無仮説 $\rho_1 = \rho_2 = \dots = \rho_m = 0$ である。また、検定用計量は下記の通りである。
$$Q(m) = T(T+2) \sum_{k=1}^m \rho_k^2 / (T-k) \dots$$
 (式4-5)
 $Q(m)$ が自由度 m のカイニ乗分布の上側 5% 点より大きいとき、または、 $Q(m)$ の p 値が 0.05 より小さいとき
帰無仮説は棄却される。結果は(付表9-1)に示すように、 m が 1 から 15 の全ての場合の帰無仮説が棄却される。つまり、いずれかの自己相関係数が $\rho_k \neq 0$ であると判断される。(1)の結果をサポートしている。

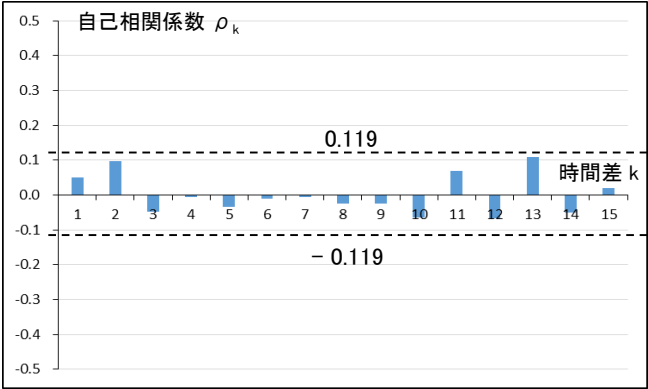
付録10. ホワイトノイズの検定(計算例)

ホワイトノイズの自己相関係数は $\rho_k(t) = 0$ ($k \neq 0$)である(本文3)。
Ljung - Box 検定(本文4)で、帰無仮説 $\rho_1 = \rho_2 = \dots = \rho_m = 0$ を棄却できず、採択することになれば、この時系列データはホワイトノイズであると言えることになる。
ここでは、(図3-2)に示す時系列データがホワイトノイズであることを確認する(使用したデータ数は 270)。

時間差を k として、自己共分散 γ_k と自己相関係数 ρ_k の計算結果を(付表10-1)に示す。また、自己相関係数 ρ_k のコレログラムを(付図10-1)に示す。

(付表10-1)検定結果 (注) $\chi^2(p)$ 上側 5% は、Excel 関数を使って、CHISQ.INV.RT(Q(p), p) で求まる。

	時間差 k	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
①個々の自己相関検定	γ_k	0.051	0.100	-0.049	-0.007	-0.035	-0.010	-0.006	-0.024	-0.025	-0.067	0.071	-0.068	0.112	-0.052	0.020
	ρ_k	0.050	0.098	-0.048	-0.007	-0.034	-0.010	-0.006	-0.024	-0.025	-0.066	0.070	-0.067	0.110	-0.052	0.019
	検定結果	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択
②まとめた自己相関検定	$Q(m)$	0.689	3.326	3.972	3.985	4.306	4.334	4.344	4.503	4.673	5.891	7.277	8.557	12.008	12.769	23.877
	$\chi^2(m)$ 5%	3.841	5.991	7.815	9.488	11.070	12.592	14.067	15.507	16.919	18.307	19.675	21.026	22.362	23.685	24.996
	p 値	0.406	0.190	0.264	0.408	0.506	0.632	0.739	0.809	0.862	0.824	0.776	0.740	0.527	0.545	0.612
	検定結果	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択	採択



(付図10-1)自己相関係数のコレログラム

- (1)時間差ごとの個々の自己相関の検定
この検定結果を(付表10-1)の①に示す。なお、 $\gamma_0 = 1.017$ である。
帰無仮説 $\rho_k = 0$ が棄却されるのは、 $|\hat{\rho}_k| > |1.96 / T^{1/2}| \dots$ (式4-4) が成立するときであるが、
サンプル数 $T = 270$ であるので、 $|1.96 / T^{1/2}| = |1.96 / 270^{1/2}| = 0.119$ となり、全ての帰無仮説 $\rho_k = 0$ は棄却されない、採択される。従って、この時系列データは、ホワイトノイズであると言える。
(付図10-1)に、棄却/採択の判定ポイントになる、0.119のラインを記入しておく。ラインを超える ρ_k はない。
- (2)ある時間差までの自己相関をまとめた検定(Ljung - Box 検定)
この検定結果を(付表10-1)の②に示す。
帰無仮説 $\rho_1 = \rho_2 = \dots = \rho_{15} = 0$ が棄却されるのは、自由度15のカイ二乗分布に従う、
 $Q(m) = T(T+2) \sum_{k=1}^m \rho_k^2 / (T-k) \dots$ (式4-5) が成立するときである。
 $Q(m)$ が自由度 m のカイ二乗分布の上側 5% 点より大きいとき、または、 $Q(m)$ の p 値が 0.05 より小さいとき
帰無仮説は棄却される。
結果は(付表10-1)に示すように、 m が 1 から 15 の全ての場合において帰無仮説は棄却されず、採択される。
つまり、この検定で、いずれの自己相関係数も $\rho_k = 0$ であると判断される。
よって、この時系列データは、ホワイトノイズであると思えることができる。

付録11. 偏自己相関係数(Partial AutoCorrelation)

AR 過程や ARMA 過程などの自己回帰分析においては、偏相関(付録E)と同様の考え方である**偏自己相関**が重要な役割を果たす。

時点 t のデータ $y(t)$ と時間差 k のデータ $y(t-k)$ の相関には、他の時点の時系列データ $y(t-1), y(t-2), \dots, y(t-(k-1))$ が影響を与えている。

そこで、時点 t のデータ $y(t)$ と時間差 k のデータ $y(t-k)$ の相関を考えるとときには、時系列データ $y(t-1), y(t-2), \dots, y(t-(k-1))$ の影響を取り除いた相関を考える必要がある。それが、偏自己相関である(付図11-1)。

・AR(p) 過程

AR(p) 過程を(式6-1)で表される。

$$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots (式6-1)$$

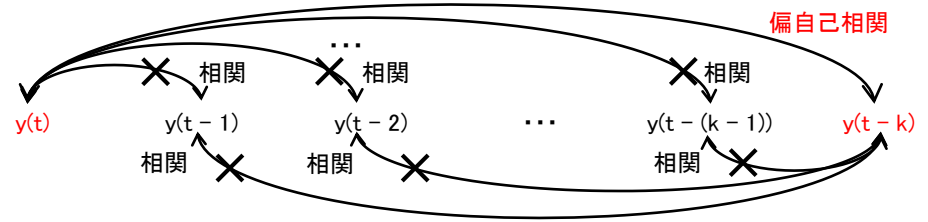
・ $y(t)$ と $y(t-k)$ の自己相関係数

ユール・ウォーカー方程式(式6-5)で表される。

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad \dots (式6-5)$$

・ $y(t)$ と $y(t-k)$ の偏自己相関係数

$y(t)$ と $y(t-k)$ の偏自己相関係数は、(式6-1)を重回帰式に見立てたときの偏回帰係数 ϕ_k になる。



(付図11-1) 偏相関係数の概念

付録12. 重回帰式の偏回帰係数の推定から、AR(p)過程のパラメータを求める

AR(p) 過程を重回帰式に見立てて(付式12-1)の偏回帰係数(AR(2) 過程のパラメータ ϕ)を最小二乗法にて推定する(別途資料「重回帰分析の概要」を参照)。

$y = Y \phi + \varepsilon \quad \dots \text{ (付式12-1)}$

$$y = \begin{bmatrix} y(t) \\ y(t-1) \\ \dots \\ y(t-N-1) \end{bmatrix}, \quad Y = \begin{bmatrix} y(t-1) & y(t-2) & \dots & y(t-p) \\ y(t-2) & y(t-3) & \dots & y(t-p-1) \\ \dots & \dots & \dots & \dots \\ y(t-N) & y(t-N-1) & \dots & y(t-p-N-1) \end{bmatrix}, \quad \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \dots \\ \phi_p \end{bmatrix}$$

パラメータ ϕ の推定値は、最小二乗法を用いて次式のように表されることが分かっている。 $\hat{\phi} = (Y^T Y)^{-1} Y^T y \quad \dots \text{ (付式12-2)}$

以下、(付式12-2)の $\hat{\phi}$ を自己相関係数 ρ で表し、Yule-Walker 法を表す(式9-4)に等しくなることを確認する。

ここで、 $Y^T = \begin{bmatrix} y(t-1) & y(t-2) & \dots & y(t-N) \\ y(t-2) & y(t-3) & \dots & y(t-N-1) \\ \dots & \dots & \dots & \dots \\ y(t-p) & y(t-p-1) & \dots & y(t-p-N-1) \end{bmatrix}$ であるので、 $y(t)$ の自己共分散 γ を使用して、

$$Y^T Y = \begin{bmatrix} \sum y^2(t-i) & \sum y(t-i) y(t-i-1) & \dots & \sum y(t-i) y(t-i-p+1) \\ \sum y(t-i-1) y(t-i) & \sum y^2(t-i-1) & \dots & \sum y(t-i-1) y(t-i-p+1) \\ \dots & \dots & \dots & \dots \\ \sum y(t-i-p+1) y(t-i) & \sum y(t-i-p+1) y(t-i-1) & \dots & \sum y^2(t-i-p+1) \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{p-2} \\ \dots & \dots & \dots & \dots \\ \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_0 \end{bmatrix}$$

両辺を γ_0 で割ると、 $Y^T Y / \gamma_0 = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}$

また、 $Y^T y / \gamma_0 = \begin{bmatrix} y(t-1) & y(t-2) & \dots & y(t-N) \\ y(t-2) & y(t-3) & \dots & y(t-N-1) \\ \dots & \dots & \dots & \dots \\ y(t-p) & y(t-p-1) & \dots & y(t-p-N-1) \end{bmatrix} \begin{bmatrix} y(t) \\ y(t-1) \\ \dots \\ y(t-N-1) \end{bmatrix} / \gamma_0 = \begin{bmatrix} \sum y(t-i) y(t-i+1) \\ \sum y(t-i-1) y(t-i+1) \\ \dots \\ \sum y(t-i-p+1) y(t-i+1) \end{bmatrix} / \gamma_0 = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_p \end{bmatrix} / \gamma_0 = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \rho_p \end{bmatrix}$

以上の結果を(付式12-2)に代入すると、(付式12-3)が導かれる。

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \dots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \dots & \dots & \dots & \dots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \dots \\ \rho_p \end{bmatrix} \quad \dots \text{ (付式12-3)}$$

この(付式12-3)は、(式9-4)と同じである。

つまり、Yule-Walker 法(そして、レビンソン・ダービンのアルゴリズム)で AR(p) 過程のパラメータを推定する方法と、AR(p) 過程を重回帰式であると見なして、AR(p) 過程のパラメータを重回帰式の偏回帰係数として求める方法は、いずれも推定残差を最小化する最小二乗法を使用する推定法であると言える。

付録13. レヴィンソン・ダービンのアルゴリズム (Levinson-Durbin) (1/3)

レヴィンソン・ダービンのアルゴリズムは、AR 過程のモデル式の係数 $\phi_1, \phi_2, \dots, \phi_p$ と $\varepsilon(t)$ の分散 σ^2 を時系列データと自己共分散から算出するアルゴリズムである。

$$y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots (式6-1)$$

なお、(式6-1)の定数 c は以下の方法で求めることができるので、(式6-1)から定数 c を除いた(付式13-1)でこのアルゴリズムを確認する。

$$y(t) = \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_p y(t-p) + \varepsilon(t) \quad \dots (付式13-1)$$

(定数 c を求める方法)

$z(t) = y(t) - \mu$ とおくと、 $E[z(t)] = E[y(t) - \mu] = E[y(t)] - \mu = \mu - \mu = 0$ であり、(式6-1)を書き換えると、

$$z(t) + \mu = c + \phi_1 (z(t-1) + \mu) + \phi_2 (z(t-2) + \mu) + \dots + \phi_p (z(t-p) + \mu) + \varepsilon(t)$$

$$\rightarrow z(t) = \phi_1 z(t-1) + \phi_2 z(t-2) + \dots + \phi_p z(t-p) + \varepsilon(t) + c - \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$$

$$\rightarrow E[z(t)] = E[\phi_1 z(t-1) + \phi_2 z(t-2) + \dots + \phi_p z(t-p)] + E[\varepsilon(t)] + E[c] - E[\mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)]$$

$$\rightarrow 0 = 0 + 0 + \dots + 0 + 0 + c - \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p) \rightarrow c = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$$

つまり、定数 c は、時系列データの期待値 μ と係数 $\phi_1, \phi_2, \dots, \phi_p$ を使って求まる。つまり、係数 $\phi_1, \phi_2, \dots, \phi_p$ を求めれば、(式6-1)の係数全てが求まる。

AR(p) 過程の分散と自己共分散の式を下記に再掲する。なお、 $\phi_m(p)$ と $\sigma(p)$ は、 p 次のAR(p) 過程での値であることを示す。

$$(式6-3)より、分散: \gamma_0 - \sum_{m=1}^p \gamma_m \phi_m(p) = \sigma^2(p) \quad \dots (付式13-2)$$

$$(式6-4)より、共分散: \gamma_k - \sum_{m=1}^p \gamma_{k-m} \phi_m(p) = 0 \quad \dots (付式13-3)$$

$\gamma_k = \gamma_{-k}$ であることに注意して、(付式13-2)と(付式13-3)をまとめて、下記のような行列の式で表す。

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_p \\ \gamma_1 & \gamma_{1-1} & \gamma_{1-2} & \dots & \gamma_{1-p} \\ \gamma_2 & \gamma_{2-1} & \gamma_{2-2} & \dots & \gamma_{2-p} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_{p-p} \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \dots \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} \sigma^2(p) \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_p \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \dots \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} \sigma^2(p) \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad \dots (付式13-4)$$

この係数行列は、**テプリッツ(Toeplitz)型の行列**と呼ばれているが、黄色で塗りつぶした要素を抜き出すと、以下の式が得られる。

$$\begin{bmatrix} \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \dots \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-2} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_1(p) \\ \phi_2(p) \\ \dots \\ \phi_p(p) \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_p \end{bmatrix}$$

同様に、AR(p-1) 過程に関しても下記の行列の黄色で塗りつぶした要素を抜き出して、

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p-1) \\ -\phi_2(p-1) \\ \dots \\ -\phi_{p-1}(p-1) \end{bmatrix} = \begin{bmatrix} \sigma^2(p-1) \\ 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \dots \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{p-2} & \gamma_{p-3} & \gamma_{p-4} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_1(p-1) \\ \phi_2(p-1) \\ \dots \\ \phi_{p-1}(p-1) \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \dots \\ \gamma_{p-1} \end{bmatrix} \quad \dots (付式13-5)$$

(付式13-5)で、例えば1行目に注目して、和の並びを並び替える。

$$\gamma_0 \phi_1(p-1) + \gamma_1 \phi_2(p-1) + \dots + \gamma_{p-3} \phi_{p-2}(p-1) + \gamma_{p-2} \phi_{p-1}(p-1) = \gamma_1 \rightarrow \gamma_{p-2} \phi_{p-1}(p-1) + \gamma_{p-3} \phi_{p-2}(p-1) + \dots + \gamma_1 \phi_2(p-1) + \gamma_0 \phi_1(p-1) = \gamma_1$$

この並び替えた要素を行列の最下段におく。次いで2行目の和の並びを並び替え下から2行目に置く。この操作を繰り返して(付式13-6)を得る。

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{p-3} \\ \dots & \dots & \dots & \dots & \dots \\ \gamma_{p-2} & \gamma_{p-3} & \gamma_{p-4} & \dots & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_{p-1}(p-1) \\ \phi_{p-2}(p-1) \\ \dots \\ \phi_1(p-1) \end{bmatrix} = \begin{bmatrix} \gamma_{p-1} \\ \gamma_{p-2} \\ \dots \\ \gamma_1 \end{bmatrix} \quad \dots (付式13-6)$$

付録13. レヴィンソン・ダービンのアルゴリズム (Levinson-Durbin) (2/3)

(付式13-4)の黄色に塗りつぶした要素を青色と緑色に塗りつぶした要素に分けた行列の式にする。

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_{p-1} & \gamma_p \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} & \gamma_{p-1} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} & \gamma_{p-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 & \gamma_1 \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \cdots \\ -\phi_{p-1}(p) \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} \sigma^2(p) \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} -\phi_1(p) \\ -\phi_2(p) \\ \cdots \\ -\phi_{p-1}(p) \end{bmatrix} = - \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \cdots \\ \gamma_{p-1} \end{bmatrix} + \phi_p(p) \begin{bmatrix} \gamma_{p-1} \\ \gamma_{p-2} \\ \cdots \\ \gamma_1 \end{bmatrix}$$

この式の①に(付式13-5)の左辺を、②に(付式13-6)の左辺を代入する。

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} -\phi_1(p) \\ -\phi_2(p) \\ \cdots \\ -\phi_{p-1}(p) \end{bmatrix} = - \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_1(p-1) \\ \phi_2(p-1) \\ \cdots \\ \phi_{p-1}(p-1) \end{bmatrix} + \phi_p(p) \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-2} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-3} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} \phi_{p-1}(p-1) \\ \phi_{p-2}(p-1) \\ \cdots \\ \phi_1(p-1) \end{bmatrix} \rightarrow \begin{bmatrix} \phi_1(p) \\ \phi_2(p) \\ \cdots \\ \phi_{p-1}(p) \end{bmatrix} = \begin{bmatrix} \phi_1(p-1) \\ \phi_2(p-1) \\ \cdots \\ \phi_{p-1}(p-1) \end{bmatrix} - \phi_p(p) \begin{bmatrix} \phi_{p-1}(p-1) \\ \phi_{p-2}(p-1) \\ \cdots \\ \phi_1(p-1) \end{bmatrix}$$

導出された行列の式は、 $\phi_j(p) = \phi_j(p-1) - \phi_p(p) \phi_{p-j}(p-1)$ 但し、 $j = 1, 2, \cdots, p-1$ $\cdots$ (付式13-7)と書き直すことができる。

ここで、 $\phi_p(p)$ は**偏自己相関係数**(Partial Auto Correlation: **PARCOR**)である(付録E)。

この式から(付式13-8)が得られる。

次いで、(付式13-8)を(付式13-4)に代入する。

$$\begin{bmatrix} 1 \\ -\phi_1(p) \\ -\phi_2(p) \\ \cdots \\ -\phi_{p-1}(p) \\ -\phi_p(p) \end{bmatrix} = \begin{bmatrix} 1 \\ -\phi_1(p-1) \\ -\phi_2(p-1) \\ \cdots \\ -\phi_{p-1}(p-1) \\ 0 \end{bmatrix} - \phi_p(p) \begin{bmatrix} 0 \\ -\phi_{p-1}(p-1) \\ -\phi_{p-2}(p-1) \\ \cdots \\ -\phi_1(p-1) \\ 1 \end{bmatrix} \cdots \text{(付式13-8)}$$

$$\begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \cdots & \gamma_p \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{p-1} & \gamma_{p-2} & \gamma_{p-3} & \cdots & \gamma_1 \\ \gamma_p & \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(p-1) \\ -\phi_2(p-1) \\ \cdots \\ -\phi_{p-1}(p-1) \\ 0 \end{bmatrix} - \phi_p(p) \begin{bmatrix} 0 \\ -\phi_{p-1}(p-1) \\ -\phi_{p-2}(p-1) \\ \cdots \\ -\phi_1(p-1) \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma^2(p) \\ 0 \\ 0 \\ \cdots \\ 0 \\ 0 \end{bmatrix} \cdots \text{(付式13-9)}$$

(付式13-9)の左辺を分解した**第1項**の各行は、

- ③第1行: $\gamma_0 - \phi_1(p-1) \gamma_1 - \phi_2(p-1) \gamma_2 - \cdots - \phi_{p-1}(p-1) \gamma_{p-1} = \gamma_0 - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_m = \sigma^2(p-1) \cdots$ (付式13-2)より
- ④第2行: $\gamma_1 - \phi_1(p-1) \gamma_0 - \phi_2(p-1) \gamma_1 - \cdots - \phi_{p-1}(p-1) \gamma_{p-2} = \gamma_1 - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{m-1} = \gamma_1 - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{1-m} = 0 \cdots$ (付式13-3)より
- ⑤第p行: $\gamma_{p-1} - \phi_1(p-1) \gamma_{p-2} - \phi_2(p-1) \gamma_{p-3} - \cdots - \phi_{p-1}(p-1) \gamma_0 = \gamma_{p-1} - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-1-m} = 0 \cdots$ (付式13-3)より
- ⑥第p+1行: $\gamma_p - \phi_1(p-1) \gamma_{p-1} - \phi_2(p-1) \gamma_{p-2} - \cdots - \phi_{p-1}(p-1) \gamma_1 = \gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \neq 0$

また、(付式13-9)の左辺を分解した**第2項**の各行は(ここでは、係数にあたる $\phi_p(p)$ の記載を割愛)、

- 第1行: $\phi_{p-1}(p-1) \gamma_1 + \phi_{p-2}(p-1) \gamma_2 + \cdots + \phi_1(p-1) \gamma_{p-1} - \gamma_p = -\gamma_p + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \neq 0 \cdots$ 並びを逆に、上記⑥と同じ
- 第2行: $\phi_{p-1}(p-1) \gamma_0 + \phi_{p-2}(p-1) \gamma_1 + \cdots + \phi_1(p-1) \gamma_{p-2} - \gamma_{p-1} = -\gamma_{p-1} + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-1-m} = 0 \cdots$ 並びを逆に、上記⑤と同じ
- 第p行: $\phi_{p-1}(p-1) \gamma_{p-2} + \phi_{p-2}(p-1) \gamma_{p-3} + \cdots + \phi_1(p-1) \gamma_0 - \gamma_1 = -\gamma_1 + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{m-1} = 0 \cdots$ 並びを逆に、上記④と同じ
- 第p+1行: $\phi_{p-1}(p-1) \gamma_{p-1} + \phi_{p-2}(p-1) \gamma_{p-2} - \cdots - \phi_1(p-1) \gamma_1 - \gamma_0 = -\gamma_0 + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_m = \sigma^2(p-1) \cdots$ 並びを逆に、上記③と同じ

付録13. レヴィンソン・ダービンのアルゴリズム (Levinson-Durbin) (3/3)

従って、(付式13-9)より、下記の(付式13-10)が得られる。

$$\begin{bmatrix} \sigma^2(p-1) \\ 0 \\ 0 \\ \vdots \\ 0 \\ \gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \end{bmatrix} - \phi_p(p) \begin{bmatrix} -\gamma_p + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \\ 0 \\ 0 \\ \vdots \\ 0 \\ \sigma^2(p-1) \end{bmatrix} = \begin{bmatrix} \sigma^2(p) \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \quad \cdots \text{ (付式13-10)}$$

(付式13-10)の最下行目から、

$$\gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} - \phi_p(p) \sigma^2(p-1) = 0 \quad \cdots \text{ (付式13-11)}$$

従って、 $\phi_p(p) = \{ \gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \} / \sigma^2(p-1) \quad \cdots \text{ (付式13-12)}$

この式は、AR(p-1) 過程の係数 $\phi_1(p-1), \phi_2(p-1), \cdots, \phi_{p-1}(p-1)$ と分散 $\sigma^2(p-1)$ を使用して、AR(p) 過程の 自己共分散 γ_p が得られたら、AR(p) 過程の偏自己相関係数 $\phi_p(p)$ を求めることができることを示している。

次いで、(付式13-10)の1行目から、

$$\sigma^2(p-1) - \phi_p(p) (-\gamma_p + \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m}) = \sigma^2(p)$$

また、(付式13-11)から、 $\gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} = \phi_p(p) \sigma^2(p-1)$ であるので、 $\sigma^2(p) = \sigma^2(p-1) - \phi_p(p) (\phi_p(p) \sigma^2(p-1)) = \sigma^2(p-1) \{1 - \phi_p^2(p)\} \quad \cdots \text{ (付式13-13)}$

この式は、AR(p-1) 過程の分散 $\sigma^2(p-1)$ と、(付式13-12)で求めた AR(p) 過程の偏自己相関係数 $\phi_p(p)$ を使用して、AR(p) 過程の分散 $\sigma^2(p)$ を求めることができることを示している。そして、(付式13-7を)再掲載すると、

$$\phi_j(p) = \phi_j(p-1) - \phi_p(p) \phi_{p-j}(p-1) \quad \text{但し、} j = 1, 2, \cdots, p-1 \quad \cdots \text{ (付式13-7)}$$

この式は、(付式13-1)で求めた AR(p) 過程の $\phi_p(p)$ と AR(p-1) 過程の係数 $\phi_j(p-1)$ と $\phi_{p-j}(p-1)$ から、AR(p) 過程の係数 $\phi_j(p)$ を求めることができることを示している。

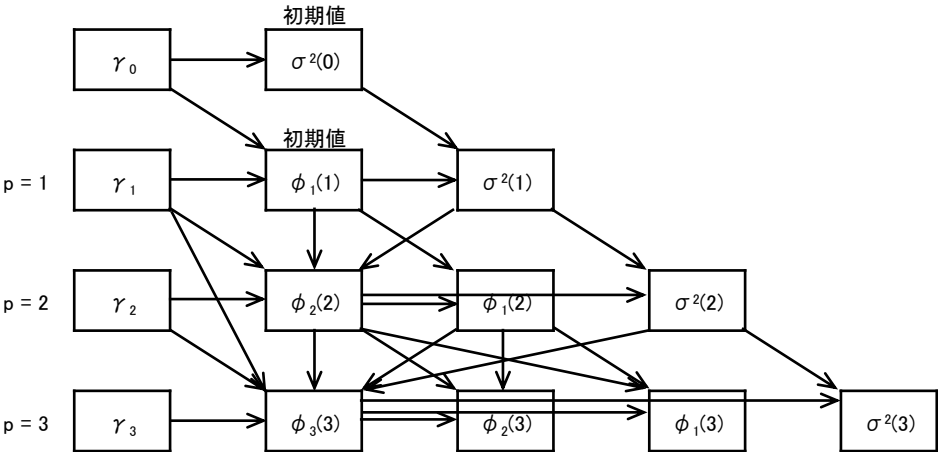
以上まとめると、(付式13-7)、(付式13-12)、(付式13-13) がパラメータの算出に使用する式で、レヴィンソン・ダービンのアルゴリズムを構成している。
 (付図13-1)に、p = 3 までのレヴィンソン・ダービンのアルゴリズムの手順を簡単に示す。AR(1) 過程の $\phi_1(1), \sigma^2(1)$ 次いで、AR(2) 過程の $\phi_1(2), \phi_2(2), \sigma^2(2)$ と順次、パラメータを求めることできる。

【レヴィンソン・ダービンのアルゴリズム】

- (1) (付式13-12) $\phi_p(p) = \{ \gamma_p - \sum_{m=1}^{p-1} \phi_m(p-1) \gamma_{p-m} \} / \sigma^2(p-1)$
- (2) (付式13-7) $\phi_j(p) = \phi_j(p-1) - \phi_p(p) \phi_{p-j}(p-1) \quad \text{但し、} j = 1, 2, \cdots, p-1$
- (3) (付式13-13) $\sigma^2(p) = \sigma^2(p-1) \{1 - \phi_p^2(p)\}$

- ・(付式13-4)において、p = 0 において得られる $\gamma_0 = \sigma^2(0)$ つまり、 $\sigma^2(0) = \gamma_0$ が $\sigma^2(0)$ の初期値になる。
- ・(付式13-4)において、p = 1 においた下記の行列の式の2行目から得られる $\gamma_1 - \phi_1(1) \gamma_0 = 0$ から、 $\phi_1(1) = \gamma_1 / \gamma_0$ が $\phi_1(1)$ の初期値になる。

$$\begin{bmatrix} \gamma_0 & \gamma_1 \\ \gamma_1 & \gamma_0 \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1(1) \end{bmatrix} = \begin{bmatrix} \sigma^2(1) \\ 0 \end{bmatrix}$$



(付図13-1)レピンソン・ダービンのアルゴリズムの処理の流れ

付録14. 重回帰式の推定式(最小二乗法)を使ったAR(1) 過程のパラメータの推定 (1/2)

最小二乗法を使用した重回帰式(付式14-1)の偏回帰係数の推定式(付式14-2)を使って、AR(1) モデルのパラメータを推定する方法を示す。
なお、(付式14-1)はベクトルと行列で表した式であり、(式6-1)において、 $p = 1$ とした式である。

$y = \phi Y + \varepsilon \quad \dots \text{(付式14-1)} \quad \hat{\phi} = (Y^T Y)^{-1} Y^T y \quad \dots \text{(付式14-2)}$

ただし、 $y = \begin{bmatrix} y(t-1) \\ y(t-2) \\ \vdots \\ y(t-N) \end{bmatrix}, \phi = \begin{bmatrix} c \\ \phi_1 \end{bmatrix}, Y = \begin{bmatrix} 1 & y(t-2) \\ 1 & y(t-3) \\ \vdots & \vdots \\ 1 & y(t-N-1) \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon(t-1) \\ \varepsilon(t-2) \end{bmatrix}$

$Y^T Y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ y(t-2) & y(t-3) & \dots & y(t-N-1) \end{bmatrix} \begin{bmatrix} 1 & y(t-2) \\ 1 & y(t-3) \\ \vdots & \vdots \\ 1 & y(t-N-1) \end{bmatrix} = \begin{bmatrix} N & \sum_{i=2}^{N+1} y(t-i) \\ \sum_{i=2}^{N+1} y(t-i) & \sum_{i=2}^{N+1} y(t-i)^2 \end{bmatrix}$ よって、 $(Y^T Y)^{-1} = \frac{1}{N \sum_{i=2}^{N+1} y(t-i)^2 - (\sum_{i=2}^{N+1} y(t-i))^2} \begin{bmatrix} \sum_{i=2}^{N+1} y(t-i)^2 & -\sum_{i=2}^{N+1} y(t-i) \\ -\sum_{i=2}^{N+1} y(t-i) & N \end{bmatrix}$

また、 $Y^T y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ y(t-2) & y(t-3) & \dots & y(t-N-1) \end{bmatrix} \begin{bmatrix} y(t-1) \\ y(t-2) \\ \vdots \\ y(t-N) \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^N y(t-i) \\ \sum_{i=1}^N y(t-i-1) y(t-i) \end{bmatrix}$

よって、 $\hat{\phi} = (Y^T Y)^{-1} Y^T y = \frac{1}{N \sum_{i=2}^{N+1} y(t-i)^2 - (\sum_{i=2}^{N+1} y(t-i))^2} \begin{bmatrix} \sum_{i=2}^{N+1} y(t-i)^2 & \sum_{i=1}^N y(t-i) & -\sum_{i=2}^{N+1} y(t-i) & \sum_{i=1}^N y(t-i-1) y(t-i) \\ -\sum_{i=2}^{N+1} y(t-i) & \sum_{i=1}^N y(t-i) & +N & \sum_{i=1}^N y(t-i-1) y(t-i) \end{bmatrix} \dots \text{(付式14-3)}$

ここで、(付式14-3)の分子の第1項 A と第2項 B を求める。

$A = \sum_{i=2}^{N+1} y(t-i)^2 \sum_{i=1}^N y(t-i) - \sum_{i=2}^{N+1} y(t-i) \sum_{i=1}^N y(t-i-1) y(t-i) = \sum_{i=1}^N y(t-i-1)^2 \sum_{i=1}^N y(t-i) - \sum_{i=1}^N y(t-i-1) \sum_{i=1}^N y(t-i-1) y(t-i) = \sum_{i=1}^N \bar{y}(t-i) \sum_{i=1}^N y(t-i-1)^2 - \sum_{i=1}^N \bar{y}(t-i-1) \sum_{i=1}^N y(t-i-1) y(t-i)$
 $= N \bar{y}(t-i) \sum_{i=1}^N y(t-i-1)^2 - 2 N^2 \bar{y}(t-i) \bar{y}(t-i-1)^2 + N^2 \bar{y}(t-i) \bar{y}(t-i-1)^2 - N \bar{y}(t-i-1) \sum_{i=1}^N y(t-i-1) y(t-i) + 2 N^2 \bar{y}(t-i) \bar{y}(t-i-1)^2 - N^2 \bar{y}(t-i) \bar{y}(t-i-1)^2$
 $= N \bar{y}(t-i) \{ \sum_{i=1}^N y(t-i-1)^2 - 2 N \bar{y}(t-i-1)^2 + N \bar{y}(t-i-1)^2 \} - N \bar{y}(t-i-1) \{ \sum_{i=1}^N y(t-i-1) y(t-i) - N \bar{y}(t-i) \bar{y}(t-i-1) - N \bar{y}(t-i) \bar{y}(t-i-1) + N \bar{y}(t-i) \bar{y}(t-i-1) \}$
 $= N \bar{y}(t-i) \{ \sum_{i=1}^N y(t-i-1)^2 - \sum_{i=1}^N 2 y(t-i-1) \bar{y}(t-i-1) + \sum_{i=1}^N \bar{y}(t-i-1)^2 \} - N \bar{y}(t-i-1) \{ \sum_{i=1}^N y(t-i-1) y(t-i) - \sum_{i=1}^N y(t-i) \bar{y}(t-i-1) - \sum_{i=1}^N \bar{y}(t-i) y(t-i-1) + \sum_{i=1}^N \bar{y}(t-i) \bar{y}(t-i-1) \}$
 $= N \bar{y}(t-i) \sum_{i=1}^N \{ y(t-i-1)^2 - 2 y(t-i-1) \bar{y}(t-i-1) + \bar{y}(t-i-1)^2 \} - N \bar{y}(t-i-1) \sum_{i=1}^N \{ y(t-i-1) y(t-i) - y(t-i) \bar{y}(t-i-1) - \bar{y}(t-i) y(t-i-1) + \bar{y}(t-i) \bar{y}(t-i-1) \}$
 $= N \bar{y}(t-i) \sum_{i=1}^N \{ y(t-i-1) - \bar{y}(t-i-1) \}^2 - N \bar{y}(t-i-1) \sum_{i=1}^N \{ y(t-i-1) - \bar{y}(t-i-1) \} (y(t-i) - \bar{y}(t-i)) \quad \dots \text{(付式14-4)}$
 $B = -\sum_{i=2}^{N+1} y(t-i) \sum_{i=1}^N y(t-i) + N \sum_{i=1}^N y(t-i-1) y(t-i) = -\sum_{i=1}^N y(t-i-1) \sum_{i=1}^N y(t-i) + N \sum_{i=1}^N y(t-i-1) y(t-i) = -N \bar{y}(t-i) \sum_{i=1}^N y(t-i-1) + N \sum_{i=1}^N y(t-i-1) y(t-i)$
 $= -N \sum_{i=1}^N \bar{y}(t-i) y(t-i-1) - N \bar{y}(t-i) N \bar{y}(t-i-1) + N \bar{y}(t-i) N \bar{y}(t-i-1) + N \sum_{i=1}^N y(t-i-1) y(t-i)$
 $= -N \sum_{i=1}^N \bar{y}(t-i) y(t-i-1) - \sum_{i=1}^N y(t-i) N \bar{y}(t-i-1) + N \sum_{i=1}^N \bar{y}(t-i) \bar{y}(t-i-1) + N \sum_{i=1}^N y(t-i-1) y(t-i) = N \sum_{i=1}^N \{ \bar{y}(t-i) y(t-i-1) + y(t-i) \bar{y}(t-i-1) - \bar{y}(t-i) \bar{y}(t-i-1) - y(t-i-1) y(t-i) \}$
 $= N \sum_{i=1}^N \{ (y(t-i) - \bar{y}(t-i)) (y(t-i-1) - \bar{y}(t-i-1)) \} \quad \dots \text{(付式14-5)}$

ついで、(付式14-3)の分母の項 C を求める。

$C = N \sum_{i=2}^{N+1} y(t-i)^2 - (\sum_{i=2}^{N+1} y(t-i))^2 = N \sum_{i=1}^N y(t-i-1)^2 - (\sum_{i=1}^N y(t-i-1))^2 = N \sum_{i=1}^N y(t-i-1)^2 - (\sum_{i=1}^N y(t-i-1))^2 - (\sum_{i=1}^N y(t-i-1))^2 + (\sum_{i=1}^N y(t-i-1))^2$
 $= N \sum_{i=1}^N y(t-i-1)^2 - 2 N \sum_{i=1}^N \bar{y}(t-i-1) y(t-i-1) + (\sum_{i=1}^N \bar{y}(t-i-1))^2 = N \sum_{i=1}^N y(t-i-1)^2 - 2 N \sum_{i=1}^N \bar{y}(t-i-1) y(t-i-1) + N \sum_{i=1}^N \bar{y}(t-i-1)^2$
 $= N \sum_{i=1}^N \{ y(t-i-1)^2 - 2 \bar{y}(t-i-1) y(t-i-1) + \bar{y}(t-i-1)^2 \} = N \sum_{i=1}^N \{ y(t-i-1) - \bar{y}(t-i-1) \}^2 \quad \dots \text{(付式14-6)}$

付録14. 重回帰式の推定式(最小二乗法)を使ったAR(1) 過程のパラメータの推定 (2/2)

(付式14-3)に、(付式14-4)、(付式14-5)、(付式14-6)を代入し、最後に、 $y(t-i)$, $i = 1 \sim N$ を $y(t)$, $t = 1 \sim T$ とおけば、

$$\begin{aligned} \hat{\varphi} = \begin{bmatrix} \hat{c} \\ \hat{\varphi}_1 \end{bmatrix} &= (Y^T Y)^{-1} Y^T y = \frac{1}{N \sum_{i=1}^N \{y(t-i-1) - \bar{y}(t-i-1)\}^2} \begin{bmatrix} N \bar{y}(t-i) \sum_{i=1}^N \{y(t-i-1) - \bar{y}(t-i-1)\}^2 - N \bar{y}(t-i-1) \sum_{i=1}^N \{y(t-i-1) - \bar{y}(t-i-1)\} (y(t-i) - \bar{y}(t-i)) \\ N \sum_{i=1}^N \{(y(t-i) - \bar{y}(t-i)) (t(t-i-1) - \bar{y}(t-i-1))\} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\bar{y}(t-i) - \bar{y}(t-i-1) \sum_{i=1}^N \{(y(t-i) - \bar{y}(t-i)) (y(t-i-1) - \bar{y}(t-i-1))\}}{\sum_{i=1}^N \{y(t-i-1) - \bar{y}(t-i-1)\}^2} \\ \frac{\sum_{i=1}^N \{(y(t-i) - \bar{y}(t-i)) (y(t-i-1) - \bar{y}(t-i-1))\}}{\sum_{i=1}^N \{y(t-i-1) - \bar{y}(t-i-1)\}^2} \end{bmatrix} = \begin{bmatrix} \bar{y}(t) - \bar{y}(t-1) \frac{\sum_{i=1}^T \{(y(t) - \bar{y}(t)) (y(t-1) - \bar{y}(t-1))\}}{\sum_{i=1}^T \{y(t-1) - \bar{y}(t-1)\}^2} \\ \frac{\sum_{i=1}^T \{(y(t) - \bar{y}(t)) (y(t-1) - \bar{y}(t-1))\}}{\sum_{i=1}^T \{y(t-1) - \bar{y}(t-1)\}^2} \end{bmatrix} \quad \dots \text{ (付式14-7)} \end{aligned}$$

従って、最小二乗法で求めた AR(1) モデルのパラメータの推定値は、次のように表すことができる。単回帰式のパラメータ推定の式になっている。

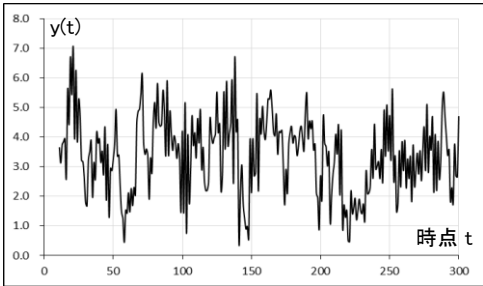
$$\hat{c} = \bar{y}(t) - \hat{\varphi}_1 \bar{y}(t-1) \quad \dots \text{ (付式14-8)}$$

$$\hat{\varphi}_1 = \frac{\sum_{i=1}^T \{(y(t) - \bar{y}(t)) (y(t-1) - \bar{y}(t-1))\}}{\sum_{i=1}^T \{y(t-1) - \bar{y}(t-1)\}^2} \quad \dots \text{ (付式14-9)}$$

付録15-1. AR(2) 過程のパラメータ推定 計算例(ユール・ウォーカー法:レヴィンソン・ダービンのアルゴリズム)

レヴィンソン・ダービンのアルゴリズム(付録13)を使用して、(付式15-1)、(付図15-1)のAR(2)過程($\mu = 3.392, \sigma^2 = 1.725$)のパラメータを求める。合わせて、偏自己相関係数を求める。

$y(t) = 1.0 + 0.2 y(t-1) + 0.5 y(t-2) + \varepsilon(t) \quad \cdots$ (付式15-1)
ただし、 $\varepsilon(t)$ は(図3-2)に示す正規ホワイトノイズ W.N. (0, 1)である。
なお、この過程の特性方程式は、 $\lambda^2 - 0.2 \lambda - 0.5 = 0$ で、 $|\lambda_1| = |0.814| < 1, |\lambda_2| = |-0.614| < 1$ この過程は定常である。
(付図15-1)のなかから 200 個の時系列データを使い、Excel の COVAR 関数で自己共分散を求めた(付表15-1)。
その上で、レヴィンソン・ダービンのアルゴリズム(付録13)を使用して、 $p = 2$ のもと、AR(2)過程のパラメータを求めた(下記)。
また、自己相関係数と偏自己相関係数 $\phi_m(m)$ (赤色)を求め(付表15-1)、コレログラムを作成した(付図15-2, 付図15-3)。



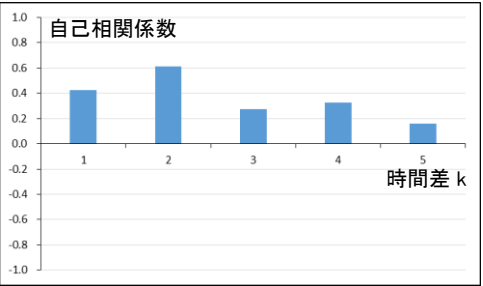
(付図15-1)AR(2)過程の時系列データ

- 1) $p = 0$ のとき
 $\gamma_0 = 1.725$
 $\sigma^2(0) = \gamma_0 = 1.725$
- 2) $p = 1$ のとき
 $\gamma_1 = 0.729$
 $\phi_1(1) = \gamma_1 / \gamma_0 = 0.729 / 1.725 = 0.423$
 $\sigma^2(1) = \sigma^2(0) (1 - \phi_1^2(1)) = 1.725 * (1 - 0.423^2) = 1.417$
- 3) $p = 2$ のとき
 $\gamma_2 = 1.056$
 $\phi_2(2) = (\gamma_2 - \phi_1(1) \gamma_1) / \sigma^2(1) = (1.056 - 0.423 * 0.729) / 1.417 = 0.527$
 $\phi_1(2) = \phi_1(1) - \phi_2(2) \phi_1(1) = 0.423 - 0.527 * 0.423 = 0.200$
 $\sigma^2(2) = \sigma^2(1) (1 - \phi_2^2(2)) = 1.417 * (1 - 0.527^2) = 1.023$
- 4) $p = 3$ のとき
 $\gamma_3 = 0.475$
 $\phi_3(3) = (\gamma_3 - (\phi_1(2) \gamma_2 + \phi_2(2) \gamma_1) / \sigma^2(2))$
 $= (0.475 - (0.200 * 1.056 + 0.527 * 0.729) / 1.023 = -0.118$
 $\phi_1(3) = \phi_1(2) - \phi_3(3) \phi_2(2) = 0.200 - (-0.118) * 0.527 = 0.262$
 $\phi_2(3) = \phi_2(2) - \phi_3(3) \phi_1(2) = 0.527 - (-0.118) * 0.200 = 0.551$
 $\sigma^2(3) = \sigma^2(2) (1 - \phi_3^2(3)) = 1.023 * (1 - (-0.118)^2) = 1.009$
- 5) $p = 4$ のとき
 $\gamma_4 = 0.562$
 $\phi_4(4) = (\gamma_4 - (\phi_1(3) \gamma_3 + \phi_2(3) \gamma_2 + \phi_3(3) \gamma_1) / \sigma^2(3))$
 $= (0.562 - (0.262 * 0.475 + 0.551 * 1.056 + (-0.118) * 0.729) / 1.009 = -0.058$
 $\phi_1(4) = \phi_1(3) - \phi_4(4) \phi_3(3) = 0.262 - (-0.058) * (-0.118) = 0.255$
 $\phi_2(4) = \phi_2(3) - \phi_4(4) \phi_2(3) = 0.551 - (-0.058) * 0.551 = 0.583$
 $\phi_3(4) = \phi_3(3) - \phi_4(4) \phi_1(3) = -0.118 - (-0.058) * 0.262 = -0.103$
 $\sigma^2(4) = \sigma^2(3) (1 - \phi_4^2(4)) = 1.009 * (1 - (-0.058)^2) = 1.005$
- 6) $p = 5$ のとき
 $\gamma_5 = 0.273$
 $\phi_5(5) = (\gamma_5 - (\phi_1(4) \gamma_4 + \phi_2(4) \gamma_3 + \phi_3(4) \gamma_2 + \phi_4(4) \gamma_1) / \sigma^2(4))$
 $= (0.273 - (0.255 * 0.562 + 0.583 * 0.475 + (-0.103) * 1.056 + (-0.058) * 0.729) / 1.005 = 0.003$

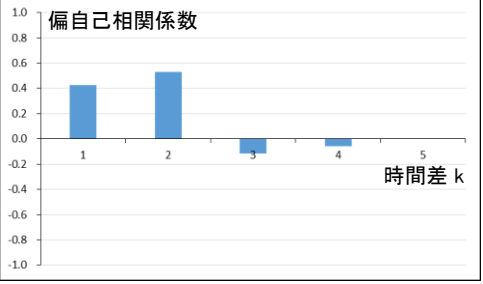
従って、AR(2)過程の推定したパラメータは、
 $\phi_1(2) = 0.200, \phi_2(2) = 0.527$, そして、(式9-2)より、 $c = \mu (1 - \phi_1(2) - \phi_2(2)) = 3.302 * (1 - 0.200 - 0.527) = 0.925$
なお、(付図15-3)から、AR(2)過程の偏自己相関係数のコレログラムは、時間差 = 3 以降はほぼ 0 になることが分かる。
このことから、この AR 過程の次数が $p = 2$ であると推察できる。
AR(2)過程では、時間差 = 3 以上離れたデータは相関がないので当然であるが、一方、(付図15-2)から分かるように自己相関係数は時間差が 3 以上であっても値を持っている。これは時間差 3 以上のデータの影響が残りに残っていることを示している。

(付表15-1)自己相関係数と偏自己相関係数

自己共分散	γ_0	γ_1	γ_2	γ_3	γ_4	γ_5
	1.725	0.729	1.056	0.475	0.562	0.273
自己相関係数	ρ_0	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5
	1.0	0.423	0.612	0.276	0.326	0.158
偏自己相関係数		$\phi_1(1)$	$\phi_2(2)$	$\phi_3(3)$	$\phi_4(4)$	$\phi_5(5)$
		0.423	0.527	-0.118	-0.058	0.003



(付図15-2)自己相関係数のコレログラム



(付図15-3)偏自己相関係数のコレログラム

付録15-2. AR(2) 過程のパラメータ推定（最小二乗法を使った計算例）

(付式15-1)、(付図15-1)に示すAR(2) 過程のパラメータを、付録12に示すように、AR(p) 過程を重回帰式に見立てて(付式12-1)、この式の偏回帰係数、つまり、AR(2) 過程のパラメータ ϕ を(付式12-2)を使って求める。正規ホワイトノイズは、(図3-2)を使う。なお、ここでは、 $p = 4$ において重回帰式分析を行っている。

$y = \phi Y + \varepsilon \quad \dots \quad$ (付式12-1) $\hat{\phi} = (Y^T Y)^{-1} Y^T y \quad \dots \quad$ (付式12-2)

ただし、 $y = \begin{bmatrix} y(t) \\ y(t-1) \\ y(t-2) \\ y(t-3) \end{bmatrix}$ 、 $\phi = \begin{bmatrix} c \\ \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 1.0 \\ 0.2 \\ 0.5 \\ 0.0 \end{bmatrix}$ 、 $Y = \begin{bmatrix} 1 & y(t-1) \\ 1 & y(t-2) \\ 1 & y(t-3) \\ 1 & y(t-4) \end{bmatrix}$ 、 $\varepsilon = \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \varepsilon(t-2) \\ \varepsilon(t-3) \end{bmatrix}$

(付図15-1)の時系列データを(付表15-2)のように並べて、Excel の回帰分析機能を使用して、AR(2) 過程のパラメータを求める(別途資料「重回帰分析の概要」を参照)。結果は(付図15-4)。なお、VIF(分散拡大要因、Variance Inflation Factor)は下記の通り、いずれも 10 より小さな値で、多重共線性のリスクはないと判断できる。

$VIF_1 = 1.678 < 10$ 、 $VIF_2 = 1.777 < 10$ 、 $VIF_3 = 1.782 < 10$ 、 $VIF_4 = 1.692 < 10$

相関行列 $R = \begin{bmatrix} 1.0 & 0.418 & 0.602 & 0.266 \\ 0.418 & 1.0 & 0.421 & 0.604 \\ 0.602 & 0.421 & 1.0 & 0.425 \\ 0.266 & 0.604 & 0.425 & 1.0 \end{bmatrix}$ 、 $R^{-1} = \begin{bmatrix} 1.678 & -0.443 & -0.911 & 0.207 \\ -0.443 & 1.777 & -0.093 & -0.916 \\ -0.911 & -0.093 & 1.782 & -0.458 \\ 0.207 & -0.916 & -0.458 & 1.692 \end{bmatrix}$

重回帰分析の結果は、 c (切片)、 X 値1 (ϕ_1)、 X 値2 (ϕ_2) の値の p 値が 0.05 より小さく有意であることを示している(但し、自由度調整済み決定係数(補正 R^2)は十分に満足できるものではない)。

(付録15-1)ユール・ウォーカー法 と (付録15-2)最小二乗法 の結果を、(付表15-3)にまとめておく。
いずれの方法でも、まずまずの推定結果が得られている。

(付表15-3)推定パラメータの比較

パラメータ		パラメータの推定値		
			ユール・ウォーカー法	最小二乗法
c	1.0	$\hat{c}$	0.925	1.080
ϕ_1	0.2	$\hat{\phi}_1$	0.200	0.255
ϕ_2	0.5	$\hat{\phi}_2$	0.527	0.579

(付表15-2)重回帰分析のための時系列データの並び替え

	目的変数	説明変数			
	y(t)	y(t - 1)	y(t - 2)	y(t - 3)	y(t - 4)
1	3.138	3.654	2.913	3.960	3.632
2	3.750	3.138	3.654	2.913	3.960
3	3.847	3.750	3.138	3.654	2.913
4	3.951	3.847	3.750	3.138	3.654
5	2.607	2.607	3.847	3.750	3.138
6	5.618	2.607	2.607	3.847	3.750
7	4.432	5.618	2.607	2.607	3.847
8	6.706	4.432	5.618	2.607	2.607
...	...	...	...	...	...

概要									
回帰統計									
重相関 R	0.646269009								
重決定 R2	0.417663632								
補正 R2	0.409193285								
標準誤差	1.011906392								
観測数	280								
分散分析表									
	自由度	変動	分散	測された分散	有意 F				
回帰	4	201.9599531	50.48998827	49.30891541	3.006E-31				
残差	275	281.5869434	1.023952522						
合計	279	483.5468965							
	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%	
切片	1.079852419	0.212410648	5.083796068	6.84929E-07	0.6616949	1.49800994	0.6616949	1.4980099	
X 値 1	0.254596287	0.060278548	4.223663257	3.27122E-05	0.1359303	0.37326232	0.1359303	0.3732623	
X 値 2	0.579157841	0.061863815	9.361819066	2.9188E-18	0.457371	0.70094467	0.457371	0.7009447	
X 値 3	-0.10141029	0.062029119	-1.63488207	0.103217891	-0.223523	0.02070196	-0.223523	0.020702	
X 値 4	-0.04977568	0.060531142	-0.82231521	0.411608944	-0.168939	0.06938761	-0.168939	0.0693876	

(付図15-4)重回帰分析の結果

付録16. AR(p) 過程のモデル次数の推定（AIC を使用）

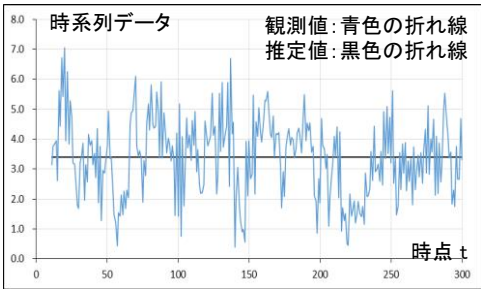
AR(p) 過程のモデル次数は、偏自己相関係数が $p + 1$ 次以降は 0 になることを利用して推定することができるが、ここでは、時系列過程が AR(p) 過程であることを前提に AIC を使用して AR(p) 過程の次数を推定する。

前掲の(付図15-1)は、(付式15-1)に示す AR(2) 過程の時系列データである。

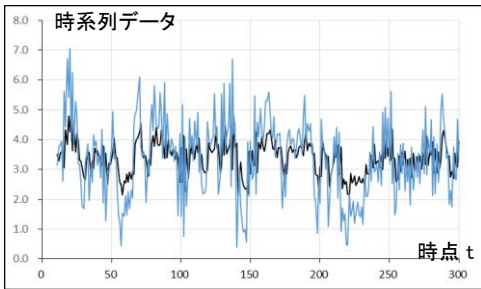
まず、(付図15-1)の時系列データを対象に、モデル次数 p ごとの AR(p) 過程のパラメータを(付録15-2)の手順で推定し、(式10-11)を使用してそれぞれの AIC を求める。
推定した次数が、 $p = 2$ になることを確認する。

推定結果を(付表16-1)にまとめる。
この表から、この AR(p) 過程のモデル次数 p は、AIC の値が最小になる $p = 2$ であると推定できる。

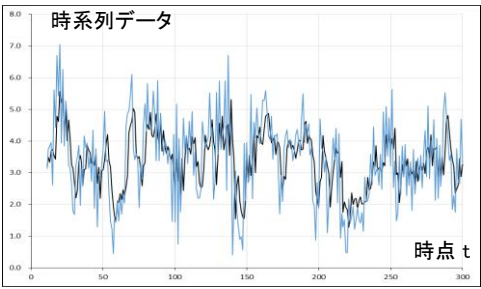
なお、次数が $p = 0$ 、 $p = 1$ 、 $p = 2$ 、 $p = 5$ のときの観測値と推定値の比較を(付図16-1)、(付図16-2)、(付図16-3)、(付図16-4)に示す。



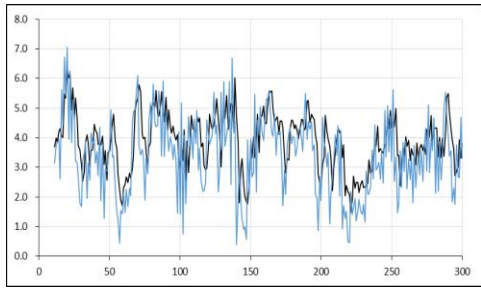
(付図16-1) 次数 $p = 0$ のときの観測値と推定値



(付図16-2) 次数 $p = 1$ のときの観測値と推定値



(付図16-3) 次数 $p = 2$ のときの観測値と推定値



(付図16-4) 次数 $p = 5$ のときの観測値と推定値

(付表16-1) AIC の計算結果

(注1) 有意な推定値(p 値 < 0.05)は赤字で示す。

次数 p 説明変数の数	残差 平方和	データ の数	AIC	自由度 調整済みR2	F検定 p 値	c の推定値 (p 値)	ϕ_1 の推定値 (p 値)	ϕ_2 の推定値 (p 値)	ϕ_3 の推定値 (p 値)	ϕ_4 の推定値 (p 値)	ϕ_5 の推定値 (p 値)
0	483.5	280	951.6								
1	397.1	280	898.5	0.176	0.00	1.959 (0.00)	0.423 (0.00)				
2	286.0	280	808.5	0.404	0.00	0.921 (0.00)	0.200 (0.00)	0.530 (0.00)			
3	330.0	280	850.6	0.410	0.00	1.027 (0.00)	0.260 (0.00)	0.552 (0.00)	- 0.114 (0.059)		
4	362.7	280	879.1	0.409	0.00	1.080 (0.00)	0.255 (0.00)	0.579 (0.00)	- 0.101 (0.103)	- 0.050 (0.412)	
5	360.7	280	879.5	0.407	0.00	1.066 (0.00)	0.255 (0.00)	0.580 (0.00)	- 0.109 (0.129)	- 0.053 (0.397)	0.013 (0.829)

付録17. AR(1) 過程の点予測

(付式17-1)に示す AR(1) 過程の点予測を行う。

$y(t) = c + \phi_1 y(t-1) + \varepsilon(t)$ 但し、 $\varepsilon(t) \sim \text{i.i.d. } N(0, \sigma^2)$... (付式17-1)

また、(式11-2)と(式11-3)の条件を前提とする。

- ① $\tau \leq t$ のとき、 $E[y(\tau) | \Omega(t)] = y(\tau)$... (式11-2) 時点 t 以前の過去の $y(t)$ の条件付き期待値は、 $y(t)$ である。
- ② $k > 0$ のとき、 $E[\varepsilon(t+k) | \Omega(t)] = 0$... (式11-3) $\varepsilon(t)$ は正規ホワイトノイズであり、時点 t より将来の $\varepsilon(t)$ の条件付き期待値は、0 である。

(1) 1期先の時点 $t+1$ での $y(t+1)$ の最適予測

(付式17-1)より、 $y(t+1) = c + \phi_1 y(t) + \varepsilon(t+1)$

一方、(式11-2)より、 $t+1$ 時点以前の t 時点の $y(t)$ の条件付き期待値は $y(t)$ であること、(式11-3)より、将来の $\varepsilon(t+1) = 0$ であることから、 $y(t+1)$ の最適予測値は、

$\hat{y}(t+1 | t) = c + \phi_1 y(t)$

従って、1期先の予測誤差は、 $\hat{\varepsilon}(t+1 | t) = y(t+1) - \hat{y}(t+1 | t) = (c + \phi_1 y(t) + \varepsilon(t+1)) - (c + \phi_1 y(t)) = \varepsilon(t+1)$

ここで、定常過程であるので時点 t に依らず、 $V[\varepsilon(t)] = E[\varepsilon(t)^2] = \sigma^2$ である。 よって、 $MSE(\hat{y}(t+1 | t)) = E[\hat{\varepsilon}(t+1 | t)^2] = E[\varepsilon(t+1)^2] = \sigma^2$

(2) 2期先の時点 $t+2$ での $y(t+2)$ の最適予測

(付式17-1)より、 $y(t+2) = c + \phi_1 y(t+1) + \varepsilon(t+2) = c + \phi_1 (c + \phi_1 y(t) + \varepsilon(t+1)) + \varepsilon(t+2) = (1 + \phi_1) c + \phi_1^2 y(t) + \varepsilon(t+2) + \phi_1 \varepsilon(t+1)$

一方、(式11-2)の条件と、(式12-3)の条件である将来の $\varepsilon(t+1) = \varepsilon(t+2) = 0$ であることから、 $y(t+2)$ の最適予測値は、

$\hat{y}(t+2 | t) = (1 + \phi_1) c + \phi_1^2 y(t)$

従って、2期先の予測誤差は、 $\hat{\varepsilon}(t+2 | t) = y(t+2) - \hat{y}(t+2 | t) = ((1 + \phi_1) c + \phi_1^2 y(t) + \varepsilon(t+2) + \phi_1 \varepsilon(t+1)) - ((1 + \phi_1) c + \phi_1^2 y(t)) = \varepsilon(t+2) + \phi_1 \varepsilon(t+1)$

ここで、定常過程であるので時点 t に依らず、 $V[\varepsilon(t)] = E[\varepsilon(t)^2] = \sigma^2$ である。 よって、

$MSE(\hat{y}(t+2 | t)) = E[\hat{\varepsilon}(t+2 | t)^2] = E[(\varepsilon(t+2) + \phi_1 \varepsilon(t+1))^2] = E[\varepsilon(t+2)^2] + 2 E[\varepsilon(t+2) \phi_1 \varepsilon(t+1)] + \phi_1^2 E[\varepsilon(t+1)^2] = \sigma^2 + 0 + \phi_1^2 \sigma^2 = (1 + \phi_1^2) \sigma^2$

(3) 一般に、 h 期先の時点 $t+h$ での $y(t+h)$ の最適予測

(1)と(2)の結果を参考にして、

(付式17-1)より、 $y(t+h) = (1 + \phi_1 + \phi_1^2 + \dots + \phi_1^{h-1}) c + \phi_1^h y(t) + \varepsilon(t+h) + \phi_1 \varepsilon(t+h-1) + \dots + \phi_1^{h-1} \varepsilon(t+1) = c \sum_{k=1}^h \phi_1^{k-1} + \phi_1^h y(t) + \sum_{k=0}^{h-1} \phi_1^k \varepsilon(t+h-k)$

一方、1時点以前の過去の $y(t)$ であることと、将来の $\varepsilon(t+1) = \varepsilon(t+2) = \dots = \varepsilon(t+h) = 0$ であることから、 $y(t+h)$ の最適予測値は、

$\hat{y}(t+h | t) = c \sum_{k=1}^h \phi_1^{k-1} + \phi_1^h y(t)$

この式の右辺の第1項は、初項 $= c \phi_1^{1-1} = c \phi_1^0 = c$ 、公比 $= \phi_1$ の等比数列であるので、等比数列の和の公式(付録A)を使って、

$\hat{y}(t+h | t) = c(1 - \phi_1^h) / (1 - \phi_1) + \phi_1^h y(t)$... (付式17-2)

また、推定値の平均二乗誤差は、

$\hat{\varepsilon}(t+h | t) = \sum_{k=0}^{h-1} \phi_1^k \varepsilon(t+h-k)$ であり、定常過程であるので時点 t に依らず、 $V[\varepsilon(t)] = E[\varepsilon(t)^2] = \sigma^2$ 、 $E[\varepsilon(i) \varepsilon(j)] = 0$ ($i \neq j$ のとき) である。 よって、

$MSE(\hat{y}(t+h | t)) = E[\hat{\varepsilon}(t+h | t)^2] = E[(\sum_{k=0}^{h-1} \phi_1^k \varepsilon(t+h-k))^2] = E[\sum_{k=0}^{h-1} \phi_1^{2k} (\varepsilon(t+h-k))^2] = \sum_{k=0}^{h-1} E[\phi_1^{2k} (\varepsilon(t+h-k))^2] = \sum_{k=0}^{h-1} \phi_1^{2k} E[\varepsilon(t+h-k)^2]$
 $= \sum_{k=0}^{h-1} \phi_1^{2k} V[\varepsilon(t+h-k)] = \sum_{k=0}^{h-1} \phi_1^{2k} \sigma^2 = \sigma^2 \sum_{k=0}^{h-1} \phi_1^{2k}$

これは、初項 $= \sigma^2 \phi_1^0 = \sigma^2$ 、公比 $= \phi_1^2$ の等比数列であるので、等比数列の和の公式(付録A)を使って、

$MSE(\hat{y}(t+h | t)) = \sigma^2 (1 - \phi_1^{2h}) / (1 - \phi_1^2)$... (付式17-3)

(付式17-2)から、 h 期先の最適予測には、現在値の $y(t)$ だけ使用し、過去の $y(t-1)$ 、 $y(t-2)$ 、... は使用していないことが分かる。

また、(式17-2)より、予測期間が長くなると($h \rightarrow \infty$ で)、 $\phi_1 < 1$ (定常の条件)であるから、 $\hat{y}(t+h | t) \rightarrow c(1 - 0) / (1 - \phi_1) + 0 * y(t) = c / (1 - \phi_1)$ となる。

この値は、(式6-2)で示す AR(1) 過程の期待値である。つまり、AR(1) 過程は長期的には期待値(平均)に向かうことを示している。平均回帰的(mean reverting)という。

さらに、(付式17-3)より、 h を大きくしていくと、 $MSE(\hat{y}(t+h | t))$ の値は徐々に大きくなりながら、 $h \rightarrow \infty$ で、 $MSE(\hat{y}(t+h | t)) \rightarrow \sigma^2 (1 - 0) / (1 - \phi_1^2) = \sigma^2 / (1 - \phi_1^2)$ となる。この値は、(式6-3)で示す AR(1) 過程の分散であり((式6-5)で、 $k=1$ のとき $\rho_1 = \phi_1$ である)、 $MSE(\hat{y}(t+h | t))$ の値の上限を示している。

つまり、 MSE の値が、際限なく大きくなることはない。これも、AR(1) 過程が平均回帰的であることを示している。

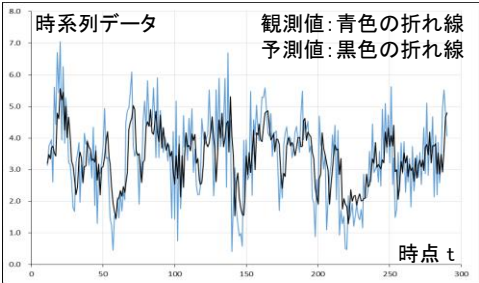
付録18. AR(p) 過程の点予測

(付式15-1)に示す AR(2) 過程を対象にして、1 期先($h = 1$)、2 期先($h = 2$)、 $\dots$ 、5 期先($h = 5$)の予測を行う。予測には(式11-9)を使用する。
なお、AR(2) のパラメータは、(付録15-2)の最小二乗法で求めた推定パラメータを使用する。また、正規ホワイトノイズは、(図3-2)を使う。

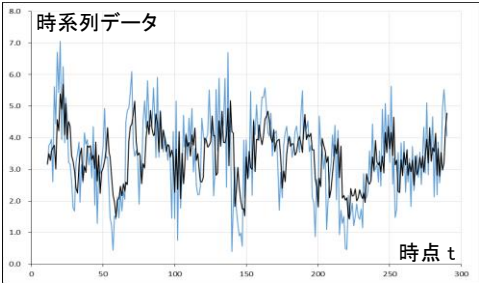
次数 $p = 2$ の予測式は、以下の通りである。

- ① 1 期先($h = 1$)は、 $\hat{y}(t + 1 | t) = \hat{c} + \hat{\phi}_1 y(t | t) + \hat{\phi}_2 y(t - 1 | t)$
- ② 2 期先($h = 2$)は、 $\hat{y}(t + 2 | t) = \hat{c} + \hat{\phi}_1 \hat{y}(t + 1 | t) + \hat{\phi}_2 y(t | t)$
- ③ 3 期先($h = 3$)は、 $\hat{y}(t + 3 | t) = \hat{c} + \hat{\phi}_1 \hat{y}(t + 2 | t) + \hat{\phi}_2 \hat{y}(t + 1 | t)$
- ④ 4 期先($h = 4$)は、 $\hat{y}(t + 4 | t) = \hat{c} + \hat{\phi}_1 \hat{y}(t + 3 | t) + \hat{\phi}_2 \hat{y}(t + 2 | t)$
- ⑤ 5 期先($h = 5$)は、 $\hat{y}(t + 5 | t) = \hat{c} + \hat{\phi}_1 \hat{y}(t + 4 | t) + \hat{\phi}_2 \hat{y}(t + 3 | t)$

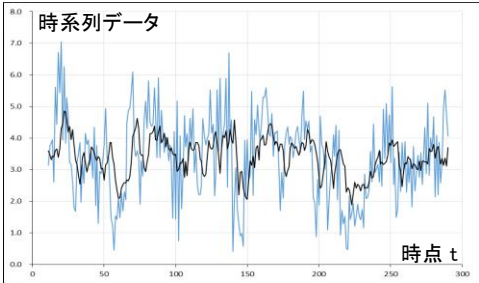
予測した結果を、(付図18-1)、(付図18-2)、(付図21-3)、(付図18-4)、(付図18-5)に示す。
1 期先($h = 1$)、2 期先($h = 2$)の予測はほどほどの結果であるが、予測が「先」になるほど、つまり、 h の値が大きくなるほど、当然のことながら予測精度が落ちる。また、予測時期が遅れていくことが分かる。



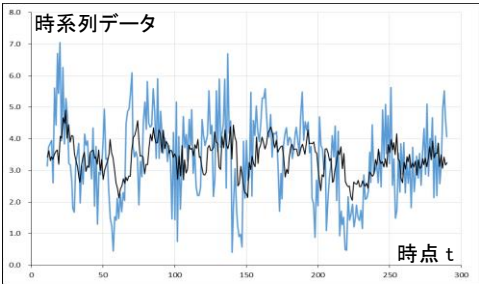
(付図18-1) 1 期先 ($h = 1$) のときの観測値と予測値



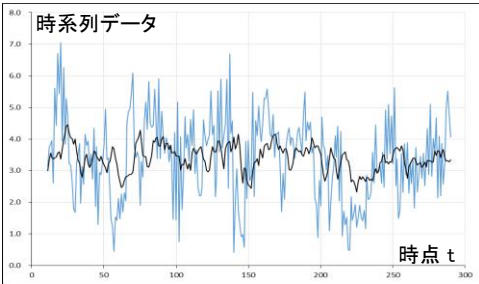
(付図18-2) 2 期先 ($h = 2$) のときの観測値と予測値



(付図18-3) 3 期先 ($h = 3$) のときの観測値と予測値



(付図18-4) 4 期先 ($h = 4$) のときの観測値と予測値



(付図18-5) 5 期先 ($h = 5$) のときの観測値と予測値

付録19. AR(p) 過程の区間予測

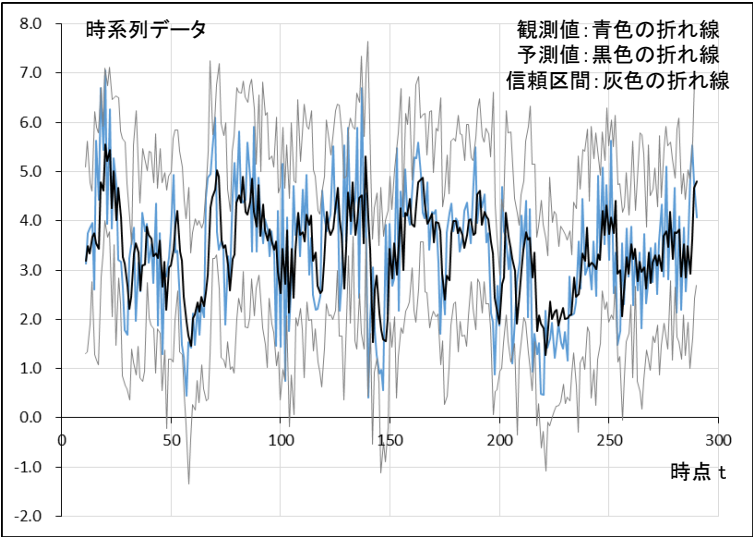
(付式15-1)に示す AR(2) 過程の 1 期先の点予測 $\hat{y}(t+1 | t)$ の区間予測を行う。
データ数 = 280, $p = 2, h = 1$ である。
その中で、 $t = 179$ 時点以前のデータを使用して $t + 1 = 180$ 時点での区間予測を行った計算結果を、一例として示す。

- 1) 1 期先の点予測 $\hat{y}(t+1 | t)$ としては、(付図18-1)の結果を使用する。
 $\hat{y}(180 | t) = 3.990$
- 2) (付式15-1)の $\varepsilon(t)$ として 10 種類の正規ホワイトノイズ $\varepsilon(t, 1), \varepsilon(t, 2), \dots, \varepsilon(t, 3)$ を用意し、
下記に示す10 種類の観測データ $y(t+1 | t, 1), y(t+1 | t, 2), \dots, y(t+1 | t, 10)$ を疑似的に作成する。
1 期先の観測データは、(付表19-1)に示す通りである。
 $y(t+1 | t, 1) = c + \phi_1 y(t) + \phi_2 y(t-1) + \varepsilon(t+1, 1)$
 $y(t+1 | t, 2) = c + \phi_1 y(t) + \phi_2 y(t-1) + \varepsilon(t+1, 2)$
...
 $y(t+1 | t, 10) = c + \phi_1 y(t) + \phi_2 y(t-1) + \varepsilon(t+1, 10)$
- 3) 平均二乗誤 (MSE) を計算する。
 $MSE(\hat{y}(180 | t)) = (1 / 10) \sum_{i=1}^{10} (y(180 | t, i) - \hat{y}(180 | t))^2 = 0.690$
- 4) (式11-11)を使って、95 % 信頼区間を求める。
上側 95% 信頼区間 = $\hat{y}(180 | t) + 1.960 (MSE(\hat{y}(180 | t)))^{1/2} = 3.990 + 1.960 \cdot 0.690^{1/2} = 5.618$
下側 95% 信頼区間 = $\hat{y}(180 | t) - 1.960 (MSE(\hat{y}(180 | t)))^{1/2} = 3.990 - 1.960 \cdot 0.690^{1/2} = 2.361$

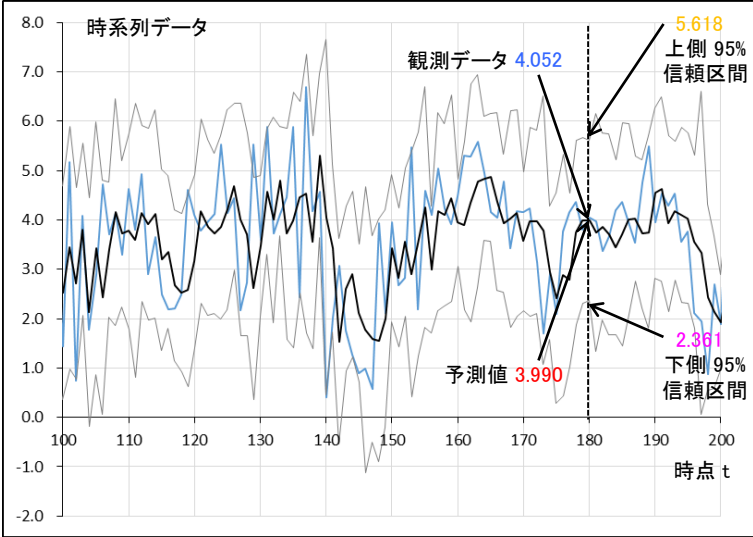
(付表19-1) 観測データ

時点 $t + 1 = 180$	観測データ
$y(t + 1 t, 1)$	4.052
$y(t + 1 t, 2)$	5.466
$y(t + 1 t, 3)$	3.002
$y(t + 1 t, 4)$	4.390
$y(t + 1 t, 5)$	3.149
$y(t + 1 t, 6)$	3.435
$y(t + 1 t, 7)$	3.676
$y(t + 1 t, 8)$	3.632
$y(t + 1 t, 9)$	5.434
$y(t + 1 t, 10)$	4.496

- 5) 信頼区間をグラフで表す。
(付図19-1)は、全時点を対象に、観測データ、予測データ、上側 95% 信頼区間、下側 95% 信頼区間を示す。
(付図19-2)は、(付図19-1)のグラフから、時点 100 ~ 200 のデータを抽出して拡大表示したものである。合わせて、 $t + 1 = 180$ 時点の計算結果も示しておく。



(付図19-1) 予測の 95 % 信頼区間
(注)観測データは、10 個のうち、1 個だけを表示



(付図19-2)
(付図19-1)の時点 100 ~ 200 を抽出

付録20. MA(q) 過程の点予測

下記に示す、MA(2) 過程 (q = 2) の点予測を行う。
計算プロセスを(付表20-1)に示す。最適予測結果は、(付図20-1)、(付図20-2)、(付図20-3)に示す。

$$y(t) = \varepsilon(t) + 0.8 \varepsilon(t-1) + 0.4 \varepsilon(t-2)$$
$$y(1) = \varepsilon(1) + 0.8 \varepsilon(0) + 0.4 \varepsilon(-1) = -1.099 + 0.8 \times 0.014 + 0.4 \times (-1.326) = -1.619$$
$$y(2) = \varepsilon(2) + 0.8 \varepsilon(1) + 0.4 \varepsilon(0) = 0.087 + 0.8 \times (-1.099) + 0.4 \times 0.014 = -0.787$$
$$\dots$$

(式12-5)を使用して、

$$\hat{\varepsilon}(1) = y(1) - \mu = -1.619 - 0.051 = -1.670$$
$$\hat{\varepsilon}(2) = y(2) - \mu - \theta_1 \hat{\varepsilon}(1) = -0.787 - 0.051 - 0.8 \times (-1.670) = 0.498$$
$$\hat{\varepsilon}(3) = y(3) - \mu - \theta_1 \hat{\varepsilon}(2) - \theta_2 \hat{\varepsilon}(1) = 0.115 - 0.051 - 0.8 \times 0.498 - 0.4 \times (-1.670) = 0.334$$
$$\hat{\varepsilon}(4) = y(4) - \mu - \theta_1 \hat{\varepsilon}(3) - \theta_2 \hat{\varepsilon}(2) = 1.419 - 0.051 - 0.8 \times 0.334 - 0.4 \times 0.498 = 0.902$$
$$\dots$$

(式12-5)から、
h = 1 のとき、 $\hat{y}(t+1 | t) = \mu + \theta_1 \hat{\varepsilon}(t) + \theta_2 \hat{\varepsilon}(t-1)$

$$\rightarrow \hat{y}(0+1 | t) = \mu + \theta_1 \hat{\varepsilon}(0) + \theta_2 \hat{\varepsilon}(-1) = 0.051 + 0.8 \times 0 + 0.4 \times 0 = 0.051$$
$$\hat{y}(1+1 | t) = \mu + \theta_1 \hat{\varepsilon}(1) + \theta_2 \hat{\varepsilon}(0) = 0.051 + 0.8 \times (-1.670) + 0.4 \times 0 = -1.285$$
$$\hat{y}(2+1 | t) = \mu + \theta_1 \hat{\varepsilon}(2) + \theta_2 \hat{\varepsilon}(1) = 0.051 + 0.8 \times 0.498 + 0.4 \times (-1.670) = -0.219$$
$$\dots$$

h = 2 のとき、 $\hat{y}(t+2 | t) = \mu + \theta_2 \hat{\varepsilon}(t)$

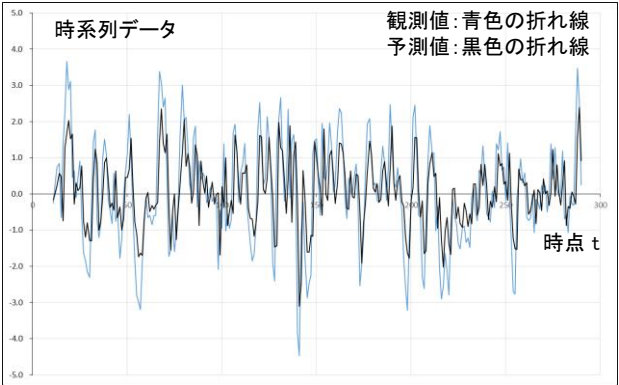
$$\rightarrow \hat{y}(0+2 | t) = \mu + \theta_2 \hat{\varepsilon}(0) = 0.051 + 0.4 \times 0 = 0.051$$
$$\hat{y}(1+2 | t) = \mu + \theta_2 \hat{\varepsilon}(1) = 0.051 + 0.4 \times (-1.670) = -0.617$$
$$\hat{y}(2+2 | t) = \mu + \theta_2 \hat{\varepsilon}(2) = 0.051 + 0.4 \times 0.498 = 0.250$$
$$\dots$$

h = 3 のとき、 $\hat{y}(t+3 | t) = \mu \rightarrow \hat{y}(0+3 | t) = \hat{y}(1+3 | t) = \hat{y}(2+3 | t) = 0.051$ (これは平均値)
h = 4, 5, ... は、h = 3 と同じ

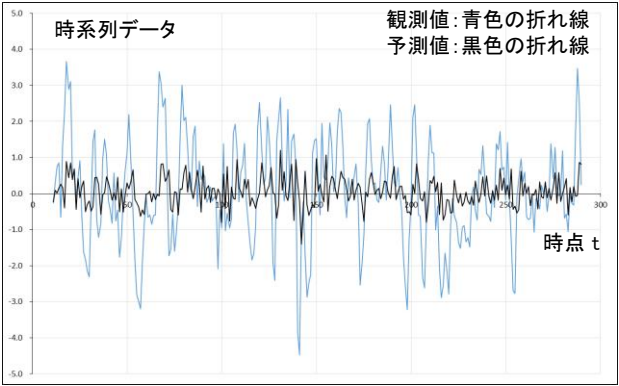
(付表20-1) 計算プロセス

t	$\varepsilon(t)$	y(t)	$\hat{\varepsilon}(t)$	$\hat{y}(t+1)$	$\hat{y}(t+2)$	$\hat{y}(t+3)$
-1	-1.326					
0	0.014					
1	-1.099	-1.619	-1.670	0.051		
2	0.087	-0.787	0.498	-1.285	0.051	
3	0.485	0.115	0.334	-0.219	-0.617	0.051
4	0.996	1.419	0.902	0.517	0.250	0.051
5	1.204	2.195	1.289	0.906	0.184	0.051
...	...	...	...	...	...	...

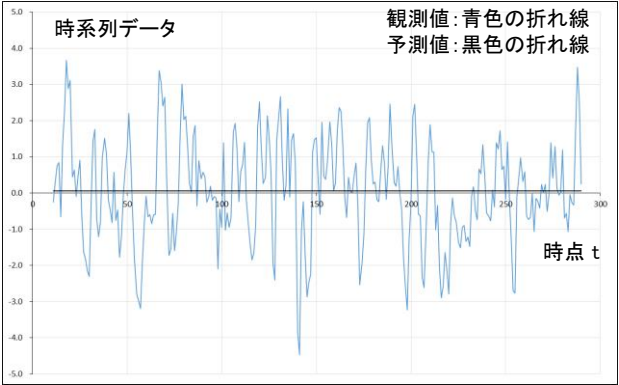
t = 11 ~ 290 の 280 個のデータを使用 y(t) の平均 = 0.051 , y(t) の分散 = 2.046



(付図20-1) 1 期先の予測



(付図20-2) 2 期先の予測



(付図20-3) 3 期以降の予測

付録21. グレンジャー因果性の検定(計算例)

- (手順1) (付式21-1)と(付式21-2)に示す、VAR(2)モデルに従う時系列データを用意する(付図21-1)。
- VAR (2) 過程のモデル $y_1(t) = c_1 + \phi_{11}(1)y_1(t-1) + \phi_{12}(2)y_2(t-1) + \varepsilon_1(t) = 0.2 + 0.6*y_1(t-1) + 0.4*y_2(t-1) + \varepsilon_1(t) \quad \dots$ (付式21-1)
- $y_2(t) = c_2 + \phi_{21}(1)y_1(t-1) + \phi_{22}(2)y_2(t-1) + \varepsilon_2(t) = 0.8 \quad + 0.5*y_2(t-1) + \varepsilon_2(t) \quad \dots$ (付式21-2)
- (手順2) (付式21-1)の時系列データを用いて、モデル(付式21-1)の係数をExcelの回帰分析ツール(最小二乗法)を使用して推定する(モデル(1)とする)。
- 推定結果(付図21-4)は、 $c_1 = 0.2257$, $\phi_{11}(1) = 0.6264$, $\phi_{12}(2) = 0.3614$
- (手順3) (付式21-1)の時系列データを用いて、(付式21-1)で $y_1(t)$ へ影響を及ぼす $y_2(t-1)$ の係数 $\phi_{12}(2)$ を0としたモデル(付式21-3)の係数をExcelの回帰分析ツール(最小二乗法)を使用して推定する(モデル(0)とする)。検定に使用する帰無仮説は、 $\phi_{12}(2) = 0$ になる。
- $y_1(t) = c_1 + \phi_{11}(1)y_1(t-1) + \phi_{12}(2)y_2(t-1) + \varepsilon_1(t) = 0.2 + 0.6*y_1(t-1) + 0*y_2(t-1) + \varepsilon_1(t) = 0.2 + 0.6*y_1(t-1) + \varepsilon_1(t) \quad \dots$ (付式21-3)
- 推定結果(付図21-5)は、 $c_1 = 0.6217$, $\phi_{11}(1) = 0.6954$
- (手順4) モデル(1)を使用して求めた推定値と(付式21-1)の時系列データの対比を(付図21-2)に示す。偏差平方和 SSR_1 を求める。 $SSR_1 = 30.01$
- (手順5) モデル(0)を使用して求めた推定値と(付式21-3)の時系列データの対比を(付図21-3)に示す。偏差平方和 SSR_0 を求める。 $SSR_0 = 35.66$
- (手順6) SSR_0 と SSR_1 を使って、(式14-4)からグレンジャーの因果性検定のための統計量 rU を求める。検定データ数 $T = 30$, 帰無仮説のパラメータ数 $r = 1$, 変量数 $k = 2$
- $rU = (SSR_0 - SSR_1) / \{SSR_1 / (T - k - 1)\} = (35.66 - 30.01) / \{30.01 / (30 - 2*1 - 1)\} = 5.082$
- (手順7) 検定統計量 U を使って、自由度1のカイ二乗検定を行う。 $U = 5.082 > \chi^2_{0.05}(1) = 3.841$
- 従って、帰無仮説 $\phi_{12}(2) = 0$ は棄却され、 $\phi_{12}(2) \neq 0$ が採択される。このケースでは、 $y_2(t-1)$ は $y_1(t)$ に影響を及ぼす、グレンジャー因果性がある、ということが分かる。

回帰統計					
重相関 R	0.749258				
重決定 R2	0.561387				
補正 R2	0.55822				
標準誤差	1.015345				
観測数	280				

分散分析表					
	自由度	変動	分散	割られた分	有意 F
回帰	2	365.5007	182.7504	177.2683	2.68E-50
残差	277	285.5663	1.030925		
合計	279	651.0671			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.225694	0.116077	1.94435	0.052865	-0.00281	0.454199	-0.00281	0.454199
X 値 1	0.626435	0.040993	15.28155	1.18E-38	0.545737	0.707132	0.545737	0.707132
X 値 2	0.361381	0.051577	7.006605	1.86E-11	0.259848	0.462914	0.259848	0.462914

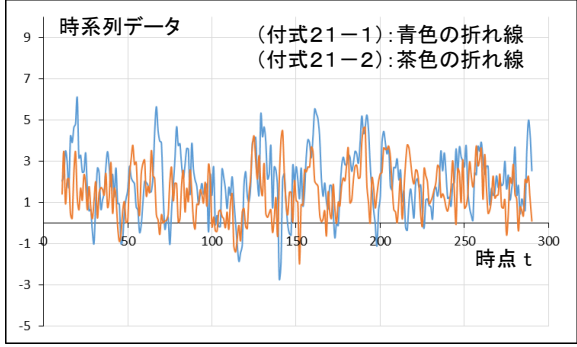
(付図21-4)モデル(1)の係数推定

回帰統計					
重相関 R	0.695451				
重決定 R2	0.483652				
補正 R2	0.481795				
標準誤差	1.099668				
観測数	280				

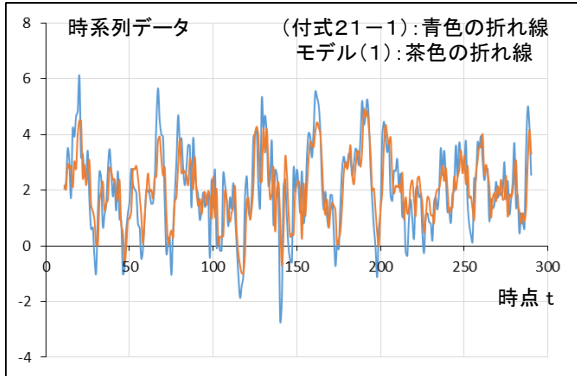
分散分析表					
	自由度	変動	分散	割られた分	有意 F
回帰	1	314.89	314.89	260.3968	8.6E-42
残差	278	336.177	1.20927		
合計	279	651.0671			

	係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
切片	0.62166	0.109811	5.661166	3.74E-08	0.405493	0.837827	0.405493	0.837827
X 値 1	0.695444	0.043097	16.13681	8.6E-42	0.610607	0.780282	0.610607	0.780282

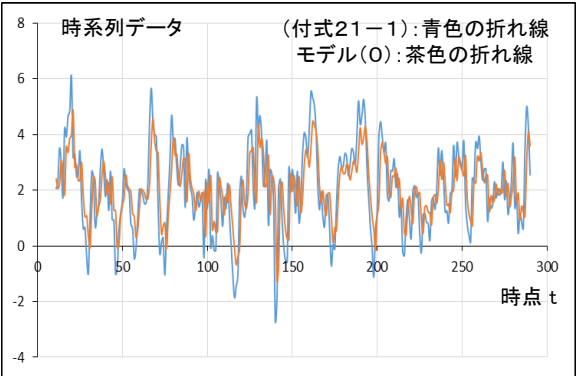
(付図21-5)モデル(0)の係数推定



(付図21-1)時系列データ



(付図21-2)モデル(1)の推定時系列データ



(付図21-3)モデル(0)の推定時系列データ

付録22. 非直交化インパルス応答関数(計算例)

(付式22-1)と(付式22-2)で表される VAR(1) 過程の非直交化インパルス応答関数の計算例を示す。

$$\left. \begin{aligned} y_1(t) &= -0.8 + 0.6 y_1(t-1) + 0.9 y_2(t-1) - 0.7 y_3(t-1) + \varepsilon_1(t) \\ y_2(t) &= 0.9 + 0.2 y_1(t-1) + 0.6 y_2(t-1) + 0.2 y_3(t-1) + \varepsilon_2(t) \\ y_3(t) &= 0.6 + 0.3 y_1(t-1) + 0.2 y_2(t-1) + 0.4 y_3(t-1) + \varepsilon_3(t) \end{aligned} \right\} \quad \cdots \text{ (付式22-1)}$$

ここで、 $\varepsilon(t)$ の分散共分散行列 $\Sigma = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix} \quad \cdots \text{ (付式22-2)}$

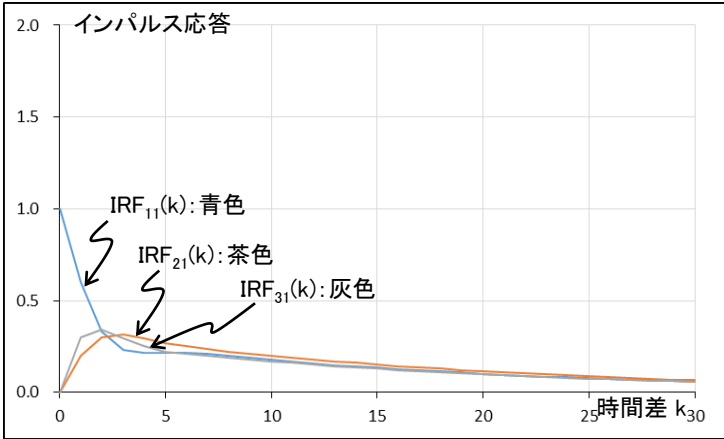
- $y_1(t)$ に 1 単位のショック(インパルス)与えたときの、時点 $k = 0$ での $y_1(t)$, $y_2(t)$, $y_3(t)$ への影響の度合いを求める。

$$\begin{aligned} \text{IRF}_{11}(0) &= \partial y_1(t) / \partial \varepsilon_1(t) = (-0.8 + 0.6 y_1(t-1) + 0.9 y_2(t-1) - 0.7 y_3(t-1) + \varepsilon_1(t)) / \partial \varepsilon_1(t) \\ &= -0.8 / \partial \varepsilon_1(t) + 0.6 y_1(t-1) / \partial \varepsilon_1(t) + 0.9 y_2(t-1) / \partial \varepsilon_1(t) - 0.7 y_3(t-1) / \partial \varepsilon_1(t) + \varepsilon_1(t) / \partial \varepsilon_1(t) = 0 + 0 + 0 - 0 + 1 = 1 \\ \text{IRF}_{21}(0) &= \partial y_2(t) / \partial \varepsilon_1(t) = (0.9 + 0.2 y_1(t-1) + 0.6 y_2(t-1) + 0.2 y_3(t-1) + \varepsilon_2(t)) / \partial \varepsilon_1(t) = 0 + 0 + 0 + 0 + 0 = 0 \\ \text{IRF}_{31}(0) &= \partial y_3(t) / \partial \varepsilon_1(t) = (0.6 + 0.3 y_1(t-1) + 0.2 y_2(t-1) + 0.4 y_3(t-1) + \varepsilon_3(t)) / \partial \varepsilon_1(t) = 0 + 0 + 0 + 0 + 0 = 0 \end{aligned}$$
- $y_1(t)$ に 1 単位のショック(インパルス)与えたときの、時点 $k = 1$ での $y_1(t+1)$, $y_2(t+1)$, $y_3(t+1)$ への影響の度合いを求める。

$$\begin{aligned} \text{IRF}_{11}(1) &= \partial y_1(t+1) / \partial \varepsilon_1(t) = (-0.8 + 0.6 y_1(t) + 0.9 y_2(t) - 0.7 y_3(t) + \varepsilon_1(t+1)) / \partial \varepsilon_1(t) = 0.6 \text{IRF}_{11}(0) + 0.9 \text{IRF}_{21}(0) - 0.7 \text{IRF}_{31}(0) = 0.6*1 + 0.9*0 - 0.7*0 = 0.6 \\ \text{IRF}_{21}(1) &= \partial y_2(t+1) / \partial \varepsilon_1(t) = (0.9 + 0.2 y_1(t) + 0.6 y_2(t) + 0.2 y_3(t) + \varepsilon_2(t+1)) / \partial \varepsilon_1(t) = 0.2 \text{IRF}_{11}(0) + 0.6 \text{IRF}_{21}(0) + 0.2 \text{IRF}_{31}(0) = 0.2*1 + 0.6*0 + 0.2*0 = 0.2 \\ \text{IRF}_{31}(1) &= \partial y_3(t+1) / \partial \varepsilon_1(t) = (0.6 + 0.3 y_1(t) + 0.2 y_2(t) + 0.4 y_3(t) + \varepsilon_3(t+1)) / \partial \varepsilon_1(t) = 0.3 \text{IRF}_{11}(0) + 0.2 \text{IRF}_{21}(0) + 0.4 \text{IRF}_{31}(0) = 0.3*1 + 0.2*0 + 0.4*0 = 0.3 \end{aligned}$$
- よって、 k 時点のインパルス応答を定式化することができる。

$$\left. \begin{aligned} \text{IRF}_{11}(k) &= 0.6 \text{IRF}_{11}(k-1) + 0.9 \text{IRF}_{21}(k-1) - 0.7 \text{IRF}_{31}(k-1) \\ \text{IRF}_{21}(k) &= 0.2 \text{IRF}_{11}(k-1) + 0.6 \text{IRF}_{21}(k-1) + 0.2 \text{IRF}_{31}(k-1) \\ \text{IRF}_{31}(k) &= 0.3 \text{IRF}_{11}(k-1) + 0.2 \text{IRF}_{21}(k-1) + 0.4 \text{IRF}_{31}(k-1) \end{aligned} \right\} \quad \cdots \text{ (付式22-3)}$$

(付式22-3)を(付図22-1)に示す。



(付図22-1)非直交化インパルス応答の計算

付録23. 直交化インパルス応答関数(計算例)

(1)コレスキー分解を使用して直交化する場合
コレスキー分解によって、 $\varepsilon(t)$ の分散共分散行列 Σ を(付式23-1)に示す三角行列に分解する(付録 I-1と付録 I-2)。

$$\Sigma = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix}^T = \mathbf{P} \mathbf{P}^T \quad \cdots \text{ (付式23-1)}$$

$\varepsilon(t) = \mathbf{P} \mathbf{v}(t)$ であるから、このコレスキー因子 $\mathbf{P}$ を使用して、
 $\varepsilon_1(t) = 2 v_1(t)$, $\varepsilon_2(t) = v_1(t) + 4 v_2(t)$, $\varepsilon_3(t) = 4 v_1(t) + 2 v_3(t)$ $\cdots$ (付式23-2)

従って、(付式22-1)は、(付式23-3)のように表すことができる。
 $y_1(t) = -0.8 + 0.6 y_1(t-1) + 0.9 y_2(t-1) - 0.7 y_3(t-1) + 2 v_1(t)$
 $y_2(t) = 0.9 + 0.2 y_1(t-1) + 0.6 y_2(t-1) + 0.2 y_3(t-1) + v_1(t) + 4 v_2(t)$
 $y_3(t) = 0.6 + 0.3 y_1(t-1) + 0.2 y_2(t-1) + 0.4 y_3(t-1) + 4 v_1(t) + 2 v_3(t)$ $\cdots$ (付式23-3)

以下、(付録22)と同じ手順で、 $y_1(t)$ に 1 単位のショック(インパルス)与えたときのインパルス応答を求めることができる。結果を、(付図23-1)に載せる。

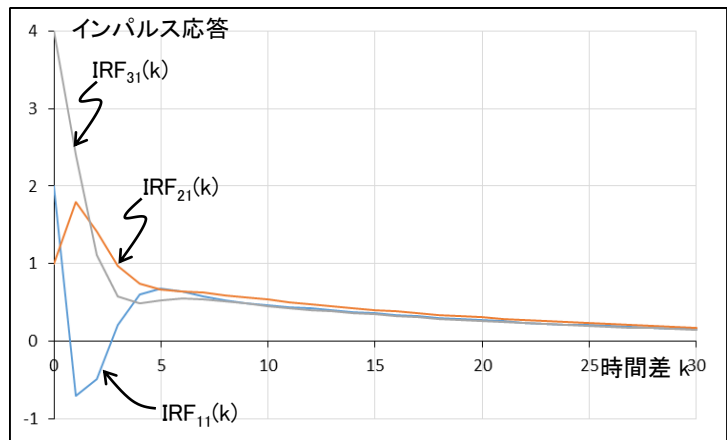
(2)修正コレスキー分解を使用して直交化する場合
修正コレスキー分解によって、 $\varepsilon(t)$ の分散共分散行列 Σ を(付式23-4)に示すように、三角行列と対角行列に分解する(付録 I-1と付録 I-2)。

$$\Sigma = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}^T = \mathbf{A} \mathbf{D} \mathbf{A}^T \quad \cdots \text{ (付式23-4)}$$

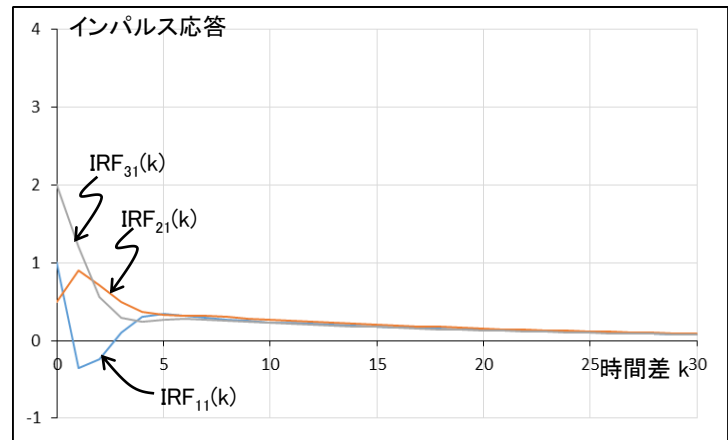
$\varepsilon(t) = \mathbf{A} \mathbf{u}(t)$ であるから、行列 $\mathbf{A}$ を使用して、
 $\varepsilon_1(t) = u_1(t)$, $\varepsilon_2(t) = 0.5 u_1(t) + u_2(t)$, $\varepsilon_3(t) = 2 u_1(t) + u_3(t)$ $\cdots$ (付式23-5)

従って、(付式22-1)は、(付式23-6)のように表すことができる。
 $y_1(t) = -0.8 + 0.6 y_1(t-1) + 0.9 y_2(t-1) - 0.7 y_3(t-1) + u_1(t)$
 $y_2(t) = 0.9 + 0.2 y_1(t-1) + 0.6 y_2(t-1) + 0.2 y_3(t-1) + 0.5 u_1(t) + u_2(t)$
 $y_3(t) = 0.6 + 0.3 y_1(t-1) + 0.2 y_2(t-1) + 0.4 y_3(t-1) + 2 u_1(t) + u_3(t)$ $\cdots$ (付式23-6)

以下、(付録22)と同じ手順で、 $y_1(t)$ に 1 単位のショック(インパルス)与えたときのインパルス応答を求めることができる。結果を、(付図23-2)に載せる。
ここでは、コレスキー分解を使った IRF の値 = 修正コレスキー分解を使った IRF の値 $\times 2$ になっていることが確認できる。



(付図23-2)直交化インパルス応答の計算(コレスキー分解)



(付図23-3)直交化インパルス応答の計算(修正コレスキー分解)

付録24. 分散分解(計算例)

インパルス応答関数の計算例で使用した(付式22-1)で表される VAR(1) 過程を対象に、分散分解の計算例を示す。

$$\left. \begin{aligned} y_1(t) &= -0.8 + 0.6 y_1(t-1) + 0.9 y_2(t-1) - 0.7 y_3(t-1) + \varepsilon_1(t) \\ y_2(t) &= 0.9 + 0.2 y_1(t-1) + 0.6 y_2(t-1) + 0.2 y_3(t-1) + \varepsilon_2(t) \\ y_3(t) &= 0.6 + 0.3 y_1(t-1) + 0.2 y_2(t-1) + 0.4 y_3(t-1) + \varepsilon_3(t) \end{aligned} \right\} \quad \cdots \text{ (付式22-1)}$$

(式13-1)に示す VAR(1) 過程のパラメータは、定数ベクトル $\mathbf{c} = \begin{pmatrix} -0.8 \\ 0.9 \\ 0.6 \end{pmatrix}$, 係数行列 $\Phi = \begin{pmatrix} 0.6 & 0.9 & -0.7 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.2 & 0.4 \end{pmatrix}$, (付式22-2)分散共分散行列 $\Sigma = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix}$

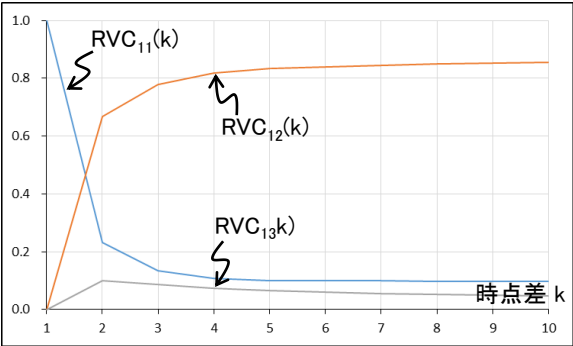
(付録 I-2)に示す通り、この Σ を修正コレスキー分解すると、 $\Sigma = \mathbf{A} \mathbf{D} \mathbf{A}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}^T$

よって、 $\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$, $\Phi \mathbf{A} = \begin{pmatrix} 0.6 & 0.9 & -0.7 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.2 & 0.4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -0.35 & 0.9 & -0.7 \\ 0.9 & 0.6 & 0.2 \\ 1.2 & 0.2 & 0.4 \end{pmatrix}$, $\Phi^2 \mathbf{A} = \begin{pmatrix} 0.6 & 0.9 & -0.7 \\ 0.2 & 0.6 & 0.2 \\ 0.3 & 0.2 & 0.4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -0.24 & 0.94 & -0.52 \\ 0.71 & 0.58 & 0.06 \\ 0.555 & 0.47 & -0.01 \end{pmatrix}$

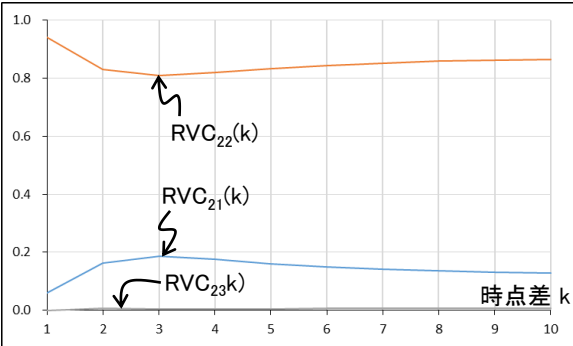
$\cdots$, $\Phi^9 \mathbf{A} = \begin{pmatrix} 0.265 & 0.461 & -0.082 \\ 0.299 & 0.526 & -0.093 \\ 0.257 & 0.445 & -0.077 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.248 & 0.438 & -0.079 \\ 0.284 & 0.497 & -0.087 \\ 0.242 & 0.422 & -0.074 \end{pmatrix}$, $\cdots$

例えば、k = 3 のとき、

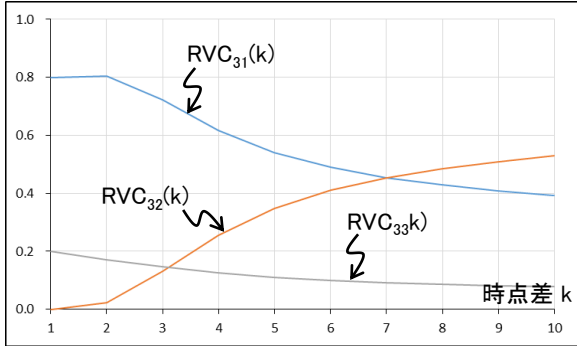
$u_1(t+1)$, $u_1(y+2)$, $u_1(t+3)$ が $\hat{y}_2(t+3)$ の予測誤差に与える部分を考慮して、 $\hat{y}_2(t+3)$ の最小二乗誤差 $\text{MSE}(\hat{y}_2(t+3))$ は、(式16-10)より、
 $\text{MSE}(\hat{y}_2(t+3)) = (0.5^2 + 0.9^2 + 0.71^2) * 4 + (1^2 + 0.6^2 + 0.58^2) * 16 + (0^2 + 0.2^2 + 0.06^2) * 4 = 6.256 + 27.142 + 0.174 = 33.573$
 従って、 $\text{RVC}_{21}(3) = 6.256 / 33.573 = 0.186$, $\text{RVC}_{22}(3) = 27.142 / 33.573 = 0.808$, $\text{RVC}_{23}(3) = 0.174 / 33.573 = 0.005$
 この計算を繰り返して、(付式22-1)で示す VAR(1) 過程における予測誤差を分散分解を行った結果を、(付図24-1)、(付図24-2)、(付図24-3)に示す。



(付図24-1) $\hat{y}_1(t+k)$ の予測誤差の分散分解



(付図24-2) $\hat{y}_2(t+k)$ の予測誤差の分散分解



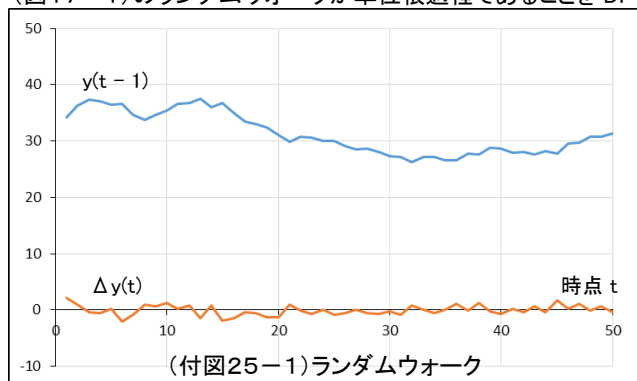
(付図24-3) $\hat{y}_3(t+k)$ の予測誤差の分散分解

付録25. DF 検定(計算例)

定数項とトレンド項を持つ過程(式18-2) $\Delta y(t) = c + \delta t + \rho y(t-1) + \varepsilon(t)$ を対象とした検定 …… 帰無仮説 $H_0: \rho = 0$ (単位根過程), 対立仮説 $H_1: \rho < 0$ (定常過程)

(1) ランダムウォーク(単位根過程)の検定

(図17-1)のランダムウォークが単位根過程であることを DF 検定を使って確認する。

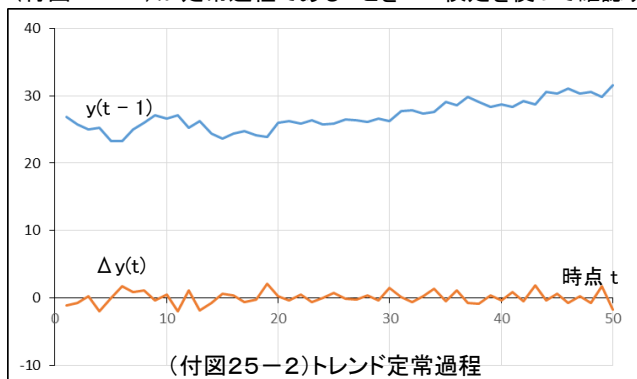


$\Delta y(t)$	t	$y(t-1)$
2.074259	1	34.22482
1.015612	2	36.29908
-0.33282	3	37.3147
-0.51309	4	36.98188
0.137274	5	36.46878
-2.00746	6	36.60606
-0.83962	7	34.5986
0.908561	8	33.75898
0.70243	9	34.66754
1.186595	10	35.36997
0.124192	11	36.55656
~	~	~
-0.06676	48	30.83087
0.583044	49	30.76411
-0.39345	50	31.34715

(付表25-1)の ADF 検定の棄却点の表(定数項、トレンドあり、サンプル数 50、有意差 5%)より、検定統計量 $-0.775954 >$ ADF 分布の棄却点 -3.495 であり、帰無仮説を棄却することはできない。よって、ランダムウォークは単位根過程である(帰無仮説を受容することになる)。

(2)トレンド定常過程の検定

(付図17-2)が定常過程であることを DF 検定を使って確認する。



$\Delta y(t)$	t	$y(t-1)$
-1.05415	1	26.80331
-0.77276	2	25.74917
0.22671	3	24.97641
-1.9687	4	25.20312
0.02661	5	23.23442
1.804151	6	23.26103
0.916359	7	25.06518
1.073981	8	25.98154
-0.37801	9	27.05552
0.457309	10	26.6775
-1.90726	11	27.13481
~	~	~
-0.76552	48	30.63016
1.690089	49	29.86465
-1.76442	50	31.55473

(付表25-2)の ADF 検定の棄却点の表(定数項、トレンドあり、サンプル数 50、有意差 5%)より、検定統計量 $-3.898179 < ADF$ 分布の棄却点 -3.495 であり、帰無仮説は棄却される。よって、この過程は定常過程である(対立仮説を採択することになる)。

(付表25-1) ADF 検定の棄却点
[Augmented Dickey-Fuller Table | Real Statistics Using Excel \(real-statistics.com\)](#)

	Model 0 - no constant, no trend				Model 1 - constant, no trend				Model 2 - constant, trend			
N	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10	0.01	0.025	0.05	0.10
25	-2.661	-2.273	-1.955	-1.609	-3.724	-3.318	-2.986	-2.633	-4.375	-3.943	-3.589	-3.238
50	-2.612	-2.246	-1.947	-1.612	-3.568	-3.213	-2.921	-2.599	-4.152	-3.791	-3.495	-3.181
100	-2.588	-2.234	-1.944	-1.614	-3.498	-3.164	-2.891	-2.582	-4.052	-3.722	-3.452	-3.153
250	-2.575	-2.227	-1.942	-1.616	-3.457	-3.136	-2.873	-2.573	-3.995	-3.683	-3.427	-3.137
500	-2.570	-2.224	-1.942	-1.616	-3.443	-3.127	-2.867	-2.570	-3.977	-3.670	-3.419	-3.132
>500	-2.567	-2.223	-1.941	-1.616	-3.434	-3.120	-2.863	-2.568	-3.963	-3.660	-3.413	-3.128

重回帰分析

概要						
回帰統計						
重相関 R	0.1541038					
重決定 R2	0.023748					
補正 R2	-0.017795					
標準誤差	0.9074732					
観測数	50					
分散分析表						
	自由度	変動	分散	測された分散	有意 F	
回帰	2	0.9415215	0.4707608	0.5716532	0.5684676	
残差	47	38.704856	0.8235076			
合計	49	39.646377				
	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	1.4516375	2.1856795	0.6641585	0.5098342	-2.945382	5.8486575
* X 値 1	-0.002693	0.0148577	-0.181267	0.8569376	-0.032583	0.0271967
* X 値 2	-0.046334	0.0597125	-0.775954	0.4416635	-0.16646	0.0737919

$$\begin{array}{cc} \hat{\delta} & \cdot \cdot \\ \hat{\rho} & \cdot \cdot \end{array}$$

重回帰分析

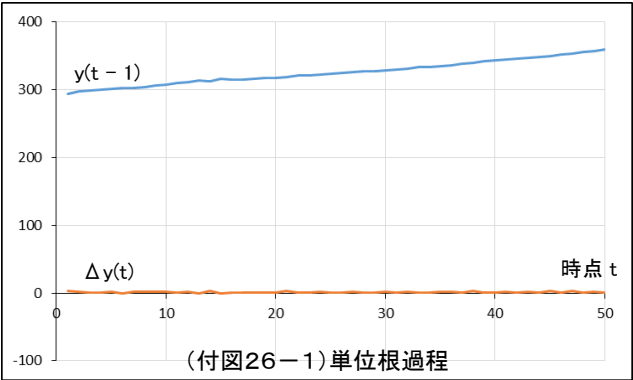
概要						
回帰統計						
重相関 R	0.5003813					
重決定 R2	0.2503814					
補正 R2	0.2184827					
標準誤差	0.866362					
観測数	50					
分散分析表						
	自由度	変動	分散	測された分散	有意 F	
回帰	2	11.783066	5.8915331	7.8492754	0.0011448	
残差	47	35.277403	0.7505831			
合計	49	47.06047				
	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	9.7287257	2.5320155	3.8422852	0.0003645	4.6349675	14.822484
X 値 1	0.0577815	0.0157607	3.6661733	0.0006254	0.026075	0.0894879
X 値 2	-0.412479	0.1058133	-3.898179	0.0003065	-0.625348	-0.19961

$$\begin{array}{cc} \hat{\delta} & \cdot \cdot \\ \hat{\rho} & \cdot \cdot \end{array}$$

付録26. ADF 検定(計算例)

定数項とトレンド項を持つ過程(式18-5) $p = 2$, $\Delta y(t) = c + \delta t + \rho y(t-1) + \zeta_1 \Delta y(t-1) + \varepsilon(t)$ 但し、 $\rho = \phi_1 + \phi_2 - 1$, $\zeta_1 = -\phi_2$ を対象とした検定
... 帰無仮説 $H_0: \rho = 0$ (単位根過程), 対立仮説 $H_1: \rho < 0$ (定常過程)

- (1) AR(2) 過程(単位根過程のとき)
(付図3-8)の AR(2) 過程が単位根過程あることを ADF 検定を使って確認する。

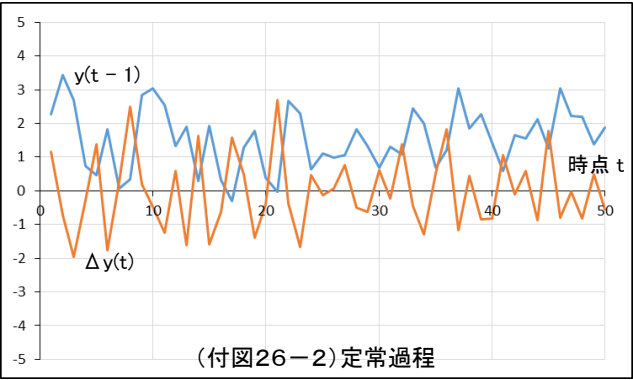


$\Delta y(t)$	t	$y(t-1)$	$\Delta y(t-1)$
3.6060092	1	294.21811	0.9206234
1.4732087	2	297.82412	3.6060092
0.9778965	3	299.29733	1.4732087
0.9957484	4	300.27523	0.9778965
1.6389742	5	301.27097	0.9957484
-0.763045	6	302.90995	1.6389742
1.365594	7	302.1469	-0.763045
2.2623236	8	303.5125	1.365594
1.6975002	9	305.77482	2.2623236
2.4075948	10	307.47232	1.6975002
1.0611539	11	309.87992	2.4075948
~	~	~	~
0.771919	48	356.07462	2.6533011
2.1742762	49	356.84654	0.771919
0.6368366	50	359.02082	2.1742762

(付表25-1)の ADF 検定の棄却点の表(定数項、トレンドあり、サンプル数 50、有意差 5%)より、検定統計量 $-0.792957 >$ ADF 分布の棄却点 -3.495 であり、帰無仮説を棄却することはできない。よって、この過程は単位根過程である(帰無仮説を受容することになる)。

$\hat{\delta}$	...				
$\hat{\rho}$	...				
$\hat{\zeta}_1$	...				

- (2) AR(2) 過程(定常過程のとき)
(図3-4)の AR(2) 過程が定常過程あることを ADF 検定を使って確認する。



$\Delta y(t)$	t	$y(t-1)$	$\Delta y(t-1)$
1.1674629	1	2.2610622	-0.233802
-0.731633	2	3.4285252	1.1674629
-1.961744	3	2.6968925	-0.731633
-0.279969	4	0.7351484	-1.961744
1.3790705	5	0.4551797	-0.279969
-1.756928	6	1.8342501	1.3790705
0.2648179	7	0.0773224	-1.756928
2.50392	8	0.3421403	0.2648179
0.1887252	9	2.8460603	2.50392
-0.479658	10	3.0347856	0.1887252
-1.233825	11	2.5551278	-0.479658
~	~	~	~
-0.818461	48	2.1985241	-0.033678
0.4995835	49	1.3800632	-0.818461
-0.549196	50	1.8796467	0.4995835

(付表25-1)の ADF 検定の棄却点の表(定数項、トレンドあり、サンプル数 50、有意差 5%)より、検定統計量 $-5.09004 <$ ADF 分布の棄却点 -3.495 であり、帰無仮説は棄却される。よって、この過程は定常過程である(対立仮説を採択することになる)。

$\hat{\delta}$	...				
$\hat{\rho}$	...				
$\hat{\zeta}_1$	...				

重回帰分析

概要					
回帰統計					
重相関 R	0.4126047				
重決定 R2	0.1702426				
補正 R2	0.116128				
標準誤差	0.9168919				
観測数	50				
分散分析表					
	自由度	変動	分散	測された分散	有意 F
回帰	3	7.9343507	2.6447836	3.1459645	0.0339293
残差	46	38.671779	0.8406908		
合計	49	46.606129			
	係数	標準誤差	t	P-値	下限 95% 上限 95%
切片	21.861689	25.472203	0.8582567	0.3952008	-29.4112 73.134575
X 値 1	0.0899012	0.105532	0.8518852	0.3986913	-0.122524 0.3023262
X 値 2	-0.068769	0.0867245	-0.792957	0.431874	-0.243336 0.1057986
X 値 3	-0.352778	0.1466293	-2.405919	0.0202081	-0.647928 -0.057629

重回帰分析

概要					
回帰統計					
重相関 R	0.641345				
重決定 R2	0.4113234				
補正 R2	0.3729314				
標準誤差	0.8903003				
観測数	50				
分散分析表					
	自由度	変動	分散	測された分散	有意 F
回帰	3	25.476365	8.4921216	10.713791	1.849E-05
残差	46	36.461192	0.7926346		
合計	49	61.937557			
	係数	標準誤差	t	P-値	下限 95% 上限 95%
切片	1.2305795	0.3625199	3.3945161	0.0014246	0.5008649 1.9602941
X 値 1	0.0056707	0.0087593	0.6473854	0.5206005	-0.011961 0.0233023
X 値 2	-0.903639	0.1775309	-5.09004	6.494E-06	-1.260991 -0.546288
X 値 3	0.1615017	0.1431199	1.1284364	0.2649867	-0.126584 0.449587

付録27. 見せかけの回帰と共和分の関係の例

(1) 見せかけの回帰

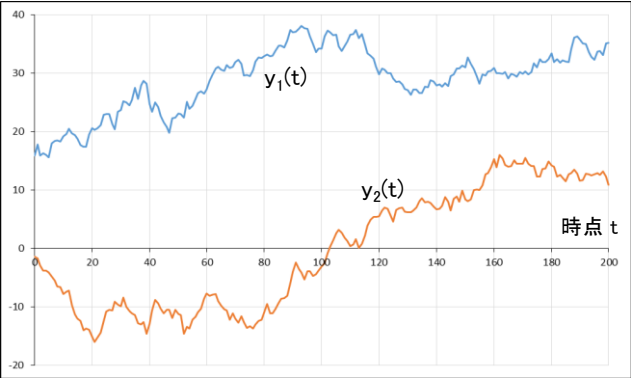
(付式27-1)、(付式27-2)のいずれもランダムウォーク(単位根過程)である(付図27-1)。

$y_1(t) = 0.1 + y_1(t-1) + \varepsilon_1(t) \quad \dots \quad (\text{付式27-1})$

$y_2(t) = 0.1 + y_2(t-1) + \varepsilon_2(t) \quad \dots \quad (\text{付式27-2})$

$y_1(t)$ を $y_2(t)$ に回帰したときの回帰分析結果を(図27-2)に示す。

$y_1(t)$ と $y_2(t)$ は互いに独立で相関がないにもかかわらず、回帰式として有意な結果(F 検定の結果、t 検定の結果)が得られている。これが見せかけの回帰である。



(付図27-1)ランダムウォーク

概要						
回帰統計						
重相関 R	0.418876					
重決定 R2	0.175457					
補正 R2	0.171293					
標準誤差	5.023788					
観測数	200					
分散分析表						
	自由度	変動	分散	割られた分散	有意 F	
回帰	1	1063.374	1063.374	42.13312	6.71E-10	
残差	198	4997.212	25.23844			
合計	199	6060.586				
	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	29.08849	0.355428	81.84083	1.3E-154	28.38758	29.7894
X 値 1	0.225486	0.034738	6.491003	6.71E-10	0.156982	0.29399

(付図27-2)回帰分析の結果

(2) 共和分の関係

(付式27-3)はランダムウォーク(単位根過程)である。また、(付式27-4)も、単位根過程である $y_1(t-1)$ を持つ式になっているので単位根過程である。

$y_1(t) = 0.1 + 1.0 y_1(t-1) + \varepsilon_1(t) \quad \dots \quad (\text{付式27-3})$

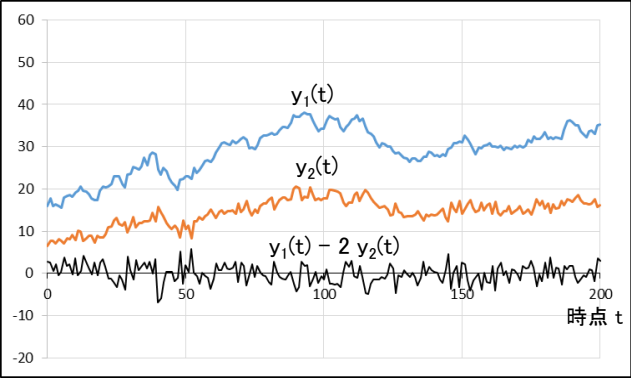
$y_2(t) = 0.1 + 0.5 y_1(t-1) + \varepsilon_2(t) \quad \dots \quad (\text{付式27-4})$

(付式27-3)と(付式27-4)より、(付式27-5)を作る。

$y_1(t) - 1/0.5 y_2(t) = (0.1 + 1.0 y_1(t-1) + \varepsilon_1(t)) - 2 (0.1 + 0.5 y_1(t-1) + \varepsilon_2(t)) = -0.1 + \varepsilon_1(t) - \varepsilon_2(t) \quad \dots \quad (\text{付式27-5})$

(付式27-5)の右辺からは単位根過程の時系列過程が消えて、定数項と誤差項(定常過程)だけが残っている。

よって、(付式27-5)は定常過程である。これが共和分の関係である。(付式27-5)が定常過程であるという結果は、(付図27-3)からも確認できる。



(付図27-1)共和分の関係

付録28. 共和分関係のある VAR(p) 過程の性質

共和分関係がある差分系列は VAR 過程では表現できないことを、以下で確認する。

(1) 共和分システムの VAR 表現とそれが単位根過程であることの確認

下式は、(付式27-3)と(付式27-4)の定数項を 0 にした共和分関係のある過程である。

$$y_1(t) = 1.0 y_1(t-1) + \varepsilon_1(t) \quad \varepsilon_1(t) \sim W.N(\sigma_1^2) \quad \dots \text{ (付式28-1)}$$

$$y_2(t) = 0.5 y_1(t-1) + \varepsilon_2(t) \quad \varepsilon_2(t) \sim W.N(\sigma_2^2) \quad \dots \text{ (付式28-2)}$$

(付式28-1)と(付式28-2)を VAR(1) 過程の形で表現する。

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix} \begin{pmatrix} y_1(t-1) \\ y_2(t-1) \end{pmatrix} + \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix} \quad \dots \text{ (付式28-3)}$$

$$\text{このVAR(1) 過程の定常性を判断する特性方程式は、(式6-7)より、} \left| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0.5 & 0 \end{pmatrix} z^{-1} \right| = \begin{vmatrix} 1 - z^{-1} & 0 \\ -0.5 z^{-1} & 1 \end{vmatrix} = 0$$

従って、 $1 - z^{-1} = 0$ から、 $z^{-1} = 1$ この共和分システム VAR(1) 過程は、単位根を持っている過程であることを確認できる。

(2) 次に、共和分システムの 差分系列は VAR 過程で表現できないことの確認

$y_1(t)$ と $y_2(t)$ の差分系列を求める。

$$\Delta y_1(t) = y_1(t) - y_1(t-1) = \varepsilon_1(t) \quad \dots \text{ (付式28-4)}$$

$$\begin{aligned} \Delta y_2(t) &= y_2(t) - y_2(t-1) = (0.5 y_1(t-1) + \varepsilon_2(t)) - (0.5 y_1(t-2) + \varepsilon_2(t-1)) = 0.5 (y_1(t-1) - y_1(t-2)) + \varepsilon_2(t) - \varepsilon_2(t-1) \\ &= 0.5 \Delta y_1(t-1) + \varepsilon_2(t) - \varepsilon_2(t-1) = 0.5 \varepsilon_1(t-1) + \varepsilon_2(t) - \varepsilon_2(t-1) = \varepsilon_2(t) + 0.5 \varepsilon_1(t-1) - \varepsilon_2(t-1) \quad \dots \text{ (付式28-5)} \end{aligned}$$

差分過程(付式28-4)と(付式28-5)を VMA(1) 過程の形で表現する。

$$\begin{pmatrix} \Delta y_1(t) \\ \Delta y_2(t) \end{pmatrix} = \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0.5 & -1 \end{pmatrix} \begin{pmatrix} \varepsilon_1(t-1) \\ \varepsilon_2(t-1) \end{pmatrix} \quad \dots \text{ (付式28-6)}$$

$$\text{このVMA(1) 過程のVAR(1)過程への反転可能性を判断する特性方程式は、(式8-2)より、} \left| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0.5 & -1 \end{pmatrix} z^{-1} \right| = \begin{vmatrix} 1 & 0 \\ -0.5 z^{-1} & 1 - z^{-1} \end{vmatrix} = 0$$

従って、 $1 - z^{-1} = 0$ から、 $z^{-1} = 1$ この共和分システムの差分過程(VMA(1) 過程)を、VAR(1)過程表現に変換することは不可能である(反転不可能である)。

付録29. Johansen(ヨハンセン)の尤度比検定統計量の導出 (1/5)

(A) 検定統計量の導出過程の見通しをよくするために、(式20-1)に示す誤差修正モデル(VECM)を簡略表記して使用する。VECM を簡略表記した式を(付式29-2)に示す。
 (付式29-2)から分かるように、このモデルのパラメータは、 Γ , Π , Λ である。

$$\Delta \mathbf{y}(t) = \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-1) + \cdots + \Gamma_{p-1} \Delta \mathbf{y}(t-(p-1)) + \Pi \mathbf{y}(t-1) + \mathbf{c} + \boldsymbol{\varepsilon}(t) \quad \cdots \text{ (式20-1)}$$

$$= \Gamma_1 \Delta \mathbf{y}(t-1) + \Gamma_2 \Delta \mathbf{y}(t-1) + \cdots + \Gamma_{p-1} \Delta \mathbf{y}(t-(p-1)) + \Gamma_p \mathbf{1} + \Pi \mathbf{y}(t-1) + \boldsymbol{\varepsilon}(t) \quad \cdots \text{ (付式29-1)}$$

$$\mathbf{z}_0(t) = \Gamma \mathbf{z}_p(t) + \Pi \mathbf{z}_1(t) + \boldsymbol{\varepsilon}(t) \quad \cdots \text{ (付式29-2)}$$

ここで、

n : 変数の数、 p : 時点(ラグ)の数、 h : 共和分関係の数、 T : サンプル数

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}_{n \times 1}, \quad \Gamma_i = \begin{pmatrix} \gamma_{i,11} & \gamma_{i,12} & \cdots & \gamma_{i,1n} \\ \gamma_{i,21} & \gamma_{i,22} & \cdots & \gamma_{i,2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{i,n1} & \gamma_{i,n2} & \cdots & \gamma_{i,nn} \end{pmatrix}_{n \times n}, \quad \Gamma_p = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}_{n \times n}, \quad \mathbf{1} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}_{n \times 1}, \quad \boldsymbol{\alpha} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1h} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2h} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nh} \end{pmatrix}_{n \times h}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1h} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2h} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nh} \end{pmatrix}_{n \times h}, \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}_{n \times 1}$$

$$\boldsymbol{\varepsilon}(t) = \begin{pmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \\ \vdots \\ \varepsilon_n(t) \end{pmatrix}_{n \times 1}, \quad \boldsymbol{\varepsilon}(t) \text{ は、独立同一分布(i.i.d. : independent and identically distributed)で、正規分布 } N(\mathbf{0}, \Lambda) \text{ に従う。}$$

$$\Lambda = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}_{n \times n} \quad (\text{対称行列})$$

$$\mathbf{z}_0(t) = \Delta \mathbf{y}(t) = \begin{pmatrix} \Delta y_1(t) \\ \Delta y_2(t) \\ \vdots \\ \Delta y_n(t) \end{pmatrix}_{n \times 1}$$

$$\mathbf{z}_p(t) = \begin{pmatrix} \Delta \mathbf{y}^T(t-1) \\ \Delta \mathbf{y}^T(t-2) \\ \vdots \\ \Delta \mathbf{y}^T(t-(p-1)) \\ 1 \end{pmatrix}_{(n(p-1)+1) \times 1} = \begin{pmatrix} \Delta y_1(t-1) \\ \Delta y_2(t-1) \\ \vdots \\ \Delta y_n(t-1) \\ \Delta y_1(t-2) \\ \Delta y_2(t-2) \\ \vdots \\ \Delta y_n(t-2) \\ \vdots \\ \Delta y_1(t-(p-1)) \\ \Delta y_2(t-(p-1)) \\ \vdots \\ \Delta y_n(t-(p-1)) \\ 1 \end{pmatrix}$$

$$\mathbf{z}_1(t) = \mathbf{y}(t-1) = \begin{pmatrix} y_1(t-1) \\ y_2(t-1) \\ \vdots \\ y_n(t-1) \end{pmatrix}_{n \times 1}$$

Γ は、(付式29-1)の $\Gamma_1 \sim \Gamma_p$ をまとめた、サイズ $n \times (n(p-1)+1)$ の行列である。
 $\Pi = \boldsymbol{\alpha} \boldsymbol{\beta}^T$ サイズ $n \times n$ $\cdots$ (付式29-3)
 この式から分かるように、 Π は $\boldsymbol{\beta}$ (共和分ベクトル)の関数である。

付録29. Johansen(ヨハンセン)の尤度比検定統計量の導出 (2/5)

(注)ここでは、和を表す $\sum_{t=1}^T$ は省略して、 Σ で表す。 また、推定値は、変数に $\hat{}$ の記号を付けて表す。

(B) 対立仮説 H_1 の下での VECM モデルのパラメータを推定する(最尤推定量を求める)。

対立仮説 H_1 の下でということとは、共和分存在は気にしないで尤度を求める、ということになる。つまり、共和分ベクトル β の関数である Π は固定値を取るものとして考える。

(B1) (式29-2)に示すVECM モデルのパラメータを推定するための尤度の計算式に使用する $\Delta y(t)$ の推定残差を求める。

(B1-1)まず、(付式29-2)のパラメータ Γ を推定する。

(付式29-2)を式変形して、 $z_0(t) - \Pi z_1(t) = \Gamma z_p(t) + \varepsilon(t) \quad \dots$ (付式29-4) が得られる。

この式を、 **$z_0(t) - \Pi z_1(t)$ を $z_p(t)$ に回帰する**回帰式であると捉え(なお、 Π は固定値)、 Γ の推定量 $\hat{\Gamma}$ を最小二乗法を使用して求める。

ここで、 $M_{ij} = (1/T) \sum z_i(t) (z_j(t))^T \quad (i, j = 0, 1, p) \quad \dots$ (付式29-5) とおく。 なお、 M_{ij} は対称行列である。

$$M_{0p} = (1/T) \sum z_0(t) (z_p(t))^T, M_{pp} = (1/T) \sum z_p(t) (z_p(t))^T, M_{p1} = (1/T) \sum z_p(t) (z_1(t))^T \text{ など}$$

(付式29-4)の正規方程式(別途資料「重回帰分析の概要」を参照)は、

$$(z_0(t) - \Pi z_1(t)) (z_p(t))^T = \hat{\Gamma} z_p(t) (z_p(t))^T \rightarrow z_0(t) (z_p(t))^T - \Pi z_1(t) (z_p(t))^T = \hat{\Gamma} z_p(t) (z_p(t))^T \rightarrow z_0(t) (z_p(t))^T = \Pi z_1(t) (z_p(t))^T + \hat{\Gamma} z_p(t) (z_p(t))^T \quad \dots \text{ (付式29-6)}$$

(付式29-5)を代入して、推定量 $\hat{\Gamma}$ を求めることができる。

$$T M_{0p} = T \Pi M_{1p} + T \hat{\Gamma} M_{pp} \rightarrow M_{0p} = \Pi M_{1p} + \hat{\Gamma} M_{pp} \rightarrow \hat{\Gamma} M_{pp} = M_{0p} - \Pi M_{1p} \rightarrow \hat{\Gamma} = (M_{0p} - \Pi M_{1p}) (M_{pp})^{-1} \quad \dots \text{ (付式29-7)}$$

(B1-2) Π を固定値にした上で、 $z_0(t)$ を $z_p(t)$ に回帰して、 $z_0(t)$ の推定残差ベクトル $u_0^{\wedge}(t)$ を求める。

$z_0(t)$ を $z_p(t)$ に回帰する式は、(付式29-2)で $z_1(t)$ の項を無視して、 $z_0(t) = \Gamma z_p(t) + \varepsilon(t) \quad \dots$ (付式29-8)

この回帰式の正規方程式(別途資料「重回帰分析の概要」を参照)は(ここでの $\hat{\Gamma}$ は、(付式29-7)の $\hat{\Gamma}$ とは異なる)、

$$z_0(t) (z_p(t))^T = \hat{\Gamma} z_p(t) (z_p(t))^T \rightarrow \hat{\Gamma} = z_0(t) (z_p(t))^T (z_p(t) (z_p(t))^T)^{-1} = T M_{0p} (T M_{pp})^{-1} = M_{0p} (M_{pp})^{-1}$$

よって、 $z_0(t)$ の推定値は(付式29-8)より、 $z_0^{\wedge}(t) = \hat{\Gamma} z_p(t) = M_{0p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-9)

従って、 $z_0(t)$ の推定残差ベクトルを $u_0^{\wedge}(t)$ とすると、 $u_0^{\wedge}(t) = z_0(t) - z_0^{\wedge}(t) = z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-10)

また、 $z_0(t) = u_0^{\wedge}(t) + M_{0p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-11)

(B1-3) Π を固定値にした上で、 $z_1(t)$ を $z_p(t)$ に回帰して、 $z_p(t)$ の推定残差ベクトル $u_1^{\wedge}(t)$ を求める。

$z_1(t)$ を $z_p(t)$ に回帰する式は、(付式29-2)で $z_0(t)$ の項を無視して、 $-\Pi z_1(t) = \Gamma z_p(t) + \varepsilon(t) \quad \dots$ (付式29-12)

この回帰式の正規方程式(別途資料「重回帰分析の概要」を参照)は(ここでの $\hat{\Gamma}$ は、(付式29-7)の $\hat{\Gamma}$ とは異なる)、

$$-\Pi z_1(t) (z_p(t))^T = \hat{\Gamma} z_p(t) (z_p(t))^T \rightarrow \hat{\Gamma} = -\Pi z_1(t) (z_p(t))^T (z_p(t) (z_p(t))^T)^{-1} = -\Pi T M_{1p} (T M_{pp})^{-1} = -\Pi M_{1p} (M_{pp})^{-1}$$

よって、 $z_1(t)$ の推定値は(付式29-12)より、 $z_1^{\wedge}(t) = -\Pi^{-1} \hat{\Gamma} z_p(t) = -\Pi^{-1} (-\Pi M_{1p} (M_{pp})^{-1}) z_p(t) = M_{1p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-13)

従って、 $z_1(t)$ の推定残差ベクトルを $v^{\wedge}(t)$ とすると、 $u_1^{\wedge}(t) = z_1(t) - z_1^{\wedge}(t) = z_1(t) - M_{1p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-14)

また、 $z_1(t) = u_1^{\wedge}(t) + M_{1p} (M_{pp})^{-1} z_p(t) \quad \dots$ (付式29-15)

(B1-4) 固定値 Π と(付式41-7)の推定値 $\hat{\Gamma}$ を使用して、 $z_0(t)$ の推定残差ベクトルを $u_0^{\wedge}(t)$ と $u_1^{\wedge}(t)$ を使って表す。

(付式29-2)より、 $z_0^{\wedge}(t) = \hat{\Gamma} z_p(t) + \Pi z_1(t)$ であるから、また、(付式29-11)と(付式29-15)を代入し

$$z_0(t) - z_0^{\wedge}(t) = z_0(t) - (\hat{\Gamma} z_p(t) + \Pi z_1(t)) = u_0^{\wedge}(t) + M_{0p} (M_{pp})^{-1} z_p(t) - \hat{\Gamma} z_p(t) - \Pi (u_1^{\wedge}(t) + M_{1p} (M_{pp})^{-1} z_p(t)) = u_0^{\wedge}(t) - \Pi u_1^{\wedge}(t) + (M_{0p} - \Pi M_{1p}) (M_{pp})^{-1} z_p(t) - \hat{\Gamma} z_p(t)$$

ここで、(付式29-7)を代入して、 $z_0(t) - z_0^{\wedge}(t) = u_0^{\wedge}(t) - \Pi u_1^{\wedge}(t) + \hat{\Gamma} z_p(t) - \hat{\Gamma} z_p(t) = u_0^{\wedge}(t) - \Pi u_1^{\wedge}(t) \quad \dots$ (付式29-16)

この式は、 $\Delta y(t) = z_0(t) - z_0^{\wedge}(t)$ であるので、 $\Delta y(t)$ の推定残差を表している式である。

付録29. Johansen(ヨハンセン)の尤度比検定統計量の導出 (3/5)

(B2) VECM モデルのパラメータ Π と Λ を推定するための対数尤度を設定する。

対数尤度 $\ln L(\mu, \Lambda)$ は、(付録G)に倣った(付式29-17)に、(付式29-16)を代入して、

$$\ln L(\Pi, \Lambda) = -(Tn/2) \ln(2\pi) - (T/2) \ln |\Lambda| - (1/2) \sum (z_0(t) - \hat{z}_0(t))^T \Lambda^{-1} (z_0(t) - \hat{z}_0(t)) \quad \dots \text{ (付式29-17)}$$

$$\ln L(\Pi, \Lambda) = -(Tn/2) \ln(2\pi) - (T/2) \ln |\Lambda| - (1/2) \sum (\hat{u}(t) - \Pi \hat{v}(t))^T \Lambda^{-1} (\hat{u}(t) - \Pi \hat{v}(t)) \quad \dots \text{ (付式29-18)}$$

(B3) 対数尤度を使用して、パラメータ Π と Λ の最尤推定量を求める。

(B3-1) 新たに、残差の積率行列(分散共分散行列)を定義する。

$$\begin{aligned} \textcircled{1} S_{00} &= (1/T) \sum u_0(t) (u_0(t))^T = (1/T) \sum (z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t)) (z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t))^T = (1/T) \sum (z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t)) ((z_0(t))^T - (z_p(t))^T (M_{pp})^{-1} (M_{0p})^T) \\ &= (1/T) \sum (z_0(t) (z_0(t))^T - z_0(t) (z_p(t))^T ((M_{pp})^{-1})^T (M_{0p})^T - M_{0p} (M_{pp})^{-1} z_p(t) (z_0(t))^T + M_{0p} (M_{pp})^{-1} z_p(t) (z_p(t))^T ((M_{pp})^{-1})^T (M_{0p})^T) \\ &= M_{00} - M_{0p} ((M_{pp})^{-1})^T (M_{0p})^T - M_{0p} (M_{pp})^{-1} M_{p0} + M_{0p} (M_{pp})^{-1} M_{pp} ((M_{pp})^{-1})^T (M_{0p})^T = M_{00} - M_{0p} ((M_{pp})^{-1})^T (M_{0p})^T - M_{0p} (M_{pp})^{-1} M_{p0} + M_{0p} ((M_{pp})^{-1})^T (M_{0p})^T = M_{00} - M_{0p} (M_{pp})^{-1} M_{p0} \\ \textcircled{2} S_{01} &= (1/T) \sum u_0(t) (u_1(t))^T = (1/T) \sum (z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t)) (z_1(t) - M_{1p} (M_{pp})^{-1} z_p(t))^T = (1/T) \sum (z_0(t) - M_{0p} (M_{pp})^{-1} z_p(t)) ((z_1(t))^T - (z_p(t))^T (M_{pp})^{-1} (M_{1p})^T) \\ &= (1/T) \sum (z_0(t) (z_1(t))^T - z_0(t) (z_p(t))^T ((M_{pp})^{-1})^T (M_{1p})^T - M_{0p} (M_{pp})^{-1} z_p(t) (z_1(t))^T + M_{0p} (M_{pp})^{-1} z_p(t) (z_p(t))^T ((M_{pp})^{-1})^T (M_{1p})^T) \\ &= M_{01} - M_{0p} ((M_{pp})^{-1})^T (M_{1p})^T - M_{0p} (M_{pp})^{-1} M_{p1} + M_{0p} (M_{pp})^{-1} M_{pp} ((M_{pp})^{-1})^T (M_{1p})^T = M_{01} - M_{0p} ((M_{pp})^{-1})^T (M_{1p})^T - M_{0p} (M_{pp})^{-1} M_{p1} + M_{0p} ((M_{pp})^{-1})^T (M_{1p})^T = M_{01} - M_{0p} (M_{pp})^{-1} M_{p1} \\ \text{同様に、} \\ \textcircled{3} S_{10} &= (1/T) \sum u_1(t) (u_0(t))^T = M_{10} - M_{1p} (M_{pp})^{-1} M_{p0} \\ \textcircled{4} S_{11} &= (1/T) \sum u_1(t) (u_1(t))^T = M_{11} - M_{1p} (M_{pp})^{-1} M_{p1} \end{aligned}$$

同様に、

$$\textcircled{3} S_{10} = (1/T) \sum u_1(t) (u_0(t))^T = M_{10} - M_{1p} (M_{pp})^{-1} M_{p0}$$

$$\textcircled{4} S_{11} = (1/T) \sum u_1(t) (u_1(t))^T = M_{11} - M_{1p} (M_{pp})^{-1} M_{p1}$$

なお、 M_{ij} は対称行列であるから、 $(S_{01})^T = (M_{01} - M_{0p} (M_{pp})^{-1} M_{p1})^T = (M_{01})^T - (M_{p1})^T ((M_{pp})^{-1})^T (M_{0p})^T = M_{10} - M_{1p} (M_{pp})^{-1} M_{p0} = S_{10}$ となり、 S_{01} と S_{10} は対称行列の関係になる。

(B3-2) 最尤推定量 Π^* を求める。

対数尤度(付式29-18)を Π で偏微分して、その値を 0 にする Π が最尤推定量になる。ベクトルと行列の微分の公式(付録C)を適用して、

$$\partial (\ln L(\Pi, \Lambda)) / \partial \Pi = (1/2) \sum (\Lambda^{-1} + (\Lambda^{-1})^T) (u_0(t) - \Pi u_1(t)) (u_1(t))^T = 0$$

よって、 $\sum (u_0(t) - \Pi u_1(t)) (u_1(t))^T = 0 \rightarrow \sum u_0(t) (u_1(t))^T - \Pi \sum u_1(t) (u_1(t))^T = 0$ より、 Π の最尤推定量 Π^* は、

$$\Pi^* = \sum u_0(t) (u_1(t))^T / \sum u_1(t) (u_1(t))^T = (T S_{01}) / (T S_{11}) = S_{01} / S_{11} = S_{01} (S_{11})^{-1} \quad \dots \text{ (付式29-19)}$$

(B3-3) 最尤推定量 Λ^* を求める。

対数尤度(付式41-17)を Λ で偏微分して、その値を 0 にする Λ が最尤推定量である。ベクトルと行列の微分の公式(付録C)を適用して、

$$\partial (\ln L(\Pi, \Lambda)) / \partial \Lambda = -(T/2) ((\Lambda^{-1})^T + (\Lambda^{-1})) (u_0(t) - \Pi^* u_1(t)) (u_0(t) - \Pi^* u_1(t))^T (\Lambda^{-1} \Lambda^{-1}) = 0$$

よって、 $-(T/2) (\Lambda^{-1})^T + (1/2) \sum (u_0(t) - \Pi^* u_1(t)) (u_0(t) - \Pi^* u_1(t))^T (\Lambda^{-1} \Lambda^{-1}) = 0 \rightarrow -(T/2) \Lambda + (1/2) \sum (u_0(t) - \Pi^* u_1(t)) (u_0(t) - \Pi^* u_1(t))^T = 0$

従って、 Λ の最尤推定量 Λ^* は、

$$\begin{aligned} \Lambda^* &= (1/T) \sum (u_0(t) - \Pi^* u_1(t)) (u_0(t) - \Pi^* u_1(t))^T = (1/T) \sum (u_0(t) - \Pi^* u_1(t)) ((u_0(t))^T - (u_1(t))^T (\Pi^*)^T) \\ &= (1/T) \sum (u_0(t) (u_0(t))^T - u_0(t) (u_1(t))^T (\Pi^*)^T - \Pi^* u_1(t) (u_0(t))^T + \Pi^* u_1(t) (u_1(t))^T (\Pi^*)^T) = S_{00} - S_{01} (\Pi^*)^T - \Pi^* S_{10} + \Pi^* S_{11} (\Pi^*)^T \end{aligned}$$

(付式29-19)を代入する。

$$\begin{aligned} \Lambda^* &= S_{00} - S_{01} (S_{01} (S_{11})^{-1})^T - S_{01} (S_{11})^{-1} S_{10} + S_{01} (S_{11})^{-1} S_{11} (S_{01} (S_{11})^{-1})^T = S_{00} - S_{01} ((S_{11})^{-1})^T (S_{01})^T - S_{01} (S_{11})^{-1} S_{10} + S_{01} (S_{11})^{-1} S_{11} ((S_{11})^{-1})^T (S_{01})^T \\ &= S_{00} - S_{01} ((S_{11})^T)^{-1} S_{10} - S_{01} (S_{11})^{-1} S_{10} + S_{01} (S_{11})^{-1} S_{11} ((S_{11})^T)^{-1} S_{10} = S_{00} - S_{01} (S_{11})^{-1} S_{10} - S_{01} (S_{11})^{-1} S_{10} + S_{01} (S_{11})^{-1} S_{11} (S_{11})^{-1} S_{10} \\ &= S_{00} - S_{01} (S_{11})^{-1} S_{10} - S_{01} (S_{11})^{-1} S_{10} + S_{01} (S_{11})^{-1} S_{10} = S_{00} - S_{01} (S_{11})^{-1} S_{10} \quad \dots \text{ (付式29-20)} \end{aligned}$$

付録29. Johansen(ヨハンセン)の尤度比検定統計量の導出 (4/5)

(C)尤度比検定統計量を導出する。

(C1) **対立仮説** H_1 の下で、VECM モデルの最大尤度を求める。

(付式29-18)の対数尤度が最大値を取るのは、パラメータ Π と Λ が最尤推定量 $\hat{\Pi}$ 、 $\hat{\Lambda}$ のとき、 $\sum (u_0^*(t) - \hat{\Pi} u_1^*(t))^T (\hat{\Lambda})^{-1} (u_0^*(t) - \hat{\Pi} u_1^*(t)) = 0$ のときである。
よって、(付式29-18)より、対数尤度の最大値は、 $\ln L(\hat{\Pi}, \hat{\Lambda}) = - (T n / 2) \ln (2 \pi) - (T / 2) \ln |\hat{\Lambda}| \rightarrow \ln L(\hat{\Pi}, \hat{\Lambda})$ であり、定数項を無視すれば、

$$\ln L(\hat{\Pi}, \hat{\Lambda}) = - (T / 2) \ln |\hat{\Lambda}| \rightarrow \ln L(\hat{\Pi}, \hat{\Lambda}) = \ln |\hat{\Lambda}|^{-T/2} \rightarrow L(\hat{\Pi}, \hat{\Lambda}) = |\hat{\Lambda}|^{-T/2}$$

従って、対立仮説 H_1 の下での最大尤度の $-2/T$ 乗は、 $(L(H_1))^{-2/T} = (L(\hat{\Pi}, \hat{\Lambda}))^{-2/T} = |\hat{\Lambda}| \quad \dots$ (付式29-21)

(C2) **帰無仮説** H_2 (h) の下で、VECM モデルの最大尤度を求める。

帰無仮説 H_2 (h) の下でということは、rank $\Pi \leq h$ ($h = 0, \dots, n-1$) として尤度を求めるということである。
 Π のランクを決める共和分ベクトル β ($n \times h$) を固定し、パラメータ α と Λ を β の関数としてとらえる。

(C2-1) 推定残差 $u^*(t)$ を共和分の項 $\beta^T v^*(t)$ に回帰して、 α の推定値 $\hat{\alpha}(\beta)$ を求める。

(付式29-2)に、(付式29-3) および、 $z_0(t) = z_0^*(t) + u_0^*(t)$ と $z_1(t) = z_1^*(t) + u_1^*(t)$ を代入する。

$$z_0^*(t) + u_0^*(t) = \Gamma z_p(t) + \alpha \beta^T (z_1^*(t) + u_1^*(t)) + \varepsilon(t) = \Gamma z_p(t) + \alpha \beta^T z_1^*(t) + \alpha \beta^T u_1^*(t) + \varepsilon(t)$$

従って、 $u_0^*(t)$ を $\beta^T u_1^*(t)$ に回帰する回帰式は、 $u_0^*(t) = \alpha \beta^T u_1^*(t) + \varepsilon(t)$

この回帰式の正規方程式(別途資料「重回帰分析の概要」を参照)は、 $u_0^*(t) (\beta^T u_1^*(t))^T = \hat{\alpha}(\beta) \beta^T u_1^*(t) (\beta^T u_1^*(t))^T$ であるから、

$$\hat{\alpha}(\beta) = u_0^*(t) (\beta^T u_1^*(t))^T (\beta^T u_1^*(t) (\beta^T u_1^*(t))^T)^{-1} = u_0^*(t) (u_1^*(t))^T \beta (\beta^T u_1^*(t) (u_1^*(t))^T \beta)^{-1} = T S_{01} \beta (\beta^T T S_{11} \beta)^{-1} = S_{01} \beta (\beta^T S_{11} \beta)^{-1} \quad \dots$$
 (付式29-22)

(C2-2) Λ の推定値 $\hat{\Lambda}(\beta)$ を求める。

(付式29-20)に(付式29-22)を反映して、

$$\begin{aligned} \hat{\Lambda}(\beta) &= S_{00} - S_{01} (S_{11})^{-1} S_{10} = S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{10} = S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T (S_{01})^T = S_{00} - \hat{\alpha}(\beta) \beta^T (S_{01})^T = S_{11} - \hat{\alpha}(\beta) (S_{01} \beta)^T \\ &= S_{00} - \hat{\alpha}(\beta) (\hat{\alpha}(\beta) \beta^T S_{11} \beta)^T = S_{00} - \hat{\alpha}(\beta) (\beta^T (S_{11})^T \beta (\hat{\alpha}(\beta))^T) = S_{00} - \hat{\alpha}(\beta) \beta^T S_{11} \beta (\hat{\alpha}(\beta))^T \quad \dots$$
 (付式29-23)

(C2-3) 帰無仮説 H_2 (h) の下での、VECM の最大尤度を求める。

対立仮説 H_1 は、帰無仮説 H_2 (h) を包含しているので、帰無仮説 H_2 (h) の下での最大尤度は、(式29-2)において β をパラメータにした尤度になる。尤度の $-2/T$ 乗は、 $(L(\hat{\beta}))^{-2/T} = |\hat{\Lambda}(\hat{\beta})| \quad \dots$ (付式29-24)

(付式29-24)に、(付式29-23)を代入し、次いで、(付式29-22)を代入して、

$$\begin{aligned} (L(\hat{\beta}))^{-2/T} &= |S_{00} - \hat{\alpha}(\beta) \beta^T S_{11} \beta (\hat{\alpha}(\beta))^T| = |S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{11} \beta (S_{01} \beta (\beta^T S_{11} \beta)^{-1})^T| \\ &= |S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{11} \beta ((\beta^T S_{11} \beta)^{-1})^T (S_{01} \beta)^T| = |S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{11} \beta ((\beta^T S_{11} \beta)^T)^{-1} (S_{01} \beta)^T| \\ &= |S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{11} \beta (\beta^T S_{11} \beta)^{-1} (S_{01} \beta)^T| = |S_{00} - S_{01} \beta (\beta^T S_{11} \beta)^{-1} \beta^T S_{10}| = |S_{00} - S_{01} \beta (\beta^{-1} (S_{11})^{-1} (\beta^T)^{-1}) \beta^T S_{10}| \\ &= |S_{00} - S_{01} (S_{11})^{-1} S_{10}| = |S_{00}| - |S_{01} (S_{11})^{-1} S_{10}| \quad \dots$$
 (付式29-25)

一般に、 A, B が正方行列のとき、 $|AB| = |A| |B|$ 、 $|A^T| = |A|$ 、 $|A^{-1}| = |A|^{-1} = 1 / |A|$ が成り立つので、また、行列式は正方行列を対象にしていることに注意して、

$$\begin{aligned} |S_{01} (S_{11})^{-1} S_{10}| &= |S_{00} (S_{00})^{-1} S_{01} (S_{11})^{-1} S_{10}| = |S_{00}| |S_{00}^{-1}| |S_{01}| |S_{11}^{-1}| |S_{10}| = |S_{00}| |S_{10}| |S_{00}^{-1}| |S_{01}| |S_{11}^{-1}| = |S_{00}| |S_{10}| |S_{00}^{-1}| |S_{01}| |S_{11}^{-1}| |\beta^T S_{11} \beta| |\beta^T S_{11} \beta|^{-1} \\ &= |S_{00}| |\beta^T S_{10}| |S_{00}^{-1}| |S_{01}| |S_{11}^{-1}| S_{11} \beta ||\beta^T S_{11} \beta|^{-1} = |S_{00}| |\beta^T S_{10} (S_{00})^{-1} S_{10} (S_{11})^{-1} S_{11} \beta| |\beta^T S_{11} \beta|^{-1} \\ &= |S_{00}| |\beta^T S_{10} (S_{00})^{-1} S_{10} \beta| |\beta^T S_{11} \beta|^{-1} \quad \dots$$
 (付式29-26)

(付式29-25)に(付式29-26)を代入して、尤度が求まる。

$$(L(\hat{\beta}))^{-2/T} = |S_{00}| - |S_{00}| |\beta^T S_{10} (S_{00})^{-1} S_{01} \beta| |\beta^T S_{11} \beta|^{-1} = |S_{00}| (1 - |\beta^T S_{10} (S_{00})^{-1} S_{01} \beta| |\beta^T S_{11} \beta|^{-1}) \quad \dots$$
 (付式29-27)

尤度 $L(\hat{\beta})$ の最大値は、 $|\beta^T S_{10} (S_{00})^{-1} S_{01} \beta| |\beta^T S_{11} \beta|^{-1}$ が最小値のときであり、そのときの $\hat{\beta}$ は、(付録J)を参考にすれば、(付式29-28)に示す固有方程式の解、つまり、行列 $(S_{11})^{-1} S_{10} (S_{00})^{-1} S_{01}$ の固有値に対応した固有ベクトルのときである。

$$|\lambda S_{11} - S_{10} (S_{00})^{-1} S_{01}| = 0 \rightarrow |\lambda - (S_{11})^{-1} S_{10} (S_{00})^{-1} S_{01}| = 0 \quad \dots$$
 (付式29-28)

(付式29-28)から求める $S_{10} (S_{00})^{-1} S_{01} = \lambda^* S_{11}$ を(付式29-27)に代入して、 $\hat{\beta}$ が行列 $(S_{11})^{-1} S_{10} (S_{00})^{-1} S_{01}$ の固有ベクトルのときの尤度の最大値を求める。

$$\begin{aligned} (L(\hat{\beta}))^{-2/T} &= |S_{00}| (1 - |\beta^T S_{10} (S_{00})^{-1} S_{01} \beta| |\beta^T S_{11} \beta|^{-1}) = |S_{00}| (1 - |\beta^T \lambda^* S_{11} \beta| |\beta^T S_{11} \beta|^{-1}) \\ &= |S_{00}| (1 - \lambda^* |\beta^T S_{11} \beta| |\beta^T S_{11} \beta|^{-1}) = |S_{00}| (1 - \lambda^*) \quad \dots$$
 (付式29-29)

従って、共和分ベクトル $\hat{\beta}$ の数が h 個あるという、帰無仮説 H_2 (h) の下での尤度は、(付式29-29)の尤度の h 個の積になる。

$$(L(H_2))^{-2/T} = \prod_{i=1}^h (L(\hat{\beta}))^{-2/T} = \prod_{i=1}^h |S_{00}| (1 - \lambda_i^*) = |S_{00}| \prod_{i=1}^h (1 - \lambda_i^*) \quad \dots$$
 (付式29-30)

付録29. Johansen(ヨハンセン)の尤度比検定統計量の導出 (5/5)

(D) 尤度比検定統計量を導出する。
尤度比検定統計量は、(付式29-28)で算出される固有値に着目した統計量である。

(付式29-28)は、帰無仮説 $H_2(h)$ のもとで共和分関係の数に相当する h 個の固有値が存在することを示す。 h 個の共和分関係(固有ベクトル)が存在しているのであれば、 h 個の固有ベクトルで回帰式が構成されるということを示し、回帰式に新たに、 $h+1, \dots, n$ 番目の固有ベクトルを加えても(式29-29)の尤度の増加は期待できない。つまり、帰無仮説 $H_2(h)$ のもとでは、固有値 $\lambda_{h+1}, \dots, \lambda_n$ は存在しない。つまり、 $\lambda_{h+1} = \dots = \lambda_n = 0$ であると見なすことができる。

(D1) **トレース検定の検定量**

対立仮説 H_1 の下での尤度は、(付式29-21)より $(L(H_1))^{-2/T} = |\Lambda^\wedge|$ であるが、これは共和分ベクトルの数が h 個あるとする帰無仮説 $H_2(h)$ において、 $h = n$ とする場合に相当する。つまり、(付式29-30)より、

$$(L(H_1))^{-2/T} = (L(H_2(n)))^{-2/T} = |\mathbf{S}_{uu}| \prod_{i=1}^n (1 - \lambda_i^\wedge) \quad \dots \text{(付式29-31)}$$

一方、帰無仮説 $H_2(h)$ の下での尤度は、(付式29-30)で表せた。

$$(L(H_2(h)))^{-2/T} = |\mathbf{S}_{00}| \prod_{i=1}^h (1 - \lambda_i^\wedge) \quad \dots \text{(付式29-30)}$$

従って、尤度比検定量 $Q(H_2(h) | H_1) = L(H_2(h)) / L(H_1)$ は、

$$(Q(H_2(h) | H_1))^{-2/T} = (L(H_2(h)))^{-2/T} / (L(H_1))^{-2/T} = |\mathbf{S}_{00}| \prod_{i=1}^h (1 - \lambda_i^\wedge) / |\mathbf{S}_{00}| \prod_{i=1}^n (1 - \lambda_i^\wedge) = 1 / \left(\prod_{i=h+1}^n (1 - \lambda_i^\wedge) \right)$$

両辺の対数をとって、 $(-2 / T) \ln Q(H_2(h) | H_1) = - \ln \left(\prod_{i=h+1}^n (1 - \lambda_i^\wedge) \right)$

従って、 $-2 \ln Q(H_2(h) | H_1) = -T \sum_{i=h+1}^n (1 - \lambda_i^\wedge) \quad \dots \text{(付式29-32)}$

(D2) **最大固有値検定の検定量**

帰無仮説は(1)のトレース検定の場合と同じで、帰無仮説 $H_2(h)$ の下では h 個の共和分関係があるとする。この帰無仮説の下での尤度は、(付式29-30)になる。

一方、対立仮説 $H_2(h+1)$ の下では $h+1$ 個の共和分関係があるとする。この対立仮説の下での尤度は、(付式29-30)を利用して求まる。

$$(L(H_2(h+1)))^{-2/T} = |\mathbf{S}_{uu}| \prod_{i=1}^{h+1} (1 - \lambda_i^\wedge) \quad \dots \text{(付式29-33)}$$

従って、尤度比検定量 $Q(H_2(h) | H_2(h+1)) = L(H_2(h)) / L(H_2(h+1))$ は、

$$Q(H_2(h) | H_2(h+1))^{2/T} = (L(H_2(h)))^{-2/T} / L(H_2(h+1))^{-2/T} = |\mathbf{S}_{00}| \prod_{i=1}^h (1 - \lambda_i^\wedge) / |\mathbf{S}_{00}| \prod_{i=1}^{h+1} (1 - \lambda_i^\wedge) = 1 / (1 - \lambda_{h+1}^\wedge)$$

両辺の対数をとって、 $(-2 / T) \ln Q(H_2(h) | H_2(h+1)) = - \ln (1 - \lambda_{h+1}^\wedge)$

従って、 $-2 \ln Q(H_2(h) | H_2(h+1)) = -T \ln (1 - \lambda_{h+1}^\wedge) \quad \dots \text{(付式29-33)}$

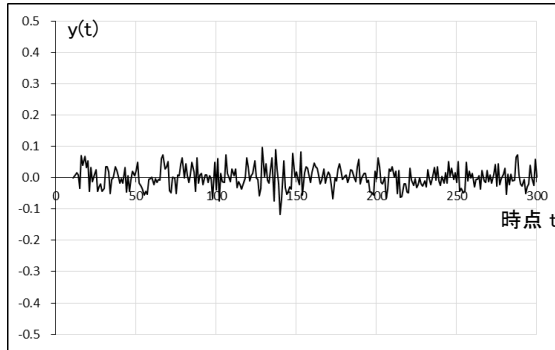
付録30. 条件付き分散不均一モデル (ARCH)

(1) ARCH (1) モデルでの計算例を載せる。

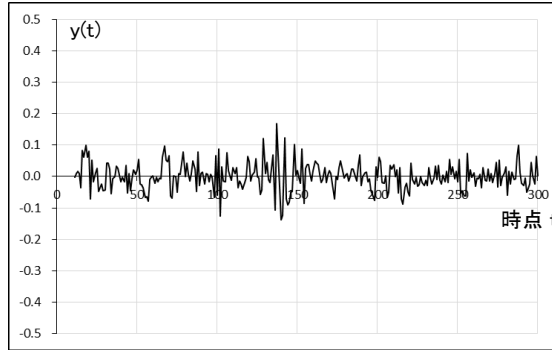
$y(t) = u(t) = h(t)^{1/2} v(t)$ $v(t) \sim \text{i.i.d. } N(0, 1)$ ここでは、正規ホワイトノイズとして、(図3-2)を使う。

$$h(t) = \omega + \alpha_1 y^2(t-1)$$

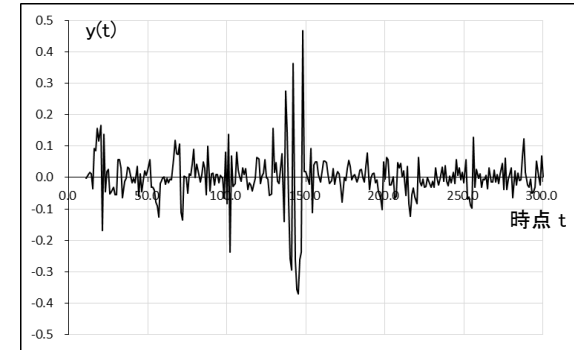
パラメータ α_1 ($0 \leq \alpha_1 < 1$) が 1 に近づくほど、ボラティリティクラスティングの現象が顕著になることが分かる。



(a) $\omega = 0.001$ $\alpha_1 = 0.1$



(b) $\omega = 0.001$ $\alpha_1 = 0.4$



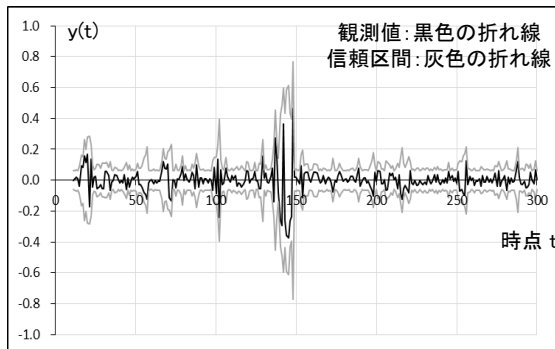
(c) $\omega = 0.001$ $\alpha_1 = 0.7$

(付図30-1) ARCH(1) モデル

(2) ARCH (1) モデルでの1期先の 95 % 区間予測の計算例を載せる。

使用する ARCH(1) モデルは上記の(1)のケース(c)である。

(式21-12)を使用して区間予測の計算を行う。



(付図30-2)

ARCH(1) モデル 95 % 区間予測

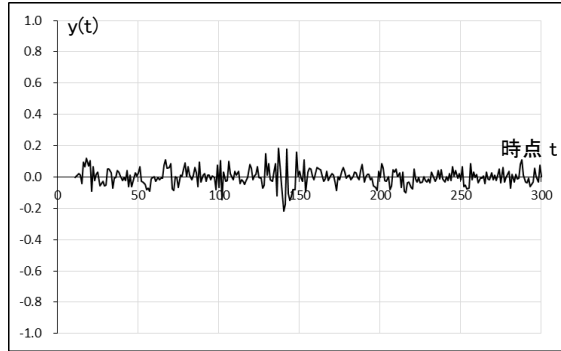
付録31. 条件付き分散不均一モデル (GARCH)

(1) GARCH (1, 1) モデルでの計算例を載せる。

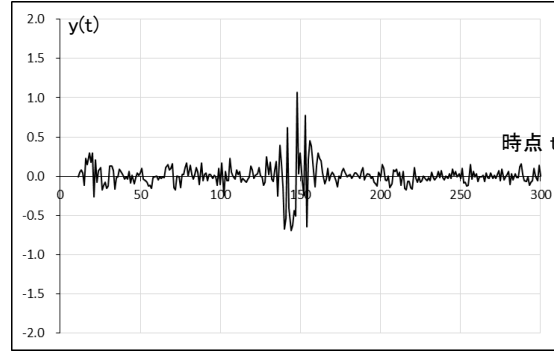
$y(t) = u(t) = h(t)^{1/2} v(t)$ $v(t) \sim \text{i.i.d. } N(0, 1)$ ここでは、正規ホワイトノイズとして、(図3-2)を使う。

$$h(t) = \omega + \alpha_1 y^2(t-1) + \beta_1 h(t-1)$$

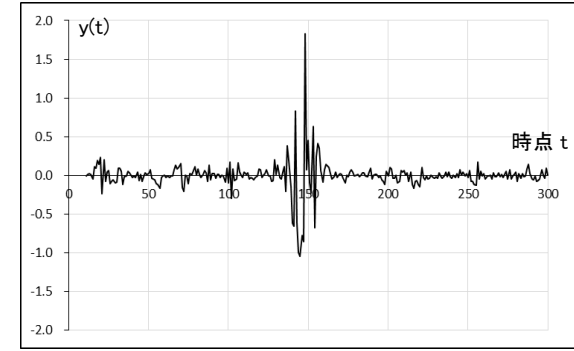
パラメータ $\alpha_1 + \beta_1$ ($0 \leq \alpha_1 + \beta_1 < 1$) が 1 に近づくほど、ボラティリティクラスターingの現象が顕著になることが分かる。。



(a) $\omega = 0.001$ $\alpha_1 = 0.3$ $\beta_1 = 0.3$ ($\alpha_1 + \beta_1 = 0.6$)



(b) $\omega = 0.001$ $\alpha_1 = 0.3$ $\beta_1 = 0.6$ ($\alpha_1 + \beta_1 = 0.9$)



(c) $\omega = 0.001$ $\alpha_1 = 0.6$ $\beta_1 = 0.3$ ($\alpha_1 + \beta_1 = 0.9$)

(付図31-1) GARCH(1, 1) モデル

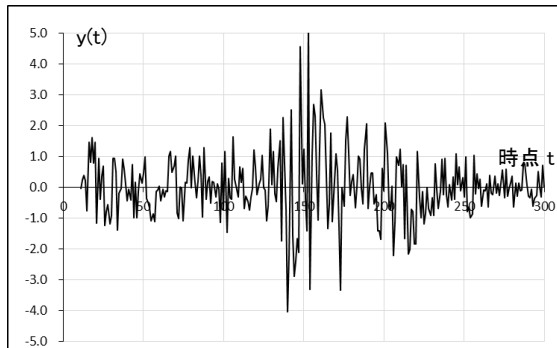
(2) $\alpha_1 + \beta_1 = 1$ のとき

条件付き分散 (ボラティリティ) $h(t)$ は恒久的な持続性を持つことになる。

(1) の GARCH(1, 1) モデルで、 $\alpha_1 + \beta_1 = 1$ とした場合を (付図31-2) に示す。

(付図31-1) に比べて大きな振れ幅は直ちに減衰せず維持されていることが分かる。

このようなモデルは、IGARCH(Integrated GARCH) モデルと呼ばれている。



(付図31-2) GARCH(1, 1) モデル

$$\omega = 0.001 \quad \alpha_1 = 0.1 \quad \beta_1 = 0.9$$

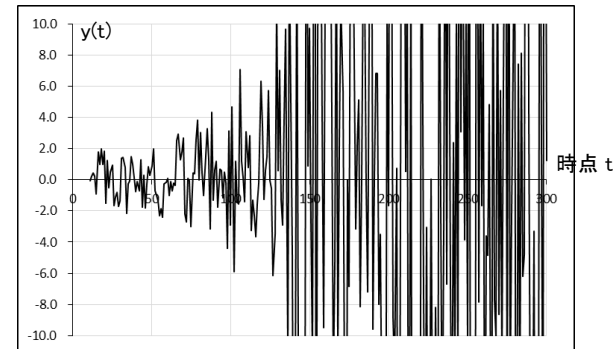
$$(\alpha_1 + \beta_1 = 1.0)$$

(3) $\alpha_1 + \beta_1 > 1$ のとき

条件付き分散 (ボラティリティ) $h(t)$ は非定常過程になる (発散する)。

(1) の GARCH(1, 1) モデルで、 $\alpha_1 + \beta_1 > 1$ とした場合を (付図31-3) に示す。

発散するモデルになる。



(付図31-3) GARCH(1, 1) モデル

$$\omega = 0.001 \quad \alpha_1 = 0.1 \quad \beta_1 = 0.93$$

$$(\alpha_1 + \beta_1 = 1.03 > 1)$$

付録32. VECモデルのパラメータ数

VEC モデルでは、条件付き分散共分散行列 $H(t)$ は以下のように表される。

$$\text{vech}(H(t)) = \mathbf{C} + \mathbf{B} \text{vech}(H(t-1)) + \mathbf{A} \text{vech}(\mathbf{u}(t-1) \mathbf{u}^T(t-1)) \quad \text{ただし、vech 作用素に関しては、(付録N)を参照のこと}$$

(1) モデル次数が 2 のときの VEC モデルのパラメータ数求める。

$$H(t) = \begin{pmatrix} h_{11}(t) & h_{12}(t) \\ h_{21}(t) & h_{22}(t) \end{pmatrix} \text{ で、 } \text{vech}(H(t-1)) = \begin{pmatrix} h_{11}(t-1) \\ h_{21}(t-1) \\ h_{22}(t-1) \end{pmatrix}$$

$$\mathbf{u}(t-1) \mathbf{u}^T(t-1) = \begin{pmatrix} u_1(t-1) \\ u_2(t-1) \end{pmatrix} \begin{pmatrix} u_1(t-1) & u_2(t-1) \end{pmatrix} = \begin{pmatrix} u_1(t-1)^2 & u_1(t-1) u_2(t-1) \\ u_1(t-1) u_2(t-1) & u_2(t-1)^2 \end{pmatrix} \text{ で、 } \text{vech}(\mathbf{u}(t-1) \mathbf{u}^T(t-1)) = \begin{pmatrix} u_1(t-1)^2 \\ u_1(t-1) u_2(t-1) \\ u_2(t-1)^2 \end{pmatrix}$$

$$\text{従って、} \begin{pmatrix} h_{11}(t) \\ h_{21}(t) \\ h_{22}(t) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} h_{11}(t-1) \\ h_{21}(t-1) \\ h_{22}(t-1) \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} u_1(t-1)^2 \\ u_1(t-1) u_2(t-1) \\ u_2(t-1)^2 \end{pmatrix} \text{ というモデル式になる。}$$

よって、2次の VEC モデルのパラメータ数は、ベクトル $\mathbf{c}$ で 3 個、行列 $\mathbf{B}$ で 9 個、行列 $\mathbf{A}$ で 9 個の、合計 $3 + 9 + 9 = 21$ 個 となる。

(2) モデル次数が 3 のときの VEC モデルのパラメータ数求める。

$$H(t) = \begin{pmatrix} h_{11}(t) & h_{12}(t) & h_{13}(t) \\ h_{21}(t) & h_{22}(t) & h_{23}(t) \\ h_{31}(t) & h_{32}(t) & h_{33}(t) \end{pmatrix} \text{ で、 } \text{vech}(H(t-1)) = \begin{pmatrix} h_{11}(t-1) \\ h_{21}(t-1) \\ h_{31}(t-1) \\ h_{22}(t-1) \\ h_{32}(t-1) \\ h_{33}(t-1) \end{pmatrix}$$

$$\mathbf{u}(t-1) \mathbf{u}^T(t-1) = \begin{pmatrix} u_1(t-1) \\ u_2(t-1) \\ u_3(t-1) \end{pmatrix} \begin{pmatrix} u_1(t-1) & u_2(t-1) & u_3(t-1) \end{pmatrix} = \begin{pmatrix} u_1(t-1)^2 & u_1(t-1) u_2(t-1) & u_1(t-1) u_3(t-1) \\ u_1(t-1) u_2(t-1) & u_2(t-1)^2 & u_2(t-1) u_3(t-1) \\ u_1(t-1) u_3(t-1) & u_2(t-1) u_3(t-1) & u_3(t-1)^2 \end{pmatrix} \text{ で、}$$

$$\text{vech}(\mathbf{u}(t-1) \mathbf{u}^T(t-1)) = \begin{pmatrix} u_1(t-1)^2 \\ u_1(t-1) u_2(t-1) \\ u_1(t-1) u_3(t-1) \\ u_2(t-1)^2 \\ u_2(t-1) u_3(t-1) \\ u_3(t-1)^2 \end{pmatrix}$$

$$\text{従って、} \begin{pmatrix} h_{11}(t) \\ h_{21}(t) \\ h_{31}(t) \\ h_{22}(t) \\ h_{32}(t) \\ h_{33}(t) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{41} & b_{41} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{51} & b_{51} & b_{54} & b_{55} & b_{56} \\ b_{61} & b_{61} & b_{61} & b_{64} & b_{65} & b_{66} \end{pmatrix} \begin{pmatrix} h_{11}(t-1) \\ h_{21}(t-1) \\ h_{31}(t-1) \\ h_{22}(t-1) \\ h_{32}(t-1) \\ h_{33}(t-1) \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{41} & b_{41} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{51} & b_{51} & b_{54} & b_{55} & b_{56} \\ b_{61} & b_{61} & b_{61} & b_{64} & b_{65} & b_{66} \end{pmatrix} \begin{pmatrix} u_1(t-1)^2 \\ u_1(t-1) u_2(t-1) \\ u_1(t-1) u_3(t-1) \\ u_2(t-1)^2 \\ u_2(t-1) u_3(t-1) \\ u_3(t-1)^2 \end{pmatrix} \text{ というモデル式になる。}$$

よって、3次の VEC モデルのパラメータ数は、ベクトル $\mathbf{c}$ で 6 個、行列 $\mathbf{B}$ で 36 個、行列 $\mathbf{A}$ で 36 個の、 $6 + 36 + 36 = 78$ 個 となる。

付録33. コピュラ (Copula, 接合分布関数)

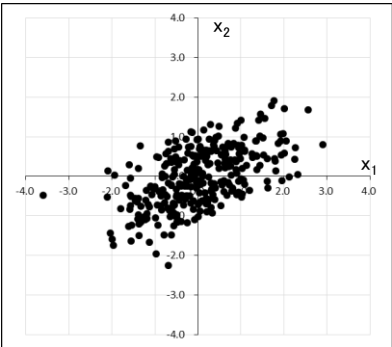
(1) **コピュラ**とは
相関係数は確率変数間の線形関係を表す指標の一つであるが、コピュラは多変数の同時分布関数とそれぞれの周辺分布関数との関係を表す関数である。
言い換えれば、コピュラは、複数の周辺分布をつなぎ合わせて(入力して)、同時分布関数を作る(出力する)関数である。日本語では**接合分布関数**と呼ばれる。
(注)コピュラはラテン語で「連結」を意味し、例えば、「He is Japanese. (彼は日本人です。)」という文章では、「He(彼)」と「Japanese(日本人)」を結びつける「is(は、です)」がコピュラにあたる。

(2) 定義
全ての周辺分布関数が区間 [0, 1] 上の一様分布(別途資料「いろいろな確率分布」を参照)に等しい n 変量同時分布関数は、 n 変量コピュラと言われる。

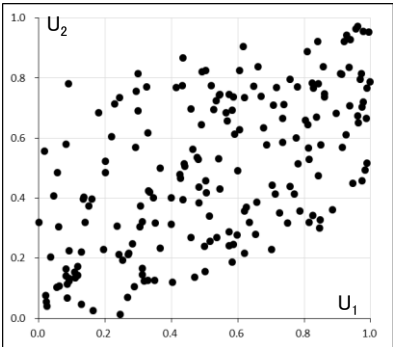
(3) **スクラーの定理** (Sklar's Theorem)
スクラーの定理は、 n 変量同時分布関数 F は、コピュラ(接合分布関数) C を用いて、以下のように表すことができるという定理である。
$$F(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$$
 ただし、 $F_i(x_i)$ は確率変数 X_i ($n = 1, 2, \dots, n$) の周辺分布関数 $\dots$ (付式32-1)
この式は、同時分布関数が、各変数の周辺分布関数とコピュラを用いて一意に分割できることを示している。
つまり、同時分布関数 = コピュラ + 周辺分布関数 であるとみなすことができ、個別の周辺分布関数の持っている情報とは別に、コピュラは同周辺分布関数の間の依存構造を表していると見ることができる。
また、(付式32-1)は、モデルとして、複雑な依存構造を持つ多変量分布関数を当てはめることができることを示している。
(例) ・正規分布(周辺分布) + ガウシアンコピュラ = 多変量元正規分布(同時分布)
 ・正規分布(周辺分布) + t コピュラ = 複雑な分布(同時分布)

(4) コピュラ算出の定式化(2変数の場合)
確率変数 X_1 と X_2 の同時分布関数 : $F(x_1, x_2)$, 確率変数 X_1 の周辺分布関数 : $F_1(x_1)$, 確率変数 X_2 の周辺分布関数 : $F_2(x_2)$
のとき、スクラーの定理より、以下の関係が成り立つような、コピュラ C が存在する。
$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$
 ここで、 $F_1(x_1) = U_1$, $F_2(x_2) = U_2$ は一様分布
コピュラは以下の式で求まる ($^{-1}$ は逆関数であることを表す)。
$$F_1(x_1) = U_1 \text{ より } x_1 = F_1^{-1}(U_1), \quad F_2(x_2) = U_2 \text{ より } x_2 = F_2^{-1}(U_2) \text{ であるから、コピュラは、 } C(U_1, U_2) = F(F_1^{-1}(U_1), F_2^{-1}(U_2))$$

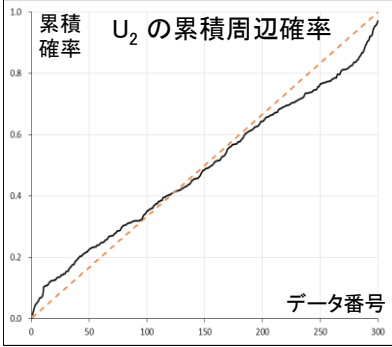
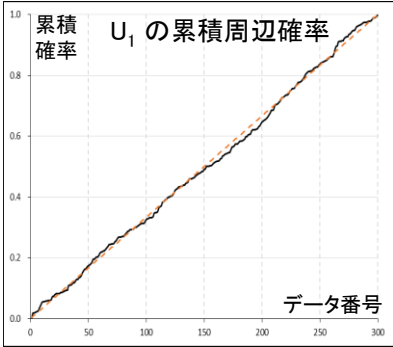
(5) 計算例
確率変数 X_1 , X_2 として、それぞれ 300 点のデータを使用。
 $f_1(x_1)$: X_1 の周辺確率密度関数、正規分布(使用データの平均 = 0.017, 分散 = 1.010) $\rightarrow$ (累積)周辺分布関数 $F_1(x_1)$ の導出 (Excel の NORMSDIST 関数を使用)
 $f_2(x_2)$: X_2 の周辺確率密度関数、正規分布(使用データの平均 = -0.031, 分散 = 0.547) $\rightarrow$ (累積)周辺分布関数 $F_2(x_2)$ の導出 (Excel の NORMSDIST 関数を使用)
確率変数 X_1 と X_2 の相関係数は、0.577 (Excel の CORREL 関数を使用)
(付図32-1)に、同時確率分布 $F(x_1, x_2)$ を示す(確率変数 X_1 と X_2 の相関関係が見て取れる)。
(付図32-2)に、コピュラ $C(U_1, U_2)$ を示す。合わせて、(付図32-3)では、 $U_1 = F_1(x_1)$ と $U_2 = F_2(x_2)$ が区間 [0, 1] 上の一様分布に従っていることを確認できる。



(付図32-1) 同時確率分布 $F(x_1, x_2)$
相関係数 0.577



(付図32-2) コピュラ $C(U_1, U_2)$



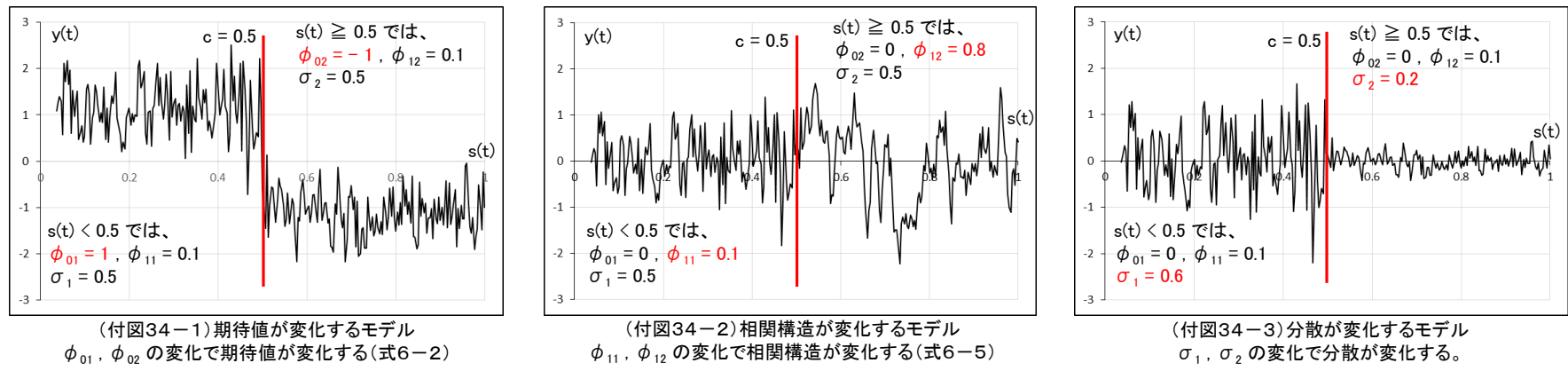
(付図32-3) 累積周辺確率 U_1, U_2
使用データ 300 点の周辺分布関数の値を大きさの順に並べ、 U_1, U_2 の一様分布を確認

付録34. 閾値モデルと平滑推移モデルの計算例

(1) 閾値モデル (式24-1)で表される閾値モデルの例を載せる。

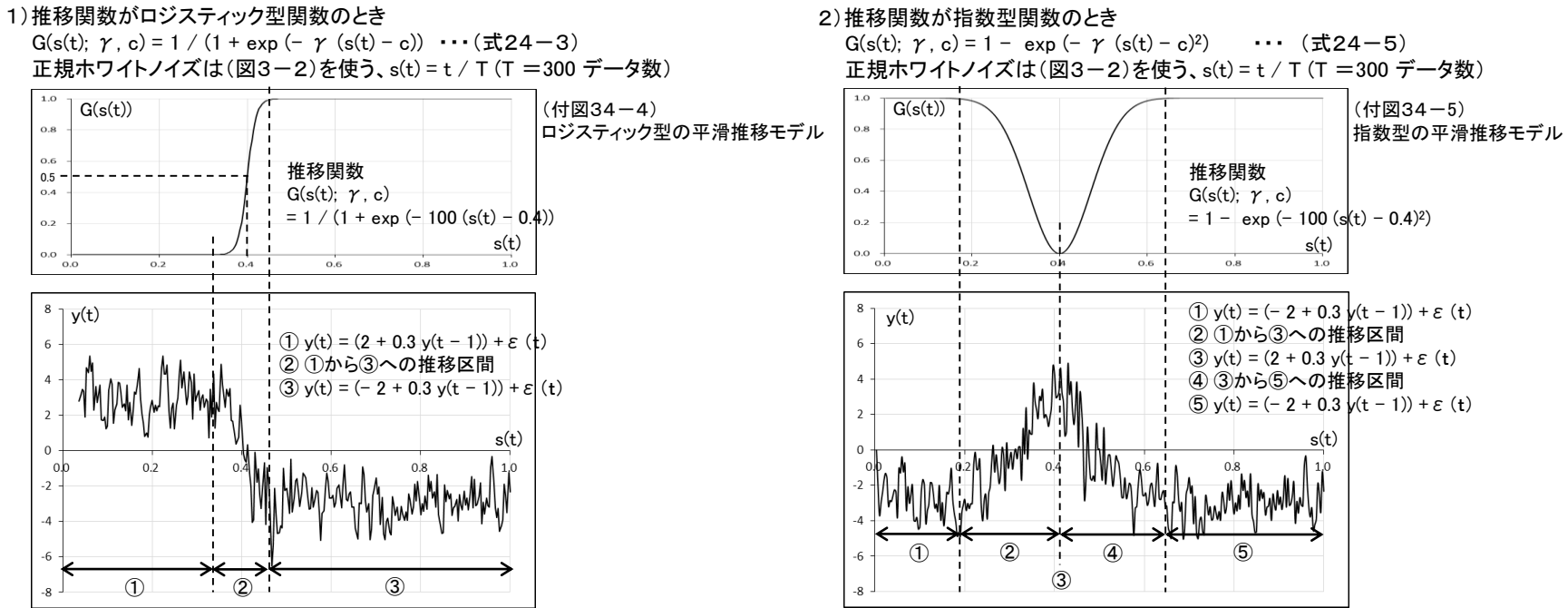
$$\left. \begin{aligned} s(t) < c \text{ のとき、 } y(t) &= \phi_{01} + \phi_{11} y(t-1) + \sigma_1 \varepsilon(t), \quad \varepsilon(t) \sim \text{i.i.d.}(0, 1) \\ s(t) \geq c \text{ のとき、 } y(t) &= \phi_{02} + \phi_{12} y(t-1) + \sigma_2 \varepsilon(t) \end{aligned} \right\} \dots \text{(式24-1)}$$

なお、正規ホワイトノイズは(図3-2)を使う。 また、 $s(t) = t / T$ ($T = 300$ データ数) とし、 $s(t) = 150 / 300 = 0.5$ つまり、 $c = 0.5$ とした。 $s(t) = 0.5$ でモデルが切り替わる。



(2) 平滑推移モデル (式24-2)で表される平滑推移モデルの例を載せる。

$$y(t) = (\phi_{01} + \phi_{11} y(t-1)) (1 - G(s(t))) + (\phi_{02} + \phi_{12} y(t-1)) G(s(t)) + \sigma \varepsilon(t), \quad \varepsilon(t) \sim \text{i.i.d.}(0, 1) \quad \dots \text{(式24-2)}$$



付録35. マルコフ転換モデルの計算例

(付録36)、(付録37)、および(付録38)の計算例を示すために、(付図35－1)で示すマルコフ転換モデルを用意する。

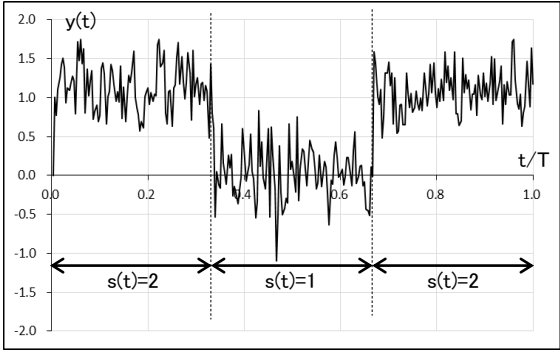
- ① $s(t) = 1$ のとき、
 時系列データ $y(t) = \phi_{01} + \phi_{11} y(t-1) + \sigma_a \varepsilon(t)$ ここで、 $\phi_{01} = 0.0$, $\phi_{11} = 0.1$, $\sigma_a = 0.3$ として、 $y(t) = 0.0 + 0.1 y(t-1) + 0.3 \varepsilon(t)$
 確率分布(確率密度関数)は、

$$f_1(y(t)) = \{1 / (2 \pi \sigma_1^2)^{1/2}\} \exp \{- (y(t) - \mu_1)^2 / (2 \sigma_1^2)\} = f_1(y(t)) = \{1 / (2 \pi (0.091)^2)^{1/2}\} \exp \{- (y(t) - 0.0)^2 / (2 (0.091)^2)\} \quad \cdots \text{ (付式35－1)}$$
 ここで、(式6－2)より、 $\mu_1 = E[y(t)] = \phi_{01} / (1 - \phi_{11}) = 0.0 / (1 - 0.1) = 0.0$
 一方、(式6－3)と(式6－5)より、 $\sigma_1 = V[y(t)] = \sigma_a^2 / (1 - \phi_{11}^2) = 0.3^2 / (1 - 0.1^2) = 0.091$
- ② $s(t) = 2$ のとき、
 時系列データ $y(t) = \phi_{02} + \phi_{12} y(t-1) + \sigma_b \varepsilon(t)$ ここで、 $\phi_{02} = 1.0$, $\phi_{12} = 0.1$, $\sigma_b = 0.3$ として、 $y(t) = 1.0 + 0.1 y(t-1) + 0.3 \varepsilon(t)$
 確率分布(確率密度関数)は、

$$f_2(y(t)) = \{1 / (2 \pi \sigma_2^2)^{1/2}\} \exp \{- (y(t) - \mu_2)^2 / (2 \sigma_2^2)\} = f_2(y(t)) = \{1 / (2 \pi (0.091)^2)^{1/2}\} \exp \{- (y(t) - 1.11)^2 / (2 (0.091)^2)\} \quad \cdots \text{ (付式35－2)}$$
 ここで、(式6－2)より、 $\mu_2 = E[y(t)] = \phi_{02} / (1 - \phi_{12}) = 1.0 / (1 - 0.1) = 1.111$
 一方、(式6－3)と(式6－5)より、 $\sigma_2 = V[y(t)] = \sigma_b^2 / (1 - \phi_{12}^2) = 0.3^2 / (1 - 0.1^2) = 0.091$
 ただし、 $\varepsilon(t) \sim \text{i.i.d. } N(0, 1)$ ここでは、(図3－2)の正規のホワイトノイズ従うとする。

時点 $t = 1 \sim T$ ($T = 300$) において、区間 $t = 1 \sim 100$ と 区間 $t = 201 \sim 300$ では $s(t) = 1$, 区間 $t = 101 \sim 200$ では $s(t) = 2$ としたときのマルコフ転換モデルの時系列データ $y(t)$ を(付図37－1)に示す。

これは、状態 $s(t)$ を既知としたときの時系列データであるが、(付録36)、(付録37)、および(付録38)では、この状態 $s(t)$ が未知であるとし、(付図35－1)の時系列データから、状態 $s(t)$ の確率を推定する計算例を示す。さらに、(付録36)、(付録37)、および(付録38)では、状態 $s(t)$ の遷移確率が(付式35－3)に示すように既知であるとして計算している。



(付図35－1) マルコフ転換モデル

$$s(t) \text{ の遷移行列 } \mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \\ 0.1 & 0.9 \end{bmatrix} \quad \cdots \text{ (付式35－3)}$$

付録36. マルコフ転換モデル（フィルター化確率の導出）

(1) フィルター化確率の導出 (状態数 = 2 のとき)
フィルター化確率の導出には、ベイズ更新の式(付式36-1)、(表36-1)を利用する(別途資料「ベイズ統計の概要」を参照)。

$P(A|B) = P(A)P(B|A) / P(B) \rightarrow \text{事後確率} = \text{事前確率} \times \text{尤度} / \text{定数} \quad \dots \text{(付式36-1)}$

この式は、**尤度の情報を得て、事前確率を事後確率に更新する式である**と捉えることができる。

ベイズ統計で、原因： $A = (s(t) = i) \ (i = 1 \text{ or } 2)$ 結果： $B = (y(t))$ であるとする。
なお、時点 t までの時系列データの集合を $\Omega(t) = \{y(t), y(t-1), \dots, y(1)\}$ で表す。

$s(t) = 1$ のとき、 事前確率： $P(A) = P(s(t) = 1) = P(s(t) = 1 | \Omega(t-1))$
事後確率： $P(A|B) = P(s(t) = 1 | y(t)) = P(s(t) = 1 | \Omega(t))$
尤度： $P(B|A) = P(y(t) | s(t) = 1) = P(y(t) | \Omega(t-1), s(t) = 1)$
 $\rightarrow f(y(t) | \Omega(t-1), s(t) = 1) = f_1(y(t) | \Omega(t-1))$

$s(t) = 2$ のとき、 事前確率： $P(A) = P(s(t) = 2) = P(s(t) = 2 | \Omega(t-1))$
事後確率： $P(A|B) = P(s(t) = 2 | y(t)) = P(s(t) = 2 | \Omega(t))$
尤度： $P(B|A) = P(y(t) | s(t) = 2) = P(y(t) | \Omega(t-1), s(t) = 2)$
 $\rightarrow f(y(t) | \Omega(t-1), s(t) = 2) = f_2(y(t) | \Omega(t-1))$

そして、 $s(t) = 1$ であろうが、 $s(t) = 2$ であろうが、つまり、 $s(t)$ の値に関係なく、次式が成り立つ。
定数： $P(B) = P(y(t) | \Omega(t-1)) \rightarrow f(y(t) | \Omega(t-1))$
これは、 $f_1(y(t) | \Omega(t-1))$ と $f_2(y(t) | \Omega(t-1))$ の混合分布である

従って、(付式36-1)は下記の式に書き直すことができる。

$s(t) = 1$ のとき、 $P(s(t) = 1 | \Omega(t)) = P(s(t) = 1 | \Omega(t-1)) f_1(y(t) | \Omega(t-1)) / f(y(t) | \Omega(t-1)) \quad \dots \text{(付式36-2)}$

$s(t) = 2$ のとき、 $P(s(t) = 2 | \Omega(t)) = P(s(t) = 2 | \Omega(t-1)) f_2(y(t) | \Omega(t-1)) / f(y(t) | \Omega(t-1)) \quad \dots \text{(付式36-3)}$

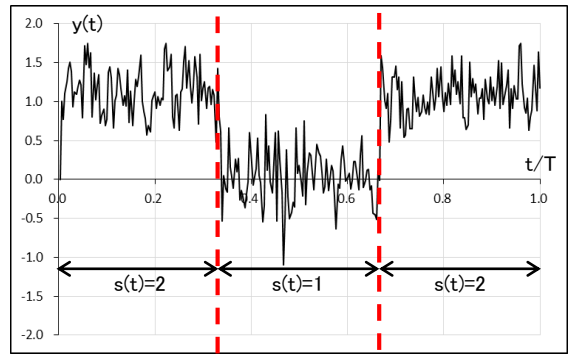
各々の確率は、混合分布 $f(y(t) | \Omega(t-1))$ のうち、各々の状態が寄与する分布の割合を求めたものであると解釈できる。
これらの式を使って、 $s(t) = 1$ では、 $P(s(t) = 1 | \Omega(t-1))$ を $P(s(t) = 1 | \Omega(t))$ に更新する。
 $s(t) = 2$ では、 $P(s(t) = 2 | \Omega(t-1))$ を $P(s(t) = 2 | \Omega(t))$ に更新する。

(表36-1)ベイズの定理の説明

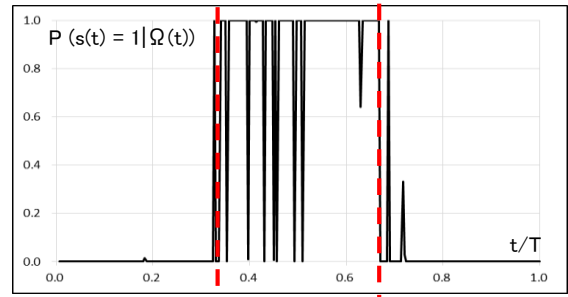
確率	名称	説明
$P(A B)$	事後確率	結果B が得られたときに、その原因が A である確率
$P(B A)$	尤度(ゆうど)	原因Aのもとで、結果Bが得られる確率
$P(A)$	事前確率	結果Bが得られる前の原因Aが発生する確率
$P(B)$	定数扱い	原因Aの有無にかかわらず、結果Bが発生する確率

(2) フィルター化確率の計算例

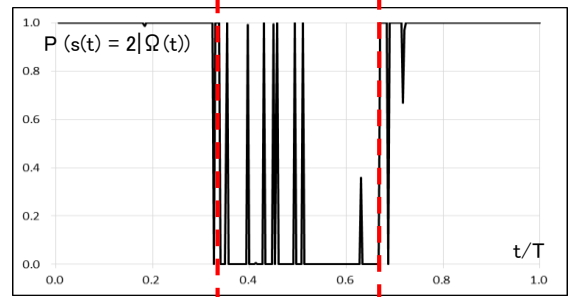
(付図35-1)のマルコフ転換モデルのフィルター化確率を求める。
初期値 $t = 1$ のとき、 $s(1) = 1$ である確率と $s(1) = 2$ である確率は、それぞれ、(付式0-6)と(付式0-7)より、
 $P(s(1) = 1 | \Omega(0)) = (1 - p_{22}) / (2 - p_{11} - p_{22}) = (1 - 0.9) / (2 - 0.8 - 0.9) = 0.333$
 $P(s(1) = 2 | \Omega(0)) = (1 - p_{11}) / (2 - p_{11} - p_{22}) = (1 - 0.8) / (2 - 0.8 - 0.9) = 0.667$
これらの初期値を使って、本文 25. の手順に従って、(付図35-1)のマルコフ転換モデルのフィルター化確率 $P(s(t) = 1 | \Omega(t))$ (付図36-1)と $P(s(t) = 2 | \Omega(t))$ (付図36-2)を求めた。
 $s(t) = 1$ と $s(t) = 2$ の状態ごとの確率が推定できていることが分かる。



(付図35-1)マルコフ転換モデル



(付図36-1)フィルター化確率 $P(s(t) = 1 | \Omega(t))$



(付図36-2)フィルター化確率 $P(s(t) = 2 | \Omega(t))$

付録37. マルコフ転換モデル（平滑化確率の導出）

(1) 平滑化確率の導出 (状態数 = 2 のとき)

1) 全観測値 $\Omega(T)$ を所与として、一般に、 $P(s(t) = 1 \mid \Omega(T)) = P(s(t) = 1, s(t+1) = 1 \mid \Omega(T)) + P(s(t) = 1, s(t+1) = 2 \mid \Omega(T)) \dots$ (付式37-1)
この式は、(左辺) $s(t) = 1$ である確率は、(右辺) $s(t) = 1$ から $s(t+1) = 1$ への遷移確率と $s(t) = 1$ から $s(t+1) = 2$ への遷移確率の和 に等しいことを表している。
同様に、 $P(s(t) = 2 \mid \Omega(T)) = P(s(t) = 1, s(t+1) = 2 \mid \Omega(T)) + P(s(t) = 2, s(t+1) = 2 \mid \Omega(T)) \dots$ (付式37-2)
 $P(s(t) = 1 \mid \Omega(T))$ を $s(t) = 1$ の平滑化確率、 $P(s(t) = 2 \mid \Omega(T))$ を $s(t) = 2$ の平滑化確率という。

2) ベイズの定理を(付式37-2)に示す(別途資料「ベイズ統計の概要」を参照)。
 $P(A \cap B) = P(B) P(A \mid B) \rightarrow$ 同時確率 = 周辺確率 $\times$ 条件付き確率 $\dots$ (付式37-2)
この式は、事象 A と事象 B の同時確率は、事象 B が起こる確率(周辺確率)に事象 B が起こる条件のもとで事象 A が起こる確率(条件付き確率)をかけたものであると捉えることができる。

事象 A = $(s(t) = j \mid \Omega(T))$ ($j = 1$ or 2) 事象 B = $(s(t+1) = k \mid \Omega(T))$ ($k = 1$ or 2) であると考える。
なお、時点 t までの時系列データの集合を $\Omega(t) = \{y(t), y(t-1), \dots, y(1)\}$ で表す。
(付式37-2)の左辺: $P(A \cap B) = P((s(t) = j \mid \Omega(T)) \cap (s(t+1) = k \mid \Omega(T))) = P(s(t) = j, s(t+1) = k \mid \Omega(T))$
(付式37-2)の右辺: $P(B) P(A \mid B) = P(s(t+1) = k \mid \Omega(T)) P((s(t) = j \mid \Omega(T)) \mid (s(t+1) = k \mid \Omega(T)))$
 $= P(s(t+1) = k \mid \Omega(T)) P(s(t) = j \mid s(t+1) = k, \Omega(T))$
 $\Omega(t)$ のもとで $s(t+1) = k$ であると判明しており、全観測値 $\Omega(T)$ は不要である。よって、
 $P(B) P(A \mid B) = P(s(t+1) = k \mid \Omega(t)) P(s(t) = j \mid s(t+1) = k, \Omega(t))$
従って、 $P(s(t) = j, s(t+1) = k \mid \Omega(T)) = P(s(t+1) = k \mid \Omega(t)) P(s(t) = j \mid s(t+1) = k, \Omega(t)) \dots$ (付式37-3)

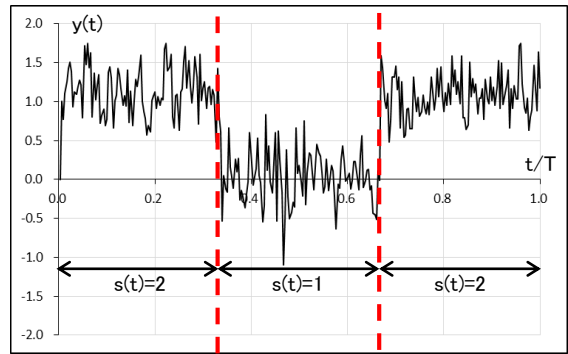
3) 改めて、ベイズ更新の式を(付式37-4)に示す(別途資料「ベイズ統計の概要」を参照)。
 $P(A \mid B) = P(A) P(B \mid A) / P(B) \rightarrow$ 事後確率 = 事前確率 $\times$ 尤度 / 定数 $\dots$ (付式37-4)
ベイズ統計で、原因: A = $(s(t) = j)$ ($j = 1$ or 2) 結果: B = $(s(t+1) = k)$ ($k = 1$ or 2) であると考える。
事前確率: $P(A) = P(s(t) = j) = P(s(t) = j \mid \Omega(t))$
事後確率: $P(A \mid B) = P(s(t) = j \mid s(t+1) = k) = P(s(t) = j \mid s(t+1) = k, \Omega(t))$
尤度: $P(B \mid A) = P(s(t+1) = k \mid s(t) = j) = p_{jk}$ これは、 $s(t) = j$ から、 $s(t+1) = k$ への遷移確率
定数: $P(B) = P(s(t+1) = k) = P(s(t+1) = k \mid \Omega(t))$
従って、 $P(s(t) = j \mid s(t+1) = k, \Omega(t)) = P(s(t) = j \mid \Omega(t)) p_{jk} / P(s(t+1) = k \mid \Omega(t)) \dots$ (付式37-5)

4) (付式37-5)を(付式37-3)に代入して、
 $P(s(t) = j, s(t+1) = k \mid \Omega(T)) = P(s(t+1) = k \mid \Omega(t)) P(s(t) = j \mid \Omega(t)) p_{jk} / P(s(t+1) = k \mid \Omega(t)) \dots$ (付式37-6)

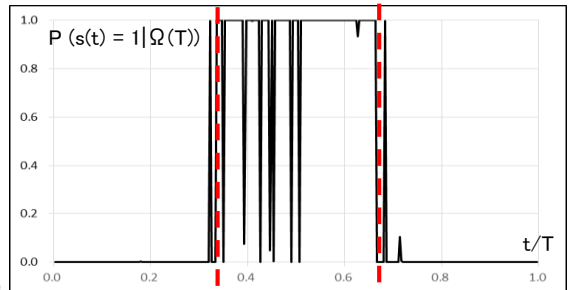
5) 平滑化確率を導出する。
① $s(t) = 1$ の平滑化確率の導出 (付式37-1)に、(式37-6)を代入する。
 $P(s(t) = 1 \mid \Omega(T)) = P(s(t+1) = 1 \mid \Omega(T)) P(s(t) = 1 \mid \Omega(t)) p_{11} / P(s(t+1) = 1 \mid \Omega(t))$
 $+ P(s(t+1) = 2 \mid \Omega(T)) P(s(t) = 1 \mid \Omega(t)) p_{12} / P(s(t+1) = 2 \mid \Omega(t)) \dots$ (付37-7)
② $s(t) = 2$ の平滑化確率の導出 (付式37-2)に、(式37-6)を代入する。
 $P(s(t) = 2 \mid \Omega(T)) = P(s(t+1) = 1 \mid \Omega(T)) P(s(t) = 2 \mid \Omega(t)) p_{21} / P(s(t+1) = 1 \mid \Omega(t))$
 $+ P(s(t+1) = 2 \mid \Omega(T)) P(s(t) = 2 \mid \Omega(t)) p_{22} / P(s(t+1) = 2 \mid \Omega(t)) \dots$ (付37-8)

(2) 平滑化確率の計算例

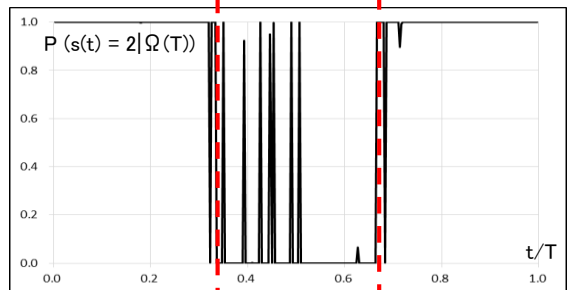
(付図35-1)のマルコフ転換モデルの平滑化確率を求める。
本文 26. の手順に従って、(付図35-1)のマルコフ転換モデルの平滑化確率 $P(s(t) = 1 \mid \Omega(T))$ (付図37-1)と $P(s(t) = 2 \mid \Omega(T))$ (付図37-2)を求めた。
フィルター化確率(付図36-1)、(付図36-2)と平滑化確率(付図37-1)、(付図37-2)を対比すると分かるように、名前が示すように、平滑化確率の推測値の方がフィルター化確率の推測値に比べて滑らかになっていることが分かる。
平滑化確率の推測値が、観測したすべての時系列データを使用した推測値になっていることの結果であると言える。



(付図35-1)マルコフ転換モデル



(付図37-1)平滑化確率 $P(s(t) = 1 \mid \Omega(T))$

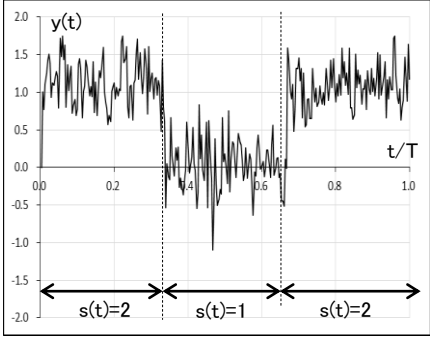


(付図37-2)平滑化確率 $P(s(t) = 2 \mid \Omega(T))$

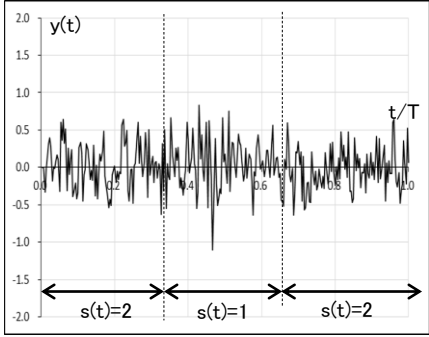
付録38. マルコフ転換モデル (いろいろなマルコフ転換モデルの平滑化確率の計算例)

様々なマルコフ転換モデルの平滑化確率の計算例を載せる。

(1) ケース1 (再掲載、 ϕ_{01} と ϕ_{02} が異なる) (2) ケース2 (全てのパラメータが同一)
 $s(t) = 1 : \phi_{01} = 0.0 \quad \phi_{11} = 0.1 \quad \sigma_a = 0.3$
 $s(t) = 2 : \phi_{02} = 1.0 \quad \phi_{12} = 0.1 \quad \sigma_b = 0.3$
 $s(t) = 1 : \phi_{01} = 0.0 \quad \phi_{11} = 0.1 \quad \sigma_a = 0.3$
 $s(t) = 2 : \phi_{02} = 0.0 \quad \phi_{12} = 0.1 \quad \sigma_b = 0.3$

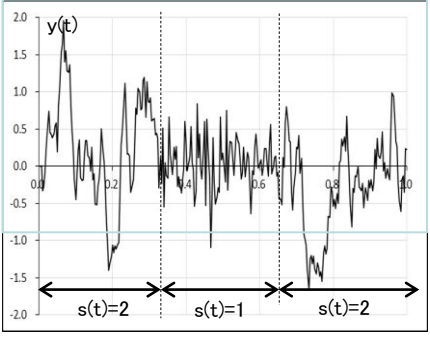


(付図35-1) マルコフ転換モデル



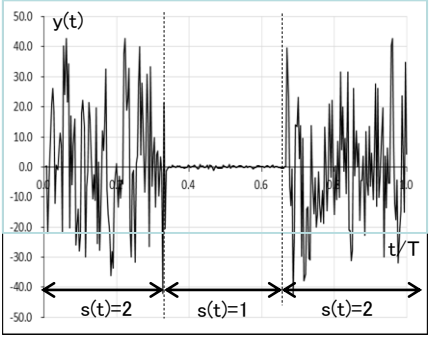
(付図38-1) マルコフ転換モデル

(3) ケース3 (ϕ_{11} と ϕ_{12} が異なる)
 $s(t) = 1 : \phi_{01} = 0.0 \quad \phi_{11} = 0.1 \quad \sigma_a = 0.3$
 $s(t) = 2 : \phi_{02} = 0.0 \quad \phi_{12} = 0.9 \quad \sigma_b = 0.3$

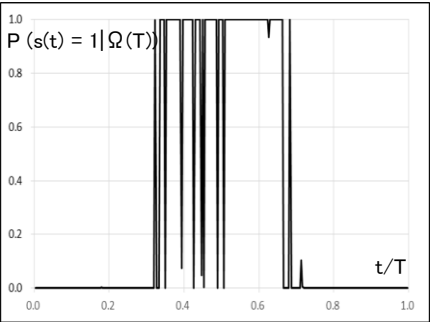


(付図38-4) マルコフ転換モデル

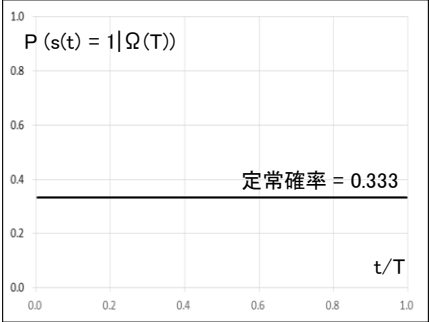
(4) ケース4 (σ_a と σ_b が異なる)
 $s(t) = 1 : \phi_{01} = 0.0 \quad \phi_{11} = 0.1 \quad \sigma_a = 0.3$
 $s(t) = 2 : \phi_{02} = 0.0 \quad \phi_{12} = 0.1 \quad \sigma_b = 2.0$



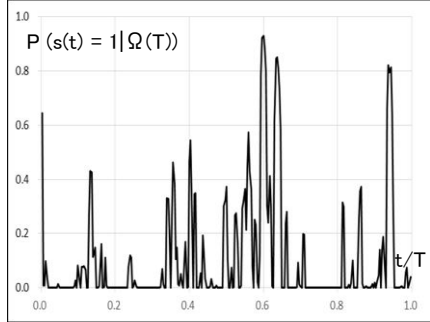
(付図38-7) マルコフ転換モデル



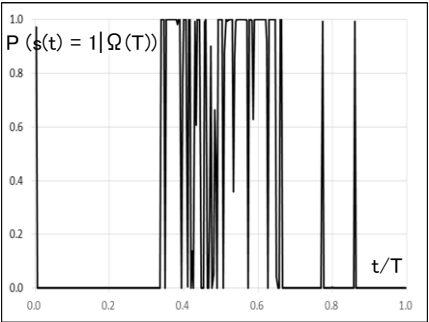
(付図37-1) 平滑化確率 $P(s(t) = 1 | \Omega(T))$



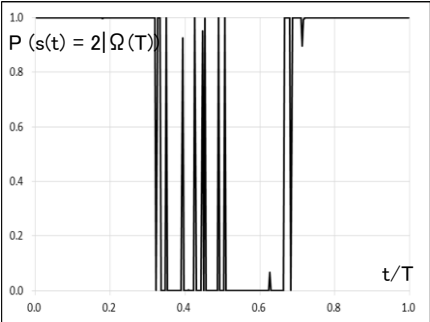
(付図38-2) 平滑化確率 $P(s(t) = 1 | \Omega(T))$



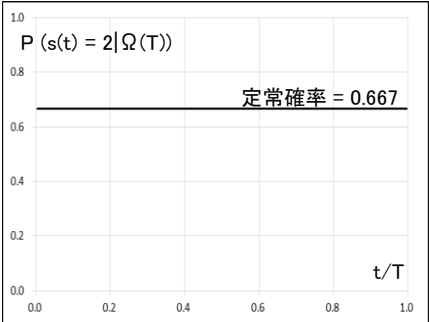
(付図38-5) 平滑化確率 $P(s(t) = 1 | \Omega(T))$



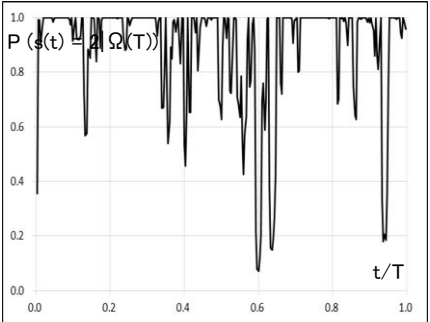
(付図38-8) 平滑化確率 $P(s(t) = 1 | \Omega(T))$



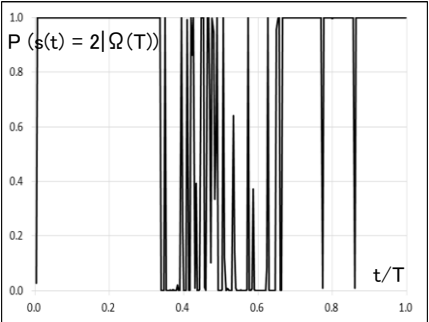
(付図37-2) 平滑化確率 $P(s(t) = 2 | \Omega(T))$



(付図38-3) 平滑化確率 $P(s(t) = 2 | \Omega(T))$



(付図38-6) 平滑化確率 $P(s(t) = 2 | \Omega(T))$



(付図38-9) 平滑化確率 $P(s(t) = 2 | \Omega(T))$

付録A. 等比数列の和と無限等比数列の和

(1) 等比数列の和

初項が a 、公比が r 、項数が n である等比数列は、 $a, ar, ar^2, \dots, ar^{n-1}$ で表される。

この等比数列の和は、 $a + ar + ar^2 + \dots + ar^{n-1} = a(r^n - 1) / (r - 1)$ となる。

(確認)

$$S_n = a + ar + ar^2 + \dots + ar^{n-1} \quad \dots \quad \textcircled{1}$$

両辺を r 倍する。

$$r S_n = ar + ar^2 + ar^3 + \dots + ar^n \quad \dots \quad \textcircled{2}$$

$\textcircled{1} - \textcircled{2}$ より、

$$(1 - r) S_n = a - ar^n$$

$$\text{従って、} S_n = (a - ar^n) / (1 - r) = a(1 - r^n) / (1 - r) = a(r^n - 1) / (r - 1)$$

(2) 無限等比数列の和

初項が a 、公比が r である無限等比数列は、 $a, ar, ar^2, \dots$ で表される。

この等比数列の和が収束するための必要十分条件は、 $a = 0$ もしくは、 $|r| < 1$ であり、

・ $a = 0$ のとき、 $a + ar + ar^2 + \dots = 0$

・ $|r| < 1$ のとき、 $a + ar + ar^2 + \dots = a / (1 - r)$

(確認)

・ $a = 0$ のときは、 $a + ar + ar^2 + \dots = 0$

・ $r = 1$ のときは、 $a + ar + ar^2 + \dots = a + a + a + \dots$ は、発散する。

・ $r \neq 1$ のときは、まず、無限等比級数の部分 and S_n を求めると、(1)より、

$$S_n = a(r^n - 1) / (r - 1)$$

無限等比級数の和は、 $\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} a(r^n - 1) / (r - 1)$ である。

・ $|r| < 1$ のとき、 $\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} a(r^n - 1) / (r - 1) = a(0 - 1) / (r - 1) = -a / (r - 1) = a / (1 - r)$

・ $|r| > 1$ のとき、 $\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} a(r^n - 1) / (r - 1) \rightarrow \text{発散}$

付録B. 期待値、分散、共分散の公式 と 分散共分散行列の公式

確率変数 ... X, Y : スカラー表記、 $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2$: ベクトル表記、 $\mathbf{X}$: 行列表記、 定数 ... a, b, c, d : スカラー表記、 $\mathbf{a}, \mathbf{b}, \mathbf{d}$: ベクトル表記、 $\mathbf{A}, \mathbf{C}$: 行列表記

- ① 共分散 $\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])]$, $\text{Cov}[\mathbf{x}, \mathbf{y}] = E[(\mathbf{x} - E[\mathbf{x}]) (\mathbf{y} - E[\mathbf{y}])^T]$
- ② 分散 $V[X] = \text{Cov}[X, X] = E[(X - E[X])^2]$, $V[\mathbf{x}] = \text{Cov}[\mathbf{x}, \mathbf{x}] = E[(\mathbf{x} - E[\mathbf{x}]) (\mathbf{x} - E[\mathbf{x}])^T]$

■期待値

- (1) $E[aX + cY] = aE[X] + cE[Y]$
- (2) $E[\mathbf{a}\mathbf{x} + \mathbf{c}\mathbf{y}] = \mathbf{a}E[\mathbf{x}] + \mathbf{c}E[\mathbf{y}]$
- (3) X と Y が独立であるとき、 $E[XY] = E[X]E[Y]$

■共分散

- (1) $\text{Cov}[X, Y] = E[X, Y] - E[X]E[Y]$
(確認) $\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY - XE[Y] - E[X]Y + E[X]E[Y]] = E[XY] - E[XE[Y]] - E[E[X]Y] + E[E[X]E[Y]]$
 $= E[XY] - E[X]E[Y] - E[X]E[Y] + E[X]E[Y] = E[XY] - E[X]E[Y]$
- (2) X と Y が独立であるとき、 $\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$
- (3) $\text{Cov}[aX + b, cY + d] = ac \text{Cov}[X, Y]$
(確認) $\text{Cov}[aX + b, cY + d] = E[(aX + b - E[aX + b])(cY + d - E[cY + d])] = E[(aX + b - aE[X] - b)(cY + d - cE[Y] - d)]$
 $= E[(aX - aE[X])(cY - cE[Y])] = E[a(X - E[X])c(Y - E[Y])] = acE[(X - E[X])(Y - E[Y])] = ac \text{Cov}[X, Y]$
- (4) $\text{Cov}[\mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{C}\mathbf{y} + \mathbf{d}] = \mathbf{A} \text{Cov}[\mathbf{x}, \mathbf{y}] \mathbf{C}^T$
(確認) $\text{Cov}[\mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{C}\mathbf{y} + \mathbf{d}] = E[(\mathbf{A}\mathbf{x} + \mathbf{b} - E[\mathbf{A}\mathbf{x} + \mathbf{b}])(\mathbf{C}\mathbf{y} + \mathbf{d} - E[\mathbf{C}\mathbf{y} + \mathbf{d}])^T] = E[(\mathbf{A}\mathbf{x} + \mathbf{b} - \mathbf{A}E[\mathbf{x}] - \mathbf{b})(\mathbf{C}\mathbf{y} + \mathbf{d} - \mathbf{C}E[\mathbf{y}] - \mathbf{d})^T]$
 $= E[(\mathbf{A}\mathbf{x} - \mathbf{A}E[\mathbf{x}])(\mathbf{C}\mathbf{y} - \mathbf{C}E[\mathbf{y}])^T] = E[\mathbf{A}(\mathbf{x} - E[\mathbf{x}])(\mathbf{C}(\mathbf{y} - E[\mathbf{y}]))^T] = E[\mathbf{A}(\mathbf{x} - E[\mathbf{x}]) (\mathbf{y} - E[\mathbf{y}])^T \mathbf{C}^T] = \mathbf{A}E[(\mathbf{x} - E[\mathbf{x}]) (\mathbf{y} - E[\mathbf{y}])^T] \mathbf{C}^T$
 $= \mathbf{A} \text{Cov}[\mathbf{x}, \mathbf{y}] \mathbf{C}^T$
- (5) $\text{Cov}[\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}_1 + \mathbf{y}_2] = \text{Cov}[\mathbf{x}_1, \mathbf{y}_1] + \text{Cov}[\mathbf{x}_1, \mathbf{y}_2] + \text{Cov}[\mathbf{x}_2, \mathbf{y}_1] + \text{Cov}[\mathbf{x}_2, \mathbf{y}_2]$
(確認) $\text{Cov}[\mathbf{x}_1 + \mathbf{x}_2, \mathbf{y}_1 + \mathbf{y}_2] = E[(\mathbf{x}_1 + \mathbf{x}_2 - E[\mathbf{x}_1 + \mathbf{x}_2])(\mathbf{y}_1 + \mathbf{y}_2 - E[\mathbf{y}_1 + \mathbf{y}_2])^T] = E[(\mathbf{x}_1 + \mathbf{x}_2 - E[\mathbf{x}_1] - E[\mathbf{x}_2])(\mathbf{y}_1 + \mathbf{y}_2 - E[\mathbf{y}_1] - E[\mathbf{y}_2])^T]$
 $= E[(\mathbf{x}_1 - E[\mathbf{x}_1] + \mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_1 - E[\mathbf{y}_1] + \mathbf{y}_2 - E[\mathbf{y}_2])^T] = E[(\mathbf{x}_1 - E[\mathbf{x}_1] + \mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_1 - E[\mathbf{y}_1])^T + (\mathbf{y}_2 - E[\mathbf{y}_2])^T]$
 $= E[(\mathbf{x}_1 - E[\mathbf{x}_1])(\mathbf{y}_1 - E[\mathbf{y}_1])^T + (\mathbf{x}_1 - E[\mathbf{x}_1])(\mathbf{y}_2 - E[\mathbf{y}_2])^T + (\mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_1 - E[\mathbf{y}_1])^T + (\mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_2 - E[\mathbf{y}_2])^T]$
 $= E[(\mathbf{x}_1 - E[\mathbf{x}_1])(\mathbf{y}_1 - E[\mathbf{y}_1])^T] + E[(\mathbf{x}_1 - E[\mathbf{x}_1])(\mathbf{y}_2 - E[\mathbf{y}_2])^T] + E[(\mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_1 - E[\mathbf{y}_1])^T] + E[(\mathbf{x}_2 - E[\mathbf{x}_2])(\mathbf{y}_2 - E[\mathbf{y}_2])^T]$
 $= \text{Cov}[\mathbf{x}_1, \mathbf{y}_1] + \text{Cov}[\mathbf{x}_1, \mathbf{y}_2] + \text{Cov}[\mathbf{x}_2, \mathbf{y}_1] + \text{Cov}[\mathbf{x}_2, \mathbf{y}_2]$

■分散

- (1) $V[aX] = a^2 V[X]$
- (2) $V[\mathbf{a}\mathbf{x}] = \mathbf{a} V[\mathbf{x}] \mathbf{a}^T$
(確認) $V[\mathbf{a}\mathbf{x}] = E[(\mathbf{a}\mathbf{x} - E[\mathbf{a}\mathbf{x}])(\mathbf{a}\mathbf{x} - E[\mathbf{a}\mathbf{x}])^T] = E[(\mathbf{a}\mathbf{x} - \mathbf{a}E[\mathbf{x}])(\mathbf{a}\mathbf{x} - \mathbf{a}E[\mathbf{x}])^T] = E[\mathbf{a}(\mathbf{x} - E[\mathbf{x}])(\mathbf{a}(\mathbf{x} - E[\mathbf{x}]))^T] = E[\mathbf{a}(\mathbf{x} - E[\mathbf{x}]) (\mathbf{x} - E[\mathbf{x}])^T \mathbf{a}^T]$
 $= \mathbf{a}E[(\mathbf{x} - E[\mathbf{x}]) (\mathbf{x} - E[\mathbf{x}])^T] \mathbf{a}^T = \mathbf{a} V[\mathbf{x}] \mathbf{a}^T$
- (3) $V[X + Y] = V[X] + V[Y] + 2 \text{Cov}(X, Y)$
(確認) $V[X + Y] = E[(X + Y - E[X + Y])^2] = E[(X + Y - E[X] - E[Y])^2] = E[((X - E[X]) + (Y - E[Y]))^2]$
 $= E[(X - E[X])^2 + (Y - E[Y])^2 + 2(X - E[X])(Y - E[Y])] = E[(X - E[X])^2] + E[(Y - E[Y])^2] + 2E[(X - E[X])(Y - E[Y])] = V[X] + V[Y] + 2 \text{Cov}[X, Y]$
- (4) X と Y が独立であるとき、 $\text{Cov}(X, Y) = 0$ であるので、 $V[X + Y] = V[X] + V[Y]$
- (5) $V[X] = E[X^2] - (E[X])^2$
(確認) $V[X] = E[(X - E[X])^2] = E[X^2 - 2XE[X] + (E[X])^2] = E[X^2] - 2E[XE[X]] + E[(E[X])^2] = E[X^2] - 2(E[X])^2 + (E[X])^2 = E[X^2] - (E[X])^2$
- (6) X と Y が独立であるとき、 $V[XY] = V[X]V[Y] + V[X](E[Y])^2 + (E[X])^2V[Y]$
(確認) $V[XY] = E[(XY)^2] - (E[XY])^2 = E[X^2Y^2] - (E[X]E[Y])^2 = E[X^2]E[Y^2] - (E[X])^2(E[Y])^2 = (V[X] + (E[X])^2)(V[Y] + (E[Y])^2) - (E[X])^2(E[Y])^2$
 $= V[X]V[Y] + V[X](E[Y])^2 + (E[X])^2V[Y] + (E[X])^2(E[Y])^2 - (E[X])^2(E[Y])^2 = V[X]V[Y] + V[X](E[Y])^2 + (E[X])^2V[Y]$

付録C. ベクトルと行列の微分の公式

ベクトル $\mathbf{a}$: サイズ $n \times 1$, $\mathbf{b}$: $n \times 1$, $\mathbf{x}$: $n \times 1$, $\mathbf{y}$: $n \times 1$, $\mathbf{z}$: $n \times 1$

行列 $\mathbf{I}$ (単位行列): $n \times n$, $\mathbf{A}$: $n \times n$, $\mathbf{B}$: $n \times n$, $\mathbf{X}$: $n \times n$

行列式 $|\mathbf{A}|$

(1) スカラーのベクトル微分

1) $\partial (\mathbf{a}^T \mathbf{x}) / \partial \mathbf{x} = \mathbf{a}$

2) $\partial (\mathbf{x}^T \mathbf{a}) / \partial \mathbf{x} = \mathbf{a}$

3) $\partial (\mathbf{x}^T \mathbf{x}) / \partial \mathbf{x} = 2 \mathbf{x}$

4) $\partial (\mathbf{x} - \mathbf{a})^T (\mathbf{x} - \mathbf{a}) / \partial \mathbf{x} = 2 (\mathbf{x} - \mathbf{a})$

5) $\partial (\mathbf{A} \mathbf{x} - \mathbf{b})^T (\mathbf{A} \mathbf{x} - \mathbf{b}) / \partial \mathbf{x} = 2 \mathbf{A}^T (\mathbf{A} \mathbf{x} - \mathbf{b})$

6) $\partial (\mathbf{x}^T \mathbf{A} \mathbf{x}) / \partial \mathbf{x} = (\mathbf{A} + \mathbf{A}^T) \mathbf{x}$

7) $\partial (\mathbf{A} \mathbf{x} - \mathbf{b})^T \mathbf{B} (\mathbf{A} \mathbf{x} - \mathbf{b}) / \partial \mathbf{x} = \mathbf{A}^T (\mathbf{B} + \mathbf{B}^T) (\mathbf{A} \mathbf{x} - \mathbf{b})$

(2) ベクトルのベクトル微分

1) $\partial \mathbf{x} / \partial \mathbf{x} = \mathbf{I}$

2) $\partial (\mathbf{A} \mathbf{x}) / \partial \mathbf{x} = \mathbf{A}^T$

3) $\partial (\mathbf{A}^T \mathbf{x}) / \partial \mathbf{x} = \mathbf{A}$

4) $\partial (\mathbf{x}^T \mathbf{A}) / \partial \mathbf{x} = \mathbf{A}$

5) $\partial \mathbf{y} / \partial \mathbf{x} = (\partial \mathbf{z} / \partial \mathbf{x}) (\partial \mathbf{y} / \partial \mathbf{z})$

(3) スカラーの行列微分

1) $\partial |\mathbf{A}| / \partial \mathbf{A} = |\mathbf{A}| (\mathbf{A}^{-1})^T$

2) $\partial \ln |\mathbf{A}| / \partial \mathbf{A} = (\mathbf{A}^{-1})^T$

(4) 行列のベクトル微分

1) $\partial \mathbf{A}^{-1} / \partial \mathbf{x} = -\mathbf{A}^{-1} (\partial \mathbf{A} / \partial \mathbf{x}) \mathbf{A}^{-1}$

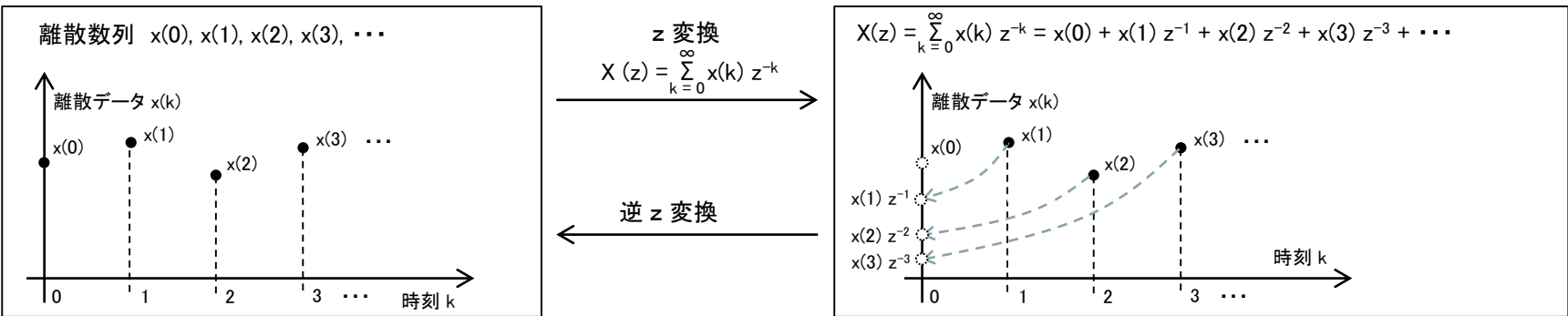
付録D-1. z 変換の定義

(1) 定義

因果的な離散数列 (負の時点の数列は 0) $x(k)$ の z 変換

$$X(z) = Z[x(k)] = \sum_{k=0}^{\infty} x(k) z^{-k} \quad \text{但し、} z = e^{a+j\omega k} = e^a e^{j\omega k} = r e^{j\omega k} \quad (r = e^a \text{ とおく})$$

z 変換のイメージを (付図D-1) に示す。 z 変換は離散データを時点 0 にずらして足し合わせたものになる (z^{-1} は、後述するように時間シフトの演算子になる)。

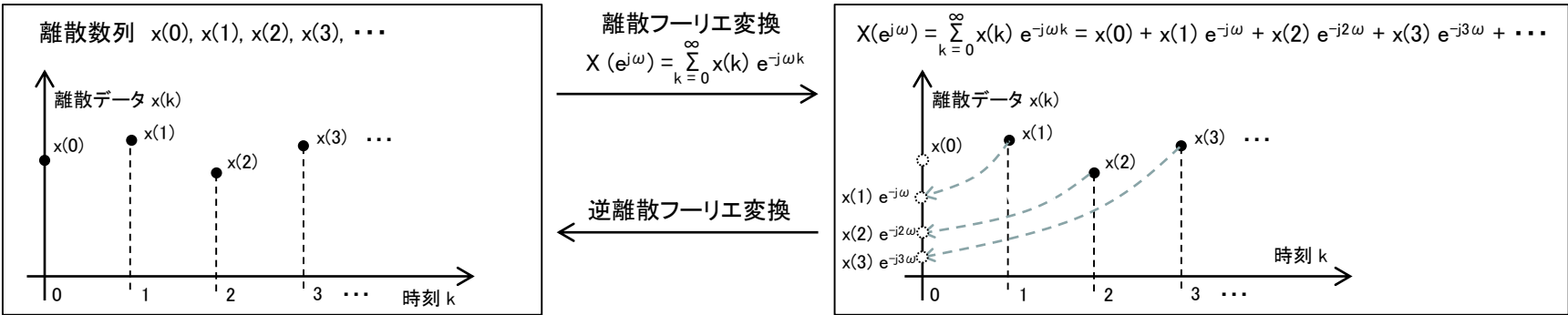


(付図D-1) z 変換

(付図D-1) の $X(z)$ は、時間領域でのラプラス変換の畳み込み積分を想起させる (別途資料「制御工学 (古典制御) の基礎知識」を参照)。
 z 変換が離散時間のラプラス変換であると言われる所以である。

(2) 離散フーリエ変換と z 変換の関係

z 変換 $X(z)$ に、 $z = e^{j\omega}$ を代入することで、離散フーリエ変換 $X(e^{j\omega})$ になる (別途資料「周波数解析 (フーリエ解析) の基礎知識」を参照)。



(付図D-2) 離散フーリエ変換

付録D-2. z 変換の性質

(1) 線形性

$$Z[a\,x_1(k)+b\,x_2(k)]=Z[a\,x_1(k)+b\,x_2(k)]=\sum_{k=0}^{\infty}(a\,x_1(k)+b\,x_2(k))\,z^{-k}=a\sum_{k=0}^{\infty}x_1(k)\,z^{-k}+b\sum_{k=0}^{\infty}x_2(k)\,z^{-k}=a\,Z[x_1(k)]+b\,Z[x_2(k)]$$

(2) 微分

$$Z[x(k)]=\sum_{k=0}^{\infty}x(k)\,z^{-k}=x(0)+x(1)\,z^{-1}+x(2)\,z^{-2}+\cdots+x(n)\,z^{-n}+\cdots$$

両辺を微分すると、

$$d/dz\,Z[x(k)]=-x(1)\,z^{-2}-2\,x(2)\,z^{-3}-\cdots-n\,x(n)\,z^{-(n+1)}-\cdots=-z^{-1}(x(1)\,z^{-1}-2\,x(2)\,z^{-2}-\cdots-n\,x(n)\,z^{-n}-\cdots)=-z^{-1}\sum_{k=0}^{\infty}k\,x(k)\,z^{-k}=-z^{-1}\,Z[k\,x(k)]$$

従って、 $Z[k\,x(k)]=-z\,d/dz\,Z[x(k)]$

(3) 指数倍

$$Z[a^k\,x(k)]=X(a^{-1}\,z)=\sum_{k=0}^{\infty}a^k\,x(k)\,z^{-k}=a^0\,x(0)+a^1\,x(1)\,z^{-1}+a^2\,x(2)\,z^{-2}+\cdots+a_n\,x(n)\,z^{-n}+\cdots=\sum_{k=0}^{\infty}x(k)\,(a^{-1}\,z)^{-k}=X(a^{-1}\,z)$$

(4) 時間軸の推移(時間遅れ) (付図D-3)

因果的な離散数列(負の時点の数列は 0)であるので、 $m>0$ として、

$$\begin{aligned} Z[x(k-m)]&=\sum_{k=0}^{\infty}x(k-m)\,z^{-k}=z^{-m}\sum_{k=0}^{\infty}x(k-m)\,z^{-(k-m)}=z^{-m}\left(\sum_{k=0}^{m-1}x(k-m)\,z^{-(k-m)}+\sum_{k=m}^{\infty}x(k-m)\,z^{-(k-m)}\right)=z^{-m}\left(0+\sum_{k=m}^{\infty}x(k-m)\,z^{-(k-m)}\right)\\ &=z^{-m}\sum_{k=m}^{\infty}x(k-m)\,z^{-(k-m)}=z^{-m}\sum_{k=0}^{\infty}x(k)\,z^{-k}=z^{-m}\,Z[x(k)] \end{aligned}$$

(5) 時間軸の推移(時間進み) (付図D-3)

$m=1$ のとき、

$$Z[x(k+1)]=\sum_{k=0}^{\infty}x(k+1)\,z^{-k}=z\sum_{k=0}^{\infty}x(k+1)\,z^{-(k+1)}=z(-x(0)+x(0)+\sum_{k=0}^{\infty}x(k+1)\,z^{-(k+1)})=z(-x(0)+\sum_{k=0}^{\infty}x(k)\,z^{-k})=z(-x(0)+X(z))=z\,X(z)-z\,x(0)$$

$m=2$ のとき、

$$\begin{aligned} Z[x(k+2)]&=\sum_{k=0}^{\infty}x(k+2)\,z^{-k}=z^2\sum_{k=0}^{\infty}x(k+2)\,z^{-(k+2)}=z^2(-x(0)-x(1)\,z^{-1}+x(0)+x(1)\,z^{-1}+\sum_{k=0}^{\infty}x(k+2)\,z^{-(k+2)})\\ &=z^2(-x(0)-x(1)\,z^{-1}+\sum_{k=0}^{\infty}x(k)\,z^{-k})=z^2(-x(0)-x(1)\,z^{-1}+X(z))=z^2\,X(z)-z^2\,x(0)-z\,x(1) \end{aligned}$$

同様にして、 $m>0$ として、

$$Z[x(k+m)]=z^m\,X(z)-z^m\,x(0)-z^{m-1}\,x(1)-\cdots-z\,x(m-1)=z^m\,Z[x(k)]-z^m\,x(0)-z^{m-1}\,x(1)-\cdots-z\,x(m-1)$$

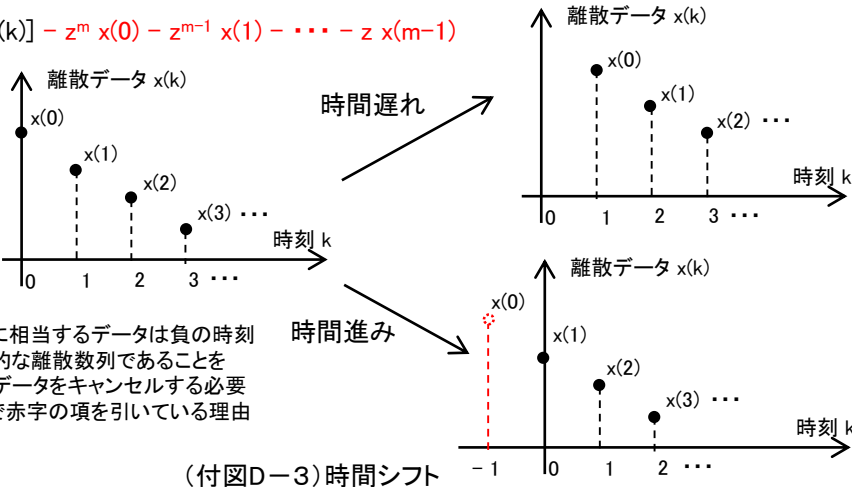
(6) 初期値定理

$$\lim_{z\rightarrow\infty}X(z)=\lim_{z\rightarrow\infty}\sum_{k=0}^{\infty}x(k)\,z^{-k}=\lim_{z\rightarrow\infty}(x(0)+z^{-1}\,x(1)+z^{-2}\,x(2)+\cdots)=x(0)$$

(7) 最終値定理 (結果のみの掲載)

$$\lim_{k\rightarrow\infty}x(k)=\lim_{z\rightarrow 1}(z-1)\,X(z)=\lim_{z\rightarrow 1}(z-1)\,Z[x(k)]$$

(注)
時間進みの図の時刻 $-k$ に相当するデータは負の時刻のデータであるので、因果的な離散数列であることを考慮して、この時刻 $-k$ のデータをキャンセルする必要がある。それが、(5) の式で赤字の項を引いている理由になる。



付録E. 偏相関係数(Partial Correlation)

一般に、**偏相関(PARCOR)**とは、他の因子の影響を取り除いた上で、関心のある2変数間の相関を表す概念である。

・偏相関係数の導出

今、確率変数 X, Y, Z があり、確率変数 X の影響を除いた確率変数 Y と Z の相関係数(偏相関係数)を求める。

別途資料「統計的検定の手順」の(付録7)単回帰式の導出を参考にして、

回帰式 $\hat{y}_i = \alpha + \beta x_i$ の係数 α, β は、 $\mu_X = E[X] = E[x_i], \mu_Y = E[Y] = E[y_i]$ として、

$$\alpha = \mu_Y - \beta \mu_X \quad \cdots \text{ (付式E-1)}$$

$$\beta = \left(\sum_{i=1}^n (x_i - \mu_X)(y_i - \mu_Y) \right) / \sum_{i=1}^n (x_i - \mu_X)^2 = \text{Cov}(X, Y) / \sigma_X^2 \quad \cdots \text{ (付式E-2)}$$

ここで、変数 Y から変数 X の影響を除くには x_i ごとの y_i と $\hat{y}_i$ の差に注目して、 $\hat{y}_i = \alpha + \beta x_i = (\mu_Y - \beta \mu_X) + \beta x_i$ であるので、
 $y'_i = y_i - \hat{y}_i = y_i - ((\mu_Y - \beta \mu_X) + \beta x_i) = (y_i - \mu_Y) - (x_i - \mu_X) \beta = (y_i - \mu_Y) - (x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2 \quad \cdots \text{ (付式E-3)}$

同様に、変数 Z から変数 X の影響を除くには z_i ごとの y_i と $\hat{y}_i$ の差に注目して、

$$z'_i = z_i - \hat{z}_i = z_i - ((\mu_Z - \beta \mu_X) + \beta x_i) = (z_i - \mu_Z) - (x_i - \mu_X) \beta = (z_i - \mu_Z) - (x_i - \mu_X) \text{Cov}(X, Z) / \sigma_X^2 \quad \cdots \text{ (付式E-4)}$$

従って、変数 X の影響を除いた変数 Y と Z の偏相関係数を $\rho_{YZ.X}$ で表せば、それは変数 Y' と Z' の相関係数 $\rho_{Y'Z'}$ になる。つまり、

$$\begin{aligned} \rho_{YZ.X} = \rho_{Y'Z'} &= E[(Y' - \mu_{Y'}) (Z' - \mu_{Z'})] / (E[(Y' - \mu_{Y'})^2] E[(Z' - \mu_{Z'})^2])^{1/2} \\ \text{分子} &= E[Y'Z'] - Y' \mu_{Z'} - \mu_{Y'} Z' + \mu_{Y'} \mu_{Z'} = E[Y'Z'] - \mu_{Z'} \mu_{Y'} - \mu_{Y'} \mu_{Z'} + \mu_{Y'} \mu_{Z'} = E[Y'Z'] - \mu_{Y'} \mu_{Z'} \\ \text{分母の第1項} &= (E[Y'^2] - 2Y' \mu_{Y'} + \mu_{Y'}^2)^{1/2} = (E[Y'^2] - 2\mu_{Y'} E[Y'] + \mu_{Y'}^2)^{1/2} = (E[Y'^2] - 2\mu_{Y'}^2 + \mu_{Y'}^2)^{1/2} = (E[Y'^2] - \mu_{Y'}^2)^{1/2} \end{aligned}$$

従って、 $\rho_{YZ.X} = (E[Y'Z'] - \mu_{Y'} \mu_{Z'}) / ((E[Y'^2] - \mu_{Y'}^2)^{1/2} (E[Z'^2] - \mu_{Z'}^2)^{1/2}) \quad \cdots \text{ (付式E-5)}$

ここで、(付式E-3)より、 $\mu_{Y'} = E[Y'] = E[(y_i - \mu_Y) - (x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2] = E[y_i - \mu_Y] - (\text{Cov}(X, Y) / \sigma_X^2) E[x_i - \mu_X] = 0 - (\text{Cov}(X, Y) / \sigma_X^2) \cdot 0 = 0 \quad \cdots \text{ (付式E-6)}$

同様に、 $\mu_{Z'} = 0 \quad \cdots \text{ (付式E-7)}$

また、(付式E-3)と(付式E-4)より、

$$\begin{aligned} E[Y'Z'] &= E[((y_i - \mu_Y) - (x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2) ((z_i - \mu_Z) - (x_i - \mu_X) \text{Cov}(X, Z) / \sigma_X^2)] \\ &= E[(y_i - \mu_Y)(z_i - \mu_Z)] - E[(y_i - \mu_Y)(x_i - \mu_X) \text{Cov}(X, Z) / \sigma_X^2] - E[(x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2 (z_i - \mu_Z)] \\ &\quad + E[(x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2 (x_i - \mu_X) \text{Cov}(X, Z) / \sigma_X^2] \\ &= E[(y_i - \mu_Y)(z_i - \mu_Z)] - (\text{Cov}(X, Z) / \sigma_X^2) E[(y_i - \mu_Y)(x_i - \mu_X)] - (\text{Cov}(X, Y) / \sigma_X^2) E[(x_i - \mu_X)(z_i - \mu_Z)] \\ &\quad + (\text{Cov}(X, Y) / \sigma_X^2) (\text{Cov}(X, Z) / \sigma_X^2) E[(x_i - \mu_X)(x_i - \mu_X)] \\ &= \text{Cov}(Y, Z) - (\text{Cov}(X, Z) / \sigma_X^2) \text{Cov}(Y, X) - (\text{Cov}(X, Y) / \sigma_X^2) \text{Cov}(X, Z) + (\text{Cov}(X, Y) / \sigma_X^2) (\text{Cov}(X, Z) / \sigma_X^2) \text{Cov}(X, X) \\ &= \text{Cov}(Y, Z) - 2 \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2 + (\text{Cov}(X, Y) / \sigma_X^2) (\text{Cov}(X, Z) / \sigma_X^2) \sigma_X^2 \\ &= \text{Cov}(Y, Z) - 2 \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2 + (\text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2) = \text{Cov}(Y, Z) - \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2 \quad \cdots \text{ (付式E-8)} \end{aligned}$$

$$\begin{aligned} E[Y'^2] &= E[(y_i - \mu_Y) - (x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2]^2 = E[(y_i - \mu_Y)^2] - E[2(y_i - \mu_Y)(x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2] + E[(x_i - \mu_X) \text{Cov}(X, Y) / \sigma_X^2]^2 \\ &= E[(y_i - \mu_Y)^2] - 2(\text{Cov}(X, Y) / \sigma_X^2) E[(y_i - \mu_Y)(x_i - \mu_X)] + (\text{Cov}(X, Y) / \sigma_X^2)^2 E[(x_i - \mu_X)^2] \\ &= \sigma_Y^2 - 2(\text{Cov}(X, Y) / \sigma_X^2) \text{Cov}(X, Y) + (\text{Cov}(X, Y) / \sigma_X^2)^2 \sigma_X^2 = \sigma_Y^2 - (\text{Cov}(X, Y))^2 / \sigma_X^2 \end{aligned}$$

ここで、 $\rho_{XY} = \text{Cov}(X, Y) / (\sigma_X^2 \sigma_Y^2)^{1/2}$ であるので、 $E[Y'^2] = \sigma_Y^2 - (\rho_{XY})^2 (\sigma_X^2 \sigma_Y^2) / \sigma_X^2 = \sigma_Y^2 (1 - (\rho_{XY})^2) \quad \cdots \text{ (付式E-9)}$

同様に、 $E[Z'^2] = \sigma_Z^2 (1 - (\rho_{XZ})^2) \quad \cdots \text{ (付式E-10)}$

以上、(付式E-5)に、(付式E-6)、(付式E-7)、(付式E-8)、(付式E-9)、(付式E-10)を代入すると、

$$\begin{aligned} \rho_{YZ.X} &= (\text{Cov}(Y, Z) - \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2 - 0 \cdot 0) / ((\sigma_Y^2 (1 - (\rho_{XY})^2) - 0)^{1/2} (\sigma_Z^2 (1 - (\rho_{XZ})^2) - 0)^{1/2}) \\ &= (\text{Cov}(Y, Z) - \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2) / (\sigma_Y (1 - (\rho_{XY})^2)^{1/2} (\sigma_Z (1 - (\rho_{XZ})^2)^{1/2})) \\ &= (\text{Cov}(Y, Z) - \text{Cov}(X, Y) \text{Cov}(X, Z) / \sigma_X^2) / (\sigma_Y \sigma_Z (1 - (\rho_{XY})^2)^{1/2} (1 - (\rho_{XZ})^2)^{1/2}) \\ &= (\text{Cov}(Y, Z) / (\sigma_Y \sigma_Z) - (\text{Cov}(X, Y) / (\sigma_X \sigma_Y)) (\text{Cov}(X, Z) / (\sigma_X \sigma_Z))) / ((1 - (\rho_{XY})^2)^{1/2} (1 - (\rho_{XZ})^2)^{1/2}) \\ &= (\rho_{YZ} - \rho_{XY} \rho_{XZ}) / ((1 - (\rho_{XY})^2)^{1/2} (1 - (\rho_{XZ})^2)^{1/2}) \quad \cdots \text{ (付式E-11)} \end{aligned}$$

(付式E-11)が、3変数 X, Y, Z で、 X の影響を除外したときの Y と Z の相関係数、つまり偏相関係数を求める式である。

付録F. 自己共分散行列から自己相関行列を求める

時間差 k の自己共分散行列 $\Gamma(k)$ を、(付式F-1)に示す。

$$\Gamma(k)=\begin{bmatrix}\sigma_{11}(k)&\sigma_{12}(k)&\cdots&\sigma_{1n}(k)\\ \sigma_{21}(k)&\sigma_{22}(k)&\cdots&\sigma_{2n}(k)\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}(k)&\sigma_{n2}(k)&\cdots&\sigma_{nn}(k)\end{bmatrix}=\begin{bmatrix}\sigma_1^2(k)&\sigma_{12}^{(k)}&\cdots&\sigma_{1n}(k)\\ \sigma_{21}^{(k)}&\sigma_2^2(k)&\cdots&\sigma_{2n}(k)\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}^{(k)}&\sigma_{n2}^{(k)}&\cdots&\sigma_n^2(k)\end{bmatrix}\qquad \cdots \text{ (付式F-1)}$$

時間差 k の自己相関行列 $\rho(k)$ を、(付式F-2)に示す。

$$\rho(k)=\begin{bmatrix}\rho_{11}(k)&\rho_{12}(k)&\cdots&\rho_{1n}(k)\\ \rho_{21}(k)&\rho_{22}(k)&\cdots&\rho_{2n}(k)\\ \cdots&\cdots&\cdots&\cdots\\ \rho_{n1}(k)&\rho_{n2}(k)&\cdots&\rho_{nn}(k)\end{bmatrix}\qquad \cdots \text{ (付式F-2)}$$

対角要素が分散 $V[y_1(t)]=\sigma_1^2(k)$, $V[y_2(t)]=\sigma_2^2(k)$, $\cdots$, $V[y_n(t)]=\sigma_n^2(k)$ からなる、対角行列 $\mathbf{D}$ を用意する。なお、 $\mathbf{D}^{1/2}\mathbf{D}^{-1/2}=\mathbf{I}$ (単位行列)になる。

$$\mathbf{D}=\begin{bmatrix}\sigma_1^2(k)&0&\cdots&0\\ 0&\sigma_2^2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&\sigma_n^2(k)\end{bmatrix},\quad \mathbf{D}^{1/2}=\begin{bmatrix}\sigma_1(k)&0&\cdots&0\\ 0&\sigma_2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&\sigma_n(k)\end{bmatrix},\quad \mathbf{D}^{-1/2}=\begin{bmatrix}1/\sigma_1(k)&0&\cdots&0\\ 0&1/\sigma_2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&1/\sigma_n(k)\end{bmatrix}$$

以下に示すように、行列 $\mathbf{D}^{-1/2}$ を使用して、自己共分散行列 $\Gamma(k)$ を自己相関行列 $\rho(k)$ に変換することができる。

$$\begin{aligned}\mathbf{D}^{-1/2}\Gamma(k)\mathbf{D}^{-1/2}&=\begin{bmatrix}1/\sigma_1(k)&0&\cdots&0\\ 0&1/\sigma_2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&1/\sigma_n(k)\end{bmatrix}\begin{bmatrix}\sigma_1^2(k)&\sigma_{12}(k)&\cdots&\sigma_{1n}(k)\\ \sigma_{21}(k)&\sigma_2^2(k)&\cdots&\sigma_{2n}(k)\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}(k)&\sigma_{n2}(k)&\cdots&\sigma_n^2(k)\end{bmatrix}\begin{bmatrix}1/\sigma_1(k)&0&\cdots&0\\ 0&1/\sigma_2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&1/\sigma_n(k)\end{bmatrix}\\&=\begin{bmatrix}\sigma_1^2(k)/\sigma_1(k)&\sigma_{12}(k)/\sigma_1(k)&\cdots&\sigma_{1n}(k)/\sigma_1(k)\\ \sigma_{21}(k)/\sigma_2(k)&\sigma_2^2(k)/\sigma_2(k)&\cdots&\sigma_{2n}(k)/\sigma_2(k)\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}(k)/\sigma_n(k)&\sigma_{n2}(k)/\sigma_n(k)&\cdots&\sigma_n^2(k)/\sigma_n(k)\end{bmatrix}\begin{bmatrix}1/\sigma_1(k)&0&\cdots&0\\ 0&1/\sigma_2(k)&\cdots&0\\ \cdots&\cdots&\cdots&\cdots\\ 0&0&\cdots&1/\sigma_n(k)\end{bmatrix}\\&=\begin{bmatrix}\sigma_1^2(k)/(\sigma_1(k)\sigma_1(k))&\sigma_{12}(k)/(\sigma_1(k)\sigma_2(k))&\cdots&\sigma_{1n}(k)/(\sigma_1(k)\sigma_n(k))\\ \sigma_{21}(k)/(\sigma_2(k)\sigma_1(k))&\sigma_2^2(k)/(\sigma_2(k)\sigma_2(k))&\cdots&\sigma_{2n}(k)/(\sigma_2(k)\sigma_n(k))\\ \cdots&\cdots&\cdots&\cdots\\ \sigma_{n1}(k)/(\sigma_n(k)\sigma_1(k))&\sigma_{n2}(k)/(\sigma_n(k)\sigma_2(k))&\cdots&\sigma_n^2(k)/(\sigma_n(k)\sigma_n(k))\end{bmatrix}=\begin{bmatrix}\rho_{11}(k)&\rho_{12}(k)&\cdots&\rho_{1n}(k)\\ \rho_{21}(k)&\rho_{22}(k)&\cdots&\rho_{2n}(k)\\ \cdots&\cdots&\cdots&\cdots\\ \rho_{n1}(k)&\rho_{n2}(k)&\cdots&\rho_{nn}(k)\end{bmatrix}=\rho(k)\qquad \cdots \text{ (付式F-3)}\end{aligned}$$

なお、以下に示すように、(付式F-3)の両辺に左右から行列 $\mathbf{D}^{1/2}$ をかけて、自己相関行列 $\rho(k)$ を自己共分散行列 $\Gamma(k)$ に変換することができる。

$$\mathbf{D}^{1/2}(\mathbf{D}^{-1/2}\Gamma(k)\mathbf{D}^{-1/2})\mathbf{D}^{1/2}=\mathbf{D}^{1/2}\rho(k)\mathbf{D}^{1/2}\quad \rightarrow \quad \mathbf{I}\Gamma(k)\mathbf{I}=\mathbf{D}^{1/2}\rho(k)\mathbf{D}^{1/2}$$

よって、

$$\Gamma(k)=\mathbf{D}^{1/2}\rho(k)\mathbf{D}^{1/2}\qquad \cdots \text{ (付式F-4)}$$

付録G. VAR の尤度関数

(1) 1 変数の対数尤度

確率変数 $y(t)$ が正規分布 $N(\mu, \sigma^2)$ に従うとき、 $y(t)$ の確率密度関数は(付式G-1)で表すことができる(別途資料「いろいろな確率分布」を参照)。

$$f(y(t)) = (1 / (2 \pi \sigma^2)^{1/2}) \exp(-(y(t) - \mu)^2 / (2 \sigma^2)) \quad \dots \text{ (付式G-1)}$$

(付式G-1)を使って、 $y(t)$ の T 個の時系列データ $y(T), y(T-1), \dots, y(1)$ の同時確率密度関数、言い換えれば、尤度 $L(\mu, \sigma^2)$ は、

$$L(\mu, \sigma^2) = f(y(T)) f(y(T-1)) \dots f(y(1)) = \prod_{t=1}^T (1 / (2 \pi \sigma^2)^{1/2}) \exp(-(y(t) - \mu)^2 / (2 \sigma^2)) \quad \dots \text{ (付式G-2)}$$

対数尤度 $\ln L(\mu, \sigma^2)$ は、

$$\begin{aligned} \ln L(\mu, \sigma^2) &= \ln \prod_{t=1}^T (1 / (2 \pi \sigma^2)^{1/2}) \exp(-(y(t) - \mu)^2 / (2 \sigma^2)) \\ &= T \ln (1 / (2 \pi \sigma^2)^{1/2}) - \sum_{t=1}^T (y(t) - \mu)^2 / (2 \sigma^2) \\ &= -(T/2) \ln(2 \pi) - (T/2) \ln(\sigma^2) - \sum_{t=1}^T (y(t) - \mu)^2 / (2 \sigma^2) \quad \dots \text{ (付式G-3)} \end{aligned}$$

(2) n 変数の対数尤度

n 個の確率密度関数 $\mathbf{y}(t) = [y_1(t) \ y_2(t) \ \dots \ y_n(t)]$ が正規分布 $N(\mu, \Sigma)$ に従うとき、 $\mathbf{y}(t)$ の確率密度関数は(付式G-1)を、 n 変数の同時確率密度関数に拡張して(付式G-4)で表すことができる(別途資料「いろいろな確率分布」を参照)。

$$f(\mathbf{y}(t)) = (1 / ((2 \pi)^{n/2} |\Sigma|^{1/2})) \exp(-(\mathbf{y}(t) - \mu)^T \Sigma^{-1} (\mathbf{y}(t) - \mu) / 2) \quad \dots \text{ (付式G-4)}$$

(付式G-4)を使って、 $\mathbf{y}(t)$ の時系列データ $\mathbf{y}(T), \mathbf{y}(T-1), \dots, \mathbf{y}(1)$ の同時確率密度関数、言い換えれば、尤度 $L(\mu, \Sigma)$ は、

$$L(\mu, \Sigma) = f(\mathbf{y}(T)) f(\mathbf{y}(T-1)) \dots f(\mathbf{y}(1)) = \prod_{i=1}^T (1 / ((2 \pi)^{n/2} |\Sigma|^{1/2})) \exp(-(\mathbf{y}(t) - \mu)^T \Sigma^{-1} (\mathbf{y}(t) - \mu) / 2) \quad \dots \text{ (付式G-5)}$$

対数尤度 $\ln L(\mu, \Sigma)$ は、

$$\begin{aligned} \ln L(\mu, \Sigma) &= \ln \prod_{i=1}^T (1 / ((2 \pi)^{n/2} |\Sigma|^{1/2})) \exp(-(\mathbf{y}(t) - \mu)^T \Sigma^{-1} (\mathbf{y}(t) - \mu) / 2) \\ &= T \ln (1 / ((2 \pi)^{n/2} |\Sigma|^{1/2})) - \sum_{t=1}^T (\mathbf{y}(t) - \mu)^T \Sigma^{-1} (\mathbf{y}(t) - \mu) / 2 \\ &= -(Tn/2) \ln(2 \pi) - (T/2) \ln|\Sigma| - (1/2) \sum_{t=1}^T (\mathbf{y}(t) - \mu)^T \Sigma^{-1} (\mathbf{y}(t) - \mu) \quad \dots \text{ (付式G-6)} \end{aligned}$$

付録 H. 行列のLU分解とLDU分解 (Decomposition)

(1) LU 分解

(付式H-1)に示すように、正方行列 (square matrix) **A** を下三角行列 (lower triangular matrix) **L** と上三角行列 (upper triangular matrix) **U** の積に分解するのが **LU 分解** である。
 なお、LU 分解ではどちらかの三角行列の対角成分を 1 にするという制約を課することが多い。

A = LU … (付式H-1)

ただし、
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \ell_{21} & 1 & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

(例1) $A = \begin{bmatrix} 4 & 2 & 3 \\ 4 & 6 & 10 \\ 8 & 12 & 7 \end{bmatrix}$ を LU 分解する。(付式H-1)より、 $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \ell_{21} & 1 & 0 \\ \ell_{31} & \ell_{32} & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ \ell_{21} u_{11} & \ell_{21} u_{12} + u_{22} & \ell_{21} u_{13} + u_{23} \\ \ell_{31} u_{11} & \ell_{31} u_{12} + \ell_{32} u_{22} & \ell_{31} u_{13} + \ell_{32} u_{23} + u_{33} \end{bmatrix}$

よって、
 $u_{11} = a_{11} = 4$
 $u_{12} = a_{12} = 2$
 $u_{13} = a_{13} = 3$
 $\ell_{21} u_{11} = a_{21} = 4$
 $\ell_{21} u_{12} + u_{22} = a_{22} = 6$
 $\ell_{21} u_{13} + u_{23} = a_{23} = 10$
 $\ell_{31} u_{11} = a_{31} = 8$
 $\ell_{31} u_{12} + \ell_{32} u_{22} = a_{32} = 12$
 $\ell_{31} u_{13} + \ell_{32} u_{23} + u_{33} = a_{33} = 7$

これらの式から、
 $\ell_{21} = 4 / u_{11} = 4 / 4 = 1$
 $u_{22} = 6 - \ell_{21} u_{12} = 6 - 1 \times 2 = 4$
 $u_{23} = 10 - \ell_{21} u_{13} = 10 - 1 \times 3 = 7$
 $\ell_{31} = 8 / u_{11} = 8 / 4 = 2$
 $\ell_{32} = (12 - \ell_{31} u_{12}) / u_{22} = (12 - 2 \times 2) / 4 = 2$
 $u_{33} = 7 - \ell_{31} u_{13} - \ell_{32} u_{23} = 7 - 2 \times 3 - 2 \times 7 = -13$

従って、行列 **A** は下記のように、
 下三角行列 **L** と上三角行列 **U** に分解できる (LU 分解できる)。

$$\begin{bmatrix} 4 & 2 & 3 \\ 4 & 6 & 10 \\ 8 & 12 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 2 & 3 \\ 0 & 4 & 7 \\ 0 & 0 & -13 \end{bmatrix}$$

(1) LDU 分解

(付式H-2)に示すように、正方行列 **A** を下三角行列 **L** と対角行列 (diagonal matrix) **D** と上三角行列 **U** の積に分解するのが **LDU 分解** である。

A = LDU … (付式H-2)

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix} = \begin{bmatrix} u_{11} & 0 & 0 & \cdots & 0 \\ 0 & u_{22} & 0 & \cdots & 0 \\ 0 & 0 & u_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12}/u_{11} & u_{13}/u_{11} & \cdots & u_{1n}/u_{11} \\ 0 & 1 & u_{23}/u_{22} & \cdots & u_{2n}/u_{22} \\ 0 & 0 & 1 & \cdots & u_{3n}/u_{33} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$
 と分解できるので、

$$D = \begin{bmatrix} u_{11} & 0 & 0 & \cdots & 0 \\ 0 & u_{22} & 0 & \cdots & 0 \\ 0 & 0 & u_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}, \quad \text{新たに、} U = \begin{bmatrix} 1 & u_{12}/u_{11} & u_{13}/u_{11} & \cdots & u_{1n}/u_{11} \\ 0 & 1 & u_{23}/u_{22} & \cdots & u_{2n}/u_{22} \\ 0 & 0 & 1 & \cdots & u_{3n}/u_{33} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$
 と置けば、(式H-1)から、(付式H-2)が得られる。

(例2) (例1)の行列 **A** を LDU 分解する。(例1)の結果を使用して、

$$\begin{bmatrix} 4 & 2 & 3 \\ 0 & 4 & 7 \\ 0 & 0 & -13 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 3/4 \\ 0 & 1 & 7/4 \\ 0 & 0 & 1 \end{bmatrix}$$
 従って、
$$\begin{bmatrix} 4 & 2 & 3 \\ 4 & 6 & 10 \\ 8 & 12 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -13 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 3/4 \\ 0 & 1 & 7/4 \\ 0 & 0 & 1 \end{bmatrix}$$

付録 I-1. 行列のコレスキー分解 (Cholesky Decomposition)

行列 A の LU 分解 (付録 H)、つまり、 $A = LU$ において、行列 A が **正定値対称行列** であるならば、 $A = LL^T$ と分解できる下三角行列 L が存在する。これは、LU 分解において、 $U = L^T$ といったものに相当する。これが、**コレスキー分解**である。以下、コレスキー分解が可能であることを帰納法で確認する。

(1) $n = 1$ のとき

行列 A は正定値行列であるので、 a を実数として、 $A = [a]$ (但し、 $a > 0$) である。
 ここで、 $C = [a^{1/2}]$ とすると、 $C^T = [a^{1/2}]$ であるので、 $A = [a] = [a^{1/2} \ a^{1/2}] = [a^{1/2}] [a^{1/2}] = C C^T$ となり、行列 A はコレスキー分解可能であることが確認できる。

(2) $n = k$ のとき

$k \times k$ の正定値対称行列 A_k はコレスキー分解可能であるとする。つまり、 $A_k = C_k C_k^T \cdots$ (付式 I-1) が成り立つとする。
 (付式 I-2) に示すように、この行列 A_k を使用して、行列 A を定義する。行列 A_k がコレスキー分解可能なとき、行列 A がコレスキー分解可能であることを確認する。
 また、(付式 I-3) に示すように、下三角行列 C_k を使用して、行列 D を定義する。
 さらに、(付式 I-4) に示すように、(付式 I-5) で示す、下三角行列 C_k と行列 A を構成するベクトル a から成るベクトル d を使用して、行列 E を定義する。

$$A = \left[\begin{array}{c|c} A_k & a \\ \hline a^T & \alpha \end{array} \right] \cdots \text{ (付式 I-2) }, \quad D = \left[\begin{array}{c|c} C_k & 0 \\ \hline 0^T & 1 \end{array} \right] \cdots \text{ (付式 I-3) }, \quad E = \left[\begin{array}{c|c} I_k & 0 \\ \hline d^T & 1 \end{array} \right] \cdots \text{ (付式 I-4)}$$

ただし、 $a : n \times 1$ のベクトル、 $\alpha : 1 \times 1$ のスカラー、 $I_k : k \times k$ の単位行列、 $d = C_k^{-1} a : n \times 1$ のベクトル $\cdots$ (付式 I-5)

また、行列 D の逆行列と、行列 E の逆行列を求めておく ($D D^{-1} = I_{k+1}$ 、 $E E^{-1} = I_{k+1}$ を確認することができる)。

$$D^{-1} = \left[\begin{array}{c|c} C_k^{-1} & 0 \\ \hline 0^T & 1 \end{array} \right] \cdots \text{ (付式 I-6) }, \quad E^{-1} = \left[\begin{array}{c|c} I_k & 0 \\ \hline -d^T & 1 \end{array} \right] \cdots \text{ (付式 I-7)}$$

まず、(付式 I-2) と (付式 I-6) を使用して、

$$D^{-1} A (D^{-1})^T = \left[\begin{array}{c|c} C_k^{-1} & 0 \\ \hline 0^T & 1 \end{array} \right] \left[\begin{array}{c|c} A_k & a \\ \hline a^T & \alpha \end{array} \right] \left[\begin{array}{c|c} C_k^{-1} & 0 \\ \hline 0^T & 1 \end{array} \right]^T = \left[\begin{array}{c|c} C_k^{-1} & 0 \\ \hline 0^T & 1 \end{array} \right] \left[\begin{array}{c|c} C_k C_k^T & a \\ \hline a^T & \alpha \end{array} \right] \left[\begin{array}{c|c} (C_k^{-1})^T & 0 \\ \hline 0^T & 1 \end{array} \right] = \left[\begin{array}{c|c} C_k^{-1} C_k C_k^T C_k^{-1} & C_k^{-1} a \\ \hline a^T & \alpha \end{array} \right] \left[\begin{array}{c|c} (C_k^{-1})^T & 0 \\ \hline 0^T & 1 \end{array} \right] = \left[\begin{array}{c|c} I_k & C_k^{-1} a \\ \hline a^T (C_k^{-1})^T & \alpha \end{array} \right] = \left[\begin{array}{c|c} I_k & d \\ \hline d^T & \alpha \end{array} \right]$$

さらに、(付式 I-7) を使用して、

$$\begin{aligned} E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T &= \left[\begin{array}{c|c} I_k & 0 \\ \hline -d^T & 1 \end{array} \right] \left[\begin{array}{c|c} I_k & d \\ \hline d^T & \alpha \end{array} \right] \left[\begin{array}{c|c} I_k & 0 \\ \hline -d^T & 1 \end{array} \right]^T = \left[\begin{array}{c|c} I_k & 0 \\ \hline -d^T & 1 \end{array} \right] \left[\begin{array}{c|c} I_k & d \\ \hline d^T & \alpha \end{array} \right] \left[\begin{array}{c|c} I_k & -d \\ \hline 0 & 1 \end{array} \right] = \left[\begin{array}{c|c} I_k & d \\ \hline -d^T I_k + d^T & -d^T d + \alpha \end{array} \right] \left[\begin{array}{c|c} I_k & -d \\ \hline 0 & 1 \end{array} \right] \\ &= \left[\begin{array}{c|c} I_k & d \\ \hline 0 & -d^T d + \alpha \end{array} \right] \left[\begin{array}{c|c} I_k & -d \\ \hline 0 & 1 \end{array} \right] = \left[\begin{array}{c|c} I_k & 0 \\ \hline 0 & -d^T d + \alpha \end{array} \right] \cdots \text{ (付式 I-8)} \end{aligned}$$

ここで、(付式 I-8) の $-d^T d + \alpha$ はスカラーであり、 $E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T$ は対角行列になり、対角行列の固有値は対角成分そのものである (付録 I-3)。
 また、 $E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T$ は、 $(E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T)^T = ((E^{-1})^T)^T (D^{-1})^T A^T (D^{-1})^T (E^{-1})^T = E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T$ となるから対称行列であり、また行列 A と合同な行列になる。
 一般に、正定値行列と合同な行列も正定値行列になる (付録 I-3) ので、正定値行列 A と合同な $E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T$ も正定値行列になる。また、正定値行列の固有値は正の値になるので (別途資料「統計的検定の手順」の付録を参照)、 $E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T$ の対角要素は全て正の値をとる。従って、 $(-d^T d + \alpha)^{1/2}$ が存在する。

$$G = \left[\begin{array}{c|c} I_k & 0 \\ \hline 0 & (-d^T d + \alpha)^{1/2} \end{array} \right] \cdots \text{ (付式 I-9) } \text{ と定義すると、(付式 I-8) より、} E^{-1} D^{-1} A (D^{-1})^T (E^{-1})^T = G G^T$$

よって、 $A = D E G G^T ((E^{-1})^T)^{-1} ((D^{-1})^T)^{-1} = D E G G^T ((E^{-1})^T)^{-1} ((D^{-1})^T)^{-1} = D E G G^T E^T D^T = D E G (D E G)^T$
 行列 D 、 E 、 G は下三角行列であるので、行列の積 $D E G$ も下三角行列になる。よって、 $(k + 1) \times (k + 1)$ の正定値行列 A はコレスキー分解できる。
 以上、帰納法により正定値行列はコレスキー分解できることが確認できた。

付録 I-2. 行列のコレスキー分解(計算例)

(付録 I-1)を参照して、次に示す実対称行列 **A** をコレスキー分解する。(付表 I-1)にコレスキー分解の手順を示す。結果は、 $\mathbf{A} = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 4 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

(付表 I-1)コレスキー分解の手順

	k = 1	k = 2
A	$\begin{pmatrix} 4 & 2 \\ 2 & 17 \end{pmatrix}$ $\mathbf{A}_1 = [4]$, $\mathbf{a} = [2]$, $\alpha = 17$	$\begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix}$ $\mathbf{A}_2 = \begin{pmatrix} 4 & 2 \\ 2 & 17 \end{pmatrix}$, $\mathbf{a} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$, $\alpha = 20$
C_k	$\mathbf{C}_1 \mathbf{C}_1^T = [c_{11}][c_{11}]^T = [c_{11} \ c_{11}] = [c_{11}^2] = [4]$ より、 $[c_{11}] = [4^{1/2}] = [2]$	$\mathbf{C}_2 \mathbf{C}_2^T = \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} c_{11} & c_{21} \\ 0 & c_{22} \end{pmatrix} = \begin{pmatrix} c_{11}^2 & c_{11} c_{21} \\ c_{11} c_{21} & c_{21}^2 + c_{22}^2 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 17 \end{pmatrix}$ より、 $\mathbf{C}_2 = \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix}$
d	$[2]^{-1} [2] = [2^{-1}] [2] = [2^{-1} \ 2] = [1]$	$\begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix}^{-1} \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} 0.5 & 0 \\ -0.125 & 0.25 \end{pmatrix} \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$
D	$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
E	$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$
$(-\mathbf{d}^T \mathbf{d} + \alpha)^{1/2}$	$(-[1] [1] + 17)^{1/2} = (-1 + 17)^{1/2} = 16^{1/2} = 4$	$(-\begin{pmatrix} 4 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} + 20)^{1/2} = (-16 + 20)^{1/2} = 4^{1/2} = 2$
G	$\begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$
求めたい 下三角行列 DEG	$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix}$
コレスキー分解 の確認 (DEG)(DEG)^T	$\begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix}^T = \begin{pmatrix} 2 & 0 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 17 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix}^T = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 4 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix}$

また、(付録H)の $\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{U}$ 分解で、 $\mathbf{U} = \mathbf{L}^T$ とした $\mathbf{A} = \mathbf{L} \mathbf{D} \mathbf{L}^T$ において、下三角行列 **L** の対角要素を 1 とした分解を、**修正コレスキー分解**という。上記の実対称行列を修正コレスキー分解する。

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 & 4 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0.5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{であるから、}$$
$$\begin{pmatrix} 4 & 2 & 8 \\ 2 & 17 & 4 \\ 8 & 4 & 20 \end{pmatrix} = \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & 0 \\ 4 & 0 & 2 \end{pmatrix}}_{\text{対角要素 = 1 の下三角行列}} \underbrace{\begin{pmatrix} 2 & 1 & 4 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}}_{\text{対角行列}} \underbrace{\begin{pmatrix} 1 & 0.5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{対角要素 = 1 の上三角行列}} = \begin{pmatrix} 1 & 0 & 0 \\ 0.5 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 \\ 0 & 16 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0.5 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

付録 I-3. コレスキー分解(付録 I-1)の補足

(1) 対角行列の固有値は対角成分である。

対角行列 $\mathbf{A}$ の固有値を λ とすると、 λ は特性方程式 $|\mathbf{A} - \lambda \mathbf{I}| = 0$ の解である($\mathbf{I}$ は単位行列)。

(確認)

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & \cdot & 0 \\ 0 & a_{22} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & a_{nn} \end{pmatrix} \text{ とおけば、} |\mathbf{A} - \lambda \mathbf{I}| = \left| \begin{pmatrix} a_{11} & 0 & \cdot & 0 \\ 0 & a_{22} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & a_{nn} \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & \cdot & 0 \\ 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 1 \end{pmatrix} \right| = \left| \begin{pmatrix} a_{11} - \lambda & 0 & \cdot & 0 \\ 0 & a_{22} - \lambda & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & a_{nn} - \lambda \end{pmatrix} \right| = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0$$

よって、 $\lambda = a_{11}, a_{22}, \dots, a_{nn}$ となり、対角行列の固有値は対角成分であることが分かる。

(2) 合同な行列の定義

行列 $\mathbf{A}, \mathbf{B}$ を $n \times n$ の正方行列とした場合、逆行列を持つ正則な行列 $\mathbf{P}$ によって、

$$\mathbf{B} = \mathbf{P}^T \mathbf{A} \mathbf{P} \quad \text{ただし、}\mathbf{P}^T \text{ は } \mathbf{P} \text{ の転置行列}$$

の関係があるとき、行列 $\mathbf{A}$ 行列 $\mathbf{B}$ と合同 (congruent) である、という。

この式は、行列 $\mathbf{A}$ を行列 $\mathbf{P}$ を使って、行列 $\mathbf{B}$ へ変換する式である。

(3) 実対称行列は合同変換しても実対称行列である。

行列 $\mathbf{A}$ が $n \times n$ の実対称行列である場合、逆行列を持つ正則な行列 $\mathbf{P}$ によって、

$$(\mathbf{P}^T \mathbf{A} \mathbf{P})^T = \mathbf{P}^T \mathbf{A} \mathbf{P} \quad \text{ただし、}\mathbf{P}^T \text{ は } \mathbf{P} \text{ の転置行列}$$

(確認)

行列 $\mathbf{A}$ が実対称行列であるとき、つまり、 $\mathbf{A} = \mathbf{A}^T$ であるとき、行列 $\mathbf{P}$ を使って合同変換すると、 $\mathbf{P}^T \mathbf{A}^T \mathbf{P}$ となる。

ここで、 $(\mathbf{P}^T \mathbf{A} \mathbf{P})^T = \mathbf{P}^T \mathbf{A}^T (\mathbf{P}^T)^T = \mathbf{P}^T \mathbf{A}^T \mathbf{P} = \mathbf{P}^T \mathbf{A} \mathbf{P}$ となることを確認できる。

つまり、実対称行列 $\mathbf{A}$ を合同変換した、 $\mathbf{P}^T \mathbf{A} \mathbf{P}$ は実対称行列であることが分かる。

(4) 正定値行列と合同な行列も正定値行列になる (別途資料「統計的検定の手順」の付録を参照)。

実対称行列 $\mathbf{A}$ が正定値行列の場合、つまり、 $\mathbf{A} > 0$ のとき、逆行列を持つ正則な行列 $\mathbf{P}$ によって、行列 $\mathbf{A}$ と合同な行列 $\mathbf{P}^T \mathbf{A} \mathbf{P}$ が正定値行列になる。

(確認)

行列 $\mathbf{A}$ が正定値行列であるとは、任意のベクトル $\mathbf{x}$ に対して、内積 $\langle \mathbf{x}, \mathbf{A} \mathbf{x} \rangle > 0$ が成立することをいう。

同じく、行列 $\mathbf{A}$ と合同な行列 $\mathbf{P}^T \mathbf{A} \mathbf{P}$ が正定値行列であるとは、任意のベクトル $\mathbf{x}$ に対して、内積 $\langle \mathbf{x}, \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{x} \rangle > 0$ が成立することをいう。

ところで、内積の性質から、内積 $\langle \mathbf{x}, \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{x} \rangle = \langle \mathbf{P} \mathbf{x}, \mathbf{A} \mathbf{P} \mathbf{x} \rangle$ が成り立つ。

ここで、 $\langle \mathbf{x}, \mathbf{A} \mathbf{x} \rangle > 0$ であるからベクトル $\mathbf{x}$ をベクトル $\mathbf{P} \mathbf{x}$ で置き換えれば、 $\langle \mathbf{P} \mathbf{x}, \mathbf{A} \mathbf{P} \mathbf{x} \rangle > 0$ が成立する。つまり、 $\langle \mathbf{x}, \mathbf{P}^T \mathbf{A} \mathbf{P} \mathbf{x} \rangle > 0$ が成立する。

言い換えれば、行列 $\mathbf{A}$ と合同な行列 $\mathbf{P}^T \mathbf{A} \mathbf{P}$ が正定値行列であることを示している。

付録J. 2次形式の比の最大値と最小値 (1/2)

(付式J-1)に示す、2次形式(別途資料「統計的検定の手順」を参照)の比の最大値と最小値を求める。

$$\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x}) \quad \cdots \text{ (付式J-1)}$$

ただし、 $\mathbf{x}$ は $n \times 1$ のベクトル、 $\mathbf{A}$ は $n \times n$ の正定対称行列

一般に、対称行列は直交行列を使用して対角化できる(別途資料「制御工学(現代制御)の基礎知識」を参照)。

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} = \mathbf{\Lambda} \quad \cdots \text{ (付式J-2)}$$

ただし、 $\mathbf{P}$ は $n \times n$ の直交行列(つまり、 $\mathbf{P}^{-1} = \mathbf{P}^T$)、 $\mathbf{\Lambda}$ は $n \times n$ の対角行列(対角要素は、行列 $\mathbf{A}$ の固有値)

また、一般に、行列 $\mathbf{D}$ が行列 $\mathbf{C}$ の平方根であるとは、 $\mathbf{D}^2 = \mathbf{D} \mathbf{D} = \mathbf{C}$ が成り立つときに言う。つまり、 $\mathbf{D} = \mathbf{C}^{1/2}$

$$(\mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1})(\mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1}) = \mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1} = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^{-1} = \mathbf{A} \text{ であるから、 } \mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1} = \mathbf{A}^{1/2} \quad \cdots \text{ (付式J-3)}$$

ここで、 $\mathbf{y} = \mathbf{P}^T \mathbf{x}$ とおくと、 $(\mathbf{P}^T)^{-1} \mathbf{y} = \mathbf{x} \rightarrow (\mathbf{P}^{-1})^{-1} \mathbf{y} = \mathbf{x} \rightarrow \mathbf{P} \mathbf{y} = \mathbf{x} \quad \cdots \text{ (付式J-4)}$

従って、(付式J-1)は、

$$\begin{aligned} \mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x}) &= \mathbf{x}^T \mathbf{A}^{1/2} \mathbf{A}^{1/2} \mathbf{x} / (\mathbf{x}^T \mathbf{x}) = \mathbf{x}^T (\mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1})(\mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1}) \mathbf{x} / ((\mathbf{P} \mathbf{y})^T \mathbf{P} \mathbf{y}) = \mathbf{x}^T (\mathbf{P} \mathbf{\Lambda}^{1/2} \mathbf{\Lambda}^{1/2} \mathbf{P}^{-1}) \mathbf{x} / (\mathbf{y}^T \mathbf{P}^T \mathbf{P} \mathbf{y}) \\ &= \mathbf{x}^T (\mathbf{P} \mathbf{\Lambda} \mathbf{P}^{-1}) \mathbf{x} / (\mathbf{y}^T \mathbf{P}^{-1} \mathbf{P} \mathbf{y}) = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} / (\mathbf{y}^T \mathbf{y}) \quad \cdots \text{ (付式J-5)} \end{aligned}$$

固有方程式(付式J-6)の行列 $\mathbf{A}$ の固有値は、特性方程式(付式J-7)から求めることができる。

行列 $\mathbf{A}$ は正定値行列で、正定値行列の固有値は全て正であるので(別途資料「統計的検定の手順」を参照)、行列 $\mathbf{A}$ の固有値は、 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$ と表すことができる。また、これらの固有値は対角行列 $\mathbf{\Lambda}$ の対角要素になる(別途資料「制御工学(現代制御)の基礎」を参照)。

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{x} \quad \cdots \text{ (付式J-6)}$$

$$|\lambda \mathbf{I} - \mathbf{A}| = 0 \quad \cdots \text{ (付式J-7)} \quad \text{但し、}\mathbf{I}\text{ は単位行列 (注) 一般に、行列 } \mathbf{M} \text{ の行列式は、} |\mathbf{M}| \text{ もしくは、} \det \mathbf{M} \text{ で表す。}$$

従って、(付式J-5)は(付式J-8)で表すことができる。

$$\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x}) = \mathbf{y}^T \mathbf{\Lambda} \mathbf{y} / (\mathbf{y}^T \mathbf{y}) = \frac{\sum_{i=1}^n \lambda_i y_i^2}{\sum_{i=1}^n y_i^2} \quad \cdots \text{ (付式J-8)}$$

ここで、 $\lambda_i = \lambda_1$ (固有値のなかで最大の固有値)とおけば、

$\sum_{i=1}^n \lambda_1 y_i^2 / \sum_{i=1}^n y_i^2 \geq \sum_{i=1}^n \lambda_i y_i^2 / \sum_{i=1}^n y_i^2$ であり、 $\sum_{i=1}^n \lambda_1 y_i^2 / \sum_{i=1}^n y_i^2 = \lambda_1 \sum_{i=1}^n y_i^2 / \sum_{i=1}^n y_i^2 = \lambda_1$ であるので、 $\lambda_1 \geq \sum_{i=1}^n \lambda_i y_i^2 / \sum_{i=1}^n y_i^2 = \mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x})$ つまり、 $\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x})$ の最大値は、行列 $\mathbf{A}$ の最大の固有値 λ_1 である。

$$\lambda_1 = \max (\mathbf{x}^T \mathbf{A} \mathbf{x}) / (\mathbf{x}^T \mathbf{x}) \quad \cdots \text{ (付式J-9)}$$

同様に、 $\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{x})$ の最小値は、行列 $\mathbf{A}$ の最小の固有値 λ_n である。

$$\lambda_n = \min (\mathbf{x}^T \mathbf{A} \mathbf{x}) / (\mathbf{x}^T \mathbf{x}) \quad \cdots \text{ (付式J-10)}$$

付録J. 2次形式の比の最大値と最小値 (2/2)

(付式J-11)に示す、2次形式(別途資料「統計的検定の手順」を参照)の比の最大値と最小値を求める。

$$\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{B} \mathbf{x}) \quad \cdots \text{ (付式J-11)}$$

ただし、 $\mathbf{x}$ は $n \times 1$ のベクトル、 $\mathbf{A}$ は $n \times n$ の対称行列、 $\mathbf{B}$ は $n \times n$ の正定値対称行列

(1) 一般化固有値の式を標準固有値の式へ変換する。

(付式J-12)は、一般化固有値の方程式である。この方程式を(付式J-6)に示すような標準固有値の方程式に変換する。

$$\mathbf{A} \mathbf{x} = \lambda \mathbf{B} \mathbf{x} \quad \cdots \text{ (付式J-12)}$$

λ : 行列 $\mathbf{A}$ 、 $\mathbf{B}$ の一般化固有値 $\mathbf{x}$: λ に対する一般化固有ベクトル

標準固有値の方程式に変換するため、行列 $\mathbf{B}$ を上三角行列 $\mathbf{U}$ の積にコレスキー分解し(付録 I-1を参照)、行列 $\mathbf{U}$ を使った行列 $\mathbf{D}$ とベクトル $\mathbf{y}$ を用意する。

$$\mathbf{B} = \mathbf{U}^T \mathbf{U} \quad \text{ただし、}\mathbf{U} \text{ は } n \times n \text{ の上三角行列、これは、コレスキー分解の式である。}$$
$$\mathbf{D} = (\mathbf{U}^{-1})^T \mathbf{A} \mathbf{U}^{-1}$$
$$\mathbf{y} = \mathbf{U} \mathbf{x}$$

}

$$\cdots \text{ (付式J-13)}$$

(付式J-12)を標準固有値の方程式(付式J-14)に変換する。

$$\begin{aligned} \mathbf{A} \mathbf{x} = \lambda \mathbf{B} \mathbf{x} &= \lambda (\mathbf{U}^T \mathbf{U}) \mathbf{x} = \lambda \mathbf{U}^T \mathbf{y} \rightarrow (1/\lambda) \mathbf{A} \mathbf{x} = \mathbf{U}^T \mathbf{y} \rightarrow (1/\lambda) (\mathbf{U}^T)^{-1} \mathbf{A} \mathbf{x} = \mathbf{y} \rightarrow (1/\lambda) (\mathbf{U}^{-1})^T \mathbf{A} \mathbf{x} = \mathbf{y} \\ &\rightarrow (1/\lambda) \mathbf{D} \mathbf{y} = \mathbf{y} \rightarrow \mathbf{D} \mathbf{y} = \lambda \mathbf{y} \quad \cdots \text{ (付式J-14)} \end{aligned}$$

従って、(付式J-12)の特性方程式は、 $|\lambda \mathbf{I} - \mathbf{D}| = 0 \quad \cdots \text{ (付式J-15)}$

(2) 一般固有値の特性方程式

$$\text{(付式J-12)から、} \mathbf{B}^{-1} \mathbf{A} \mathbf{x} = \lambda \mathbf{x} \quad \cdots \text{ (付式J-16)}$$

行列 $\mathbf{A}$ と $\mathbf{B}$ は対称行列、つまり正方行列であるので、特性方程式は、 $|\lambda \mathbf{I} - \mathbf{B}^{-1} \mathbf{A}| = 0 \quad \cdots \text{ (付式J-17)}$

(付式K-16)を式変形すると、 $|\mathbf{B}^{-1} (\lambda \mathbf{B} - \mathbf{A})| = |\mathbf{B}^{-1}| |\lambda \mathbf{B} - \mathbf{A}| = 0$

ここで、 $|\mathbf{B}^{-1}| \neq 0$ であるから、 $|\lambda \mathbf{B} - \mathbf{A}| = 0 \quad \cdots \text{ (付式J-18)}$

(3) 2次形式の比の最大値と最小値

(付式J-14)の行列 $\mathbf{D}$ とベクトル $\mathbf{y}$ を使用した2次形式の比は、(付式J-13)を使用して、

$$\mathbf{y}^T \mathbf{D} \mathbf{y} / (\mathbf{y}^T \mathbf{y}) = (\mathbf{U} \mathbf{x})^T ((\mathbf{U}^{-1})^T \mathbf{A} \mathbf{U}^{-1}) (\mathbf{U} \mathbf{x}) / ((\mathbf{U} \mathbf{x})^T (\mathbf{U} \mathbf{x})) = \mathbf{x}^T \mathbf{U}^T (\mathbf{U}^T)^{-1} \mathbf{A} \mathbf{U}^{-1} \mathbf{U} \mathbf{x} / (\mathbf{x}^T \mathbf{U}^T \mathbf{U} \mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{B} \mathbf{x})$$

従って、一般化固有値の式に関する2次形式の比である $\mathbf{x}^T \mathbf{A} \mathbf{x} / (\mathbf{x}^T \mathbf{B} \mathbf{x})$ の最大値と最小値を求めることは、それを変換した標準固有値の式に関する2次形式の比である $\mathbf{y}^T \mathbf{D} \mathbf{y} / (\mathbf{y}^T \mathbf{y})$ の最大値と最小値を求めることに等しいことが分かる。

よって、(付式J-11)の最大値と最小値は、特性方程式(付式J-17)、もしくは(付式J-18)から求まる最大の固有値と最小の固有値になる。

付録K. ガンマ関数の漸近展開(近似式)

ガンマ関数の近似式を求める。

ガンマ関数は、(付式K-1)で定義される(別途資料「いろいろな確率分布」参照)。

$$s > 0 \text{ に対して、 } \Gamma(s) = \int_0^\infty x^{s-1} \exp(-x) dx \quad \cdots \text{ (付式K-1)}$$

(付式K-1)に対して、 $x = \exp(t)$ とする置換積分を行う。 $dx/dt = \exp(t)$, $x: 0 \rightarrow \infty$ に対して、 $t: -\infty \rightarrow \infty$ であるから、

$$\begin{aligned} \Gamma(s) &= \int_{-\infty}^\infty \exp(t(s-1)) \exp(-\exp(t)) \exp(t) dt = \int_{-\infty}^\infty \exp(st) \exp(-t) \exp(-\exp(t)) \exp(t) dt = \int_{-\infty}^\infty \exp(st) \exp(-\exp(t)) dt \\ &= \int_{-\infty}^\infty \exp(st - \exp(t)) dt \quad \cdots \text{ (付式K-2)} \end{aligned}$$

ここで、 $f(t) = st - \exp(t)$ $\cdots$ (付式K-3) とおき、この $f(t)$ の近似式を求める。

$f(t)$ の最大値を求める。そのために、 $f(t)$ を微分して、 $f'(t) = 0$ にする t を求める。

$$f'(t) = s - \exp(t) = 0 \text{ より、 } \exp(t) = s \text{ つまり、 } t = \ln s \text{ のとき、 } f(t) \text{ は最大値 } f(\ln s) = s \ln s - \exp(\ln s) = s \ln s - s \text{ をとる。}$$

従って、(付式K-2)の被積分関数 $\exp(f(t)) = \exp(st - \exp(t))$ は、 $t = \ln s$ のとき、(付式K-4)に示す最大値を取る。

$$\exp(f(\ln s)) = \exp(s \ln s - s) = \exp(\ln s^s - s) = \exp(\ln s^s) / \exp(s) = s^s / \exp(s) = s^s / e^s = (s/e)^s \quad \cdots \text{ (付式K-4)}$$

(付図K-1)に $f(t)$ のグラフを示す。 $s \gg 1$ であれば、 $f(t)$ は最大値をとる $t = \ln s$ 以外のところでは急激に小さな値を取ることが分かる。よって、 $s \gg 1$ であれば、 $f(t)$ を $t = \ln s$ の周辺でテーラー展開で $f(t)$ を近似することができる(別途資料「いろいろな確率分布」を参照)。

まず、 $f(t)$ の導関数を求める。

$$\begin{aligned} f(t) &= st - \exp(t) \rightarrow f(\ln s) = s \ln s - \exp(\ln s) = s \ln s - s \\ f'(t) &= s - \exp(t) \rightarrow f'(\ln s) = s - \exp(\ln s) = s - s = 0 \\ f''(t) &= f'''(t) = \cdots = -\exp(t) \rightarrow f''(\ln s) = f'''(\ln s) = \cdots = -\exp(\ln s) = -s \end{aligned}$$

2 階微分の項までを近似に使用する。(付式K-2)の被積分項は、(付式K-4)を使って、

$$\begin{aligned} \exp(f(t)) &= \exp((s \ln s - s) - s/2! (t - \ln s)^2 - s/3! (t - \ln s)^3 - \cdots) \\ &\doteq \exp((s \ln s - s) - s/2! (t - \ln s)^2) \\ &= \exp(s \ln s - s) \exp((-s/2) (t - \ln s)^2) = (s/e)^s \exp((-s/2) (t - \ln s)^2) \end{aligned}$$

$$\text{よって、 } \Gamma(s) \doteq \int_{-\infty}^\infty (s/e)^s \exp((-s/2) (t - \ln s)^2) dt = (s/e)^s \int_{-\infty}^\infty \exp((-s/2) (t - \ln s)^2) dt$$

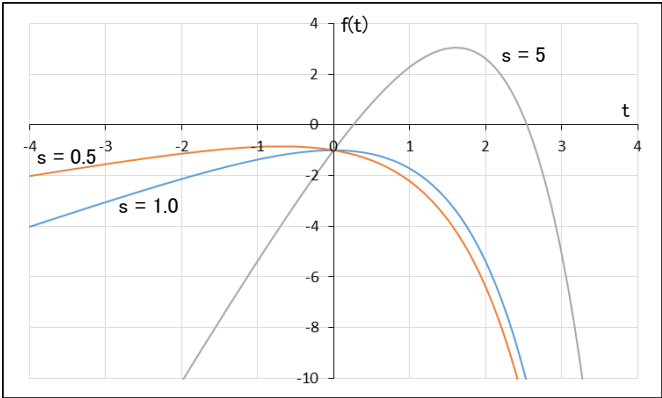
$x = t - \ln s$ とする置換積分を行う。 $dx = dt$, $t: -\infty \rightarrow \infty$ に対して、 $x: -\infty \rightarrow \infty$ であるから、ガウス積分(別途「いろいろな確率分布」を参照)を使って、

$$\Gamma(s) \doteq (s/e)^s \int_{-\infty}^\infty \exp((-s/2) x^2) dx = (s/e)^s (\pi / (s/2))^{1/2} = (s/e)^s (2\pi / s)^{1/2} \quad \cdots \text{ (付式K-5)} \quad \text{(付図K-1) } f(t) \text{ のグラフ}$$

(付式K-5)がガンマ関数の漸近展開(近似式)になる。

なお、(付式K-5)は、スターリングの公式(別途資料「いろいろな確率分布」を参照)を使って、直ちに導くこともできる。

$$\begin{aligned} \Gamma(s) &= (s-1)! \rightarrow s \Gamma(s) = s! \text{ (別途資料「いろいろな確率分布」を参照) であるので、スターリングの公式を使って、} \\ s \Gamma(s) &= s! x! \doteq (2\pi s)^{1/2} (s/e)^s \rightarrow \Gamma(s) \doteq (2\pi s)^{1/2} (s/e)^s s^{-1} = (s/e)^s (2\pi / s)^{1/2} \end{aligned}$$



付録L. 自由度が無限大のときの F 分布は、 χ^2 分布になる

確率密度関数 $f(x)$ が自由度 (m, n) の F 分布に従うとき、自由度 $n \rightarrow \infty$ にすれば、確率密度関数 $m f(x)$ は自由度 m のカイ二乗分布に従う。
つまり、 $f(x) \sim F(m, n)$ のとき、 $n \rightarrow \infty$ にすれば、 $m f(x) \sim \chi^2(m)$ が成り立つ。

(確認)
自由度 (m, n) の F 分布 $F(m, n)$ に従う確率密度関数は、(付式L-1)で表される(別途資料「いろいろな確率分布」を参照)。

$$F(m, n) = f(x) = \left\{ \frac{\Gamma((m+n)/2)}{\Gamma(m/2) \Gamma(n/2)} \right\} (m/n)^{m/2} x^{(m-2)/2} / \{1 + (m/n)x\}^{(m+n)/2} \quad \dots \text{(付式L-1)}$$

$m f(x)$ の確率密度関数を求めるため、新たな確率変数 y を $y = mx$ として、確率変数 x を確率変数 y へ変数変換する。
変数変換のためのヤコビアン $J(y)$ は、 $x = (1/m)y$ であるので、 $J(y) = |\partial x / \partial y| = 1/m = m^{-1}$
よって、確率変数 y の確率密度関数は、

$$g(y) = \left\{ \frac{\Gamma((m+n)/2)}{\Gamma(m/2) \Gamma(n/2)} \right\} (m/n)^{m/2} (y/m)^{(m-2)/2} / \{1 + (m/n)(y/m)\}^{(m+n)/2} m^{-1} \\ = \left\{ \frac{\Gamma((m+n)/2)}{\Gamma(m/2) \Gamma(n/2)} \right\} n^{-m/2} y^{(m-2)/2} / (1+y/n)^{(m+n)/2} \quad \dots \text{(付式L-2)}$$

ここで、(付式L-2)で表される $g(y)$ つまり、 $m f(x)$ が、 $n \rightarrow \infty$ のときの値を求める。

(1) (付録K)の(付式K-5)を使用して、 $\Gamma((m+n)/2) \doteq ((m+n)/2)^{(m+n)/2} (2\pi / ((m+n)/2))^{1/2}$

(2) 同様に、(付録K)の(付式K-5)を使用して、 $\Gamma(n/2) \doteq (n/2)^{n/2} (2\pi / (n/2))^{1/2}$

よって、 $\left\{ \frac{\Gamma((m+n)/2)}{\Gamma(n/2)} \right\} n^{-m/2} \doteq ((m+n)/2)^{(m+n)/2} (2\pi / ((m+n)/2))^{1/2} ((n/2)^{n/2} (2\pi / (n/2))^{1/2})^{-1} n^{-m/2}$
 $= (m+n)^{(m+n-1)/2} n^{-(m+n-1)/2} 2^{-m/2} e^{-m/2} = (m/n+1)^{(m+n-1)/2} 2^{-m/2} e^{-m/2} \quad \dots \text{(付式L-3)}$

ところで、ネイピア数 e に関して、 $\lim_{n \rightarrow \infty} (1 + x/n)^n = e^x$ が成り立つ(別途資料「いろいろな確率分布」を参照)。

従って、(付式L-3)において、 $n \rightarrow \infty$ とすると、

$$\lim_{n \rightarrow \infty} \left\{ \frac{\Gamma((m+n)/2)}{\Gamma(n/2)} \right\} n^{-m/2} = \lim_{n \rightarrow \infty} (m/n+1)^{(m+n-1)/2} 2^{-m/2} e^{-m/2} = 2^{-m/2} e^{-m/2} \lim_{n \rightarrow \infty} (m/n+1)^{(m+n-1)/2} = 2^{-m/2} e^{-m/2} \lim_{n \rightarrow \infty} \left\{ (m/n+1)^n (m/n+1)^{(m-1)/2} \right\}^{1/2} \\ = 2^{-m/2} e^{-m/2} \left\{ \lim_{n \rightarrow \infty} (m/n+1)^n \lim_{n \rightarrow \infty} (m/n+1)^{(m-1)/2} \right\}^{1/2} = 2^{-m/2} e^{-m/2} (e^m * 1)^{1/2} = 2^{-m/2} e^{-m/2} e^{m/2} = 2^{-m/2} \quad \dots \text{(付式L-4)}$$

以上から、 $n \rightarrow \infty$ のとき、(付式L-2)に(付式L-4)を使って、

$$\lim_{n \rightarrow \infty} g(y) = \lim_{n \rightarrow \infty} \left\{ \frac{\Gamma((m+n)/2)}{\Gamma(m/2) \Gamma(n/2)} \right\} n^{-m/2} y^{(m-2)/2} / (1+y/n)^{(m+n)/2} \\ = 2^{-m/2} \lim_{n \rightarrow \infty} \left\{ \frac{1}{\Gamma(m/2)} \right\} y^{(m-2)/2} / (1+y/n)^{(m+n)/2} = y^{(m-2)/2} / (2^{m/2} \Gamma(m/2)) \lim_{n \rightarrow \infty} (1+y/n)^{-(m+n)/2} \\ = y^{(m-2)/2} / (2^{m/2} \Gamma(m/2)) \lim_{n \rightarrow \infty} (1+y/n)^{-m/2} \lim_{n \rightarrow \infty} (1+y/n)^{-n/2} = y^{(m-2)/2} / (2^{m/2} \Gamma(m/2)) (1+0)^{-m/2} \lim_{n \rightarrow \infty} (1+y/n)^{-n/2} \\ = y^{(m-2)/2} / (2^{m/2} \Gamma(m/2)) \lim_{n \rightarrow \infty} (1+(y/2)/(n/2))^{-n/2} = y^{(m-2)/2} / (2^{m/2} \Gamma(m/2)) e^{-y/2} = y^{(m-2)/2} e^{-y/2} / (2^{m/2} \Gamma(m/2)) \quad \dots \text{(付式L-5)}$$

(付式L-5)は、自由度 m のカイ二乗分布の確率密度関数である(別途資料「いろいろな確率分布」を参照)。

従って、確率密度関数 $f(x)$ が自由度 (m, n) F 分布に従うとき、自由度 $n \rightarrow \infty$ にすれば、確率密度関数 $m f(x)$ は自由度 m のカイ二乗分布に従う、ことが確認できた。

付録M. 条件付き期待値と期待値の繰り返しの公式

離散型確率変数を例にとって、条件付き期待値と期待値の繰り返しの公式に関してまとめる(別途資料「いろいろな確率分布」付録を参照)。

(表記)

- ・確率変数は大文字で表す。
- ・実現値は小文字で表す。
- ・確率 $P(x)$ は、確率変数 X が実現値 x を取る確率を表す。
- ・条件付き確率 $P(y | x)$ は、確率変数 X が実現値 x を取るという条件(前提)で、確率変数 Y が実現値 y を取る確率を表す。

(1) 期待値の定義

$$E[X] = \sum_{i=1}^n x_i P(x_i) \quad \dots \quad (\text{付式M-1})$$

確率変数 X の期待値は、確率変数 X の実現値 x_i をその実現値を取り得る確率 $P(x_i)$ で加重平均したもの、つまり平均的に取る値を示す。

(2) 条件付き期待値の定義

$$E[X | y] = \sum_{i=1}^n x_i P(x_i | y) \quad \dots \quad (\text{付式M-2})$$

確率変数 X の条件付き期待値は、確率変数 Y の実現値が y のとき(これが条件)、
確率変数 X の実現値 x_i をその実現値を取り得る確率 $P(x_i | y)$ で加重平均したもの、つまり平均的に取る値を示す。

(3) 条件付き確率

$$P(x) = \sum_{j=1}^m P(x | y_j) P(y_j) \quad \dots \quad (\text{付式M-3})$$

確率変数 Y の実現値 y_j ごとの確率変数 X の実現値 x の確率を足し合わせた確率(右辺)は、確率変数 Y の全ての実現値に対する確率変数 X の実現値 x の確率(左辺)に等しい。つまり、確率変数 X の実現値が x になる確率は、確率変数 Y には依存しない。

(4) 期待値の繰り返しの公式

$$E[E[X | Y]] = E[X] \quad \dots \quad (\text{付式M-4})$$

(確認)

(付式M-1)の左辺に(付式M-2)を代入し、次いで、(付式M-2)の定義を使って、

$$E[E[X | Y]] = \sum_{j=1}^m (E[X | y_j]) P(y_j) = \sum_{j=1}^m \left(\sum_{i=1}^n x_i P(x_i | y_j) \right) P(y_j) = \sum_{i=1}^n \sum_{j=1}^m x_i P(x_i | y_j) P(y_j) = \sum_{i=1}^n x_i \sum_{j=1}^m P(x_i | y_j) P(y_j)$$

ここで、(付式M-3)を代入して、(付式M-4)が導かれる。

$$E[E[X | Y]] = \sum_{i=1}^n x_i P(x_i) = E[X]$$

付録N. vec オペレータと hvec オペレータ

(1) vec オペレータ

vec オペレータは、任意の行列の1行目から順に1列に並べたベクトルに変換するオペレータである。

(例)

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{のとき、} \text{vec}(\mathbf{A}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \\ a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

(2) hvec オペレータ

hvec (vec ハーフ) オペレータは、任意の正方行列の下三角部分にある要素を1列目から順に並べたベクトルに変換するオペレータである。

(例)

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \text{のとき、} \text{hvec}(\mathbf{A}) = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{22} \\ a_{32} \\ a_{33} \end{pmatrix}$$

付録O. マルコフ連鎖 (Markov Chain)

(1) マルコフ連鎖の定義

確率変数 $s(t)$ は、 $1, 2, \dots, M$ のうちの1つの値をとるとする。この値の集合を状態 $S = \{1, 2, \dots, M\}$ で表す。
このとき、(付式O-1)を満たす確率変数の列 $s(1), s(2), \dots$ をマルコフ連鎖という(離散的な時間軸でのマルコフ過程をマルコフ連鎖という)。
 $P(s(t) = j \mid s(t-1) = i, s(t-2) = j, s(t-3) = k, \dots) = P(s(t) = j \mid s(t-1) = i) = p_{ij} \quad \dots$ (付式O-1)

この式は、時点 t で $s(t) = j$ となる確率は 時点 $t-1$ での $s(t-1) = i$ の値だけに依存し、時点 $t-2$ 以前の $s(t-2), s(t-3), \dots$ の値には依存しないことを表している。
つまり、マルコフ過程は、未来の状態は過去の状態に依らず、現在の状態のみによって決まる過程を表している、と言える。

(2) 状態遷移図

(付式O-1)で表す状態 i から状態 j へ移る確率 p_{ij} を遷移確率といい、これをまとめた $M \times M$ の行列を遷移確率行列 P という。なお、 $p_{11} + p_{12} + \dots + p_{1M} = 1$ になる。
(付図O-1)に $M = 3$ のときの状態遷移図を表す。

(3) チャップマン・コルモゴロフ方程式 (Chapman - Kolmogorov Equation)

n 時点経過後の推移確率行列を $P^{(n)}$ で表したとき、 $P^{(n_1 + n_2)} = P^{(n_1)} P^{(n_2)} \quad \dots$ (付式O-2) が成立する。
(付式O-2)は、チャップマン・コルモゴロフ方程式と呼ばれている。

(確認) $n = 1$ のときは、定義より、 $P^{(1)} = P$ が成立する。
 $n = 2$ のときは、 $P^{(2)}_{ij} = P(s(t+2) = j \mid s(t) = i) = \sum_{k \in S} P(s(t+2) = j \mid s(t+1) = k) P(s(t+1) = k \mid s(t) = i) = \sum_{k \in S} p_{kj} p_{ik} = P^2_{ij}$
同様に、 $P^{(n)}_{ij} = P^n_{ij}$ を確認することができる。つまり、 $P^{(n)} = P^n \quad \dots$ (付式O-3) が成立する。
従って、 $P^{(n_1 + n_2)} = P^{n_1 + n_2} = P^{n_1} P^{n_2} = P^{(n_1)} P^{(n_2)}$ が成立する。

(4) 定常分布

(付式O-4)を満足する確率ベクトル π を定常分布という(確率ベクトルは、要素が確率変数で、要素の和が1になる)。
 $\pi P = \pi \quad \dots$ (付式O-4)
つまり、定常分布は、「一度その状態に入ったら、ずっとその状態であり続ける」分布である。

(例) 遷移確率行列 $P = \begin{bmatrix} 0.2 & 0.8 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.4 & 0.6 \end{bmatrix}$ のときの定常分布 π を求める。

$$\pi = [\pi_1 \quad \pi_2 \quad \pi_3] \text{ とおくと、} \pi P = [\pi_1 \quad \pi_2 \quad \pi_3] \begin{bmatrix} 0.2 & 0.8 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.4 & 0.6 \end{bmatrix} = [\pi_1 \quad \pi_2 \quad \pi_3] \quad \text{そして、} \pi \text{ は確率ベクトルであるから、} \pi_1 + \pi_2 + \pi_3 = 1$$

この連立方程式を解けば、定常分布は、 $\pi = [0.6/3 \quad 1.6/3 \quad 0.8/3] = [0.2 \quad 0.533 \quad 0.267] \quad \dots$ (付式O-5)

(参考) (付式O-4)を使用して、2 次元の推移確率行列の定常分布を求める。

$$[\pi_1 \quad \pi_2] \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = [\pi_1 \quad \pi_2], \quad p_{11} + p_{12} = 1, \quad p_{21} + p_{22} = 1, \quad \pi_1 + \pi_2 = 1 \quad \text{より、}$$
$$p_{11} \pi_1 + p_{21} \pi_2 = \pi_1 \rightarrow (1 - p_{11}) \pi_1 + p_{21} (1 - \pi_1) = 0 \rightarrow (1 - p_{11} + p_{21}) \pi_1 = p_{21}$$
$$\rightarrow (1 - p_{11} + (1 - p_{22})) \pi_1 = 1 - p_{22} \rightarrow \pi_1 = (1 - p_{22}) / (2 - p_{11} - p_{22}) \quad \dots \text{ (付式O-6)}$$
$$p_{12} \pi_1 + p_{22} \pi_2 = \pi_2 \rightarrow p_{12} (1 - \pi_2) + (p_{22} - 1) \pi_2 = 0 \rightarrow (1 + p_{12} - p_{22}) \pi_2 = p_{12}$$
$$\rightarrow (1 + (1 - p_{11}) - p_{22}) \pi_2 = 1 - p_{11} \rightarrow \pi_2 = (1 - p_{11}) / (2 - p_{11} - p_{22}) \quad \dots \text{ (付式O-7)}$$

(5) 極限分布

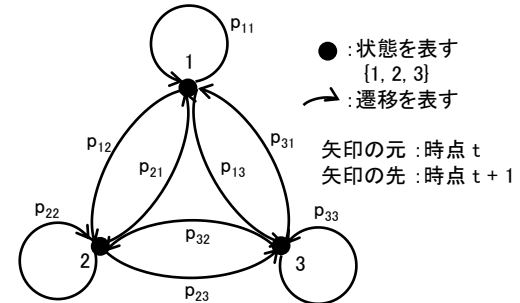
(付式O-8)の条件を満たす確率ベクトル π を極限分布という。
任意の確率ベクトル x に対して、 $\lim_{n \rightarrow \infty} x P^n = \pi \quad \dots$ (付式O-8)

つまり、極限分布は、
「どんな状態からスタートしても時間が十分に経過すれば、最終的にその状態にたどり着く」という分布である。
極限分布は定常分布である。 遷移行列が上記の例のときの極限分布を(付図O-2)に示す。

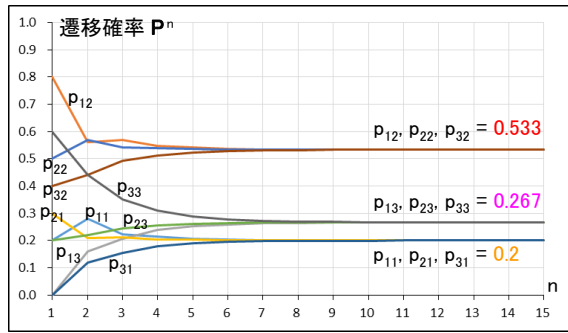
(確認)

$$\pi P = (\lim_{n \rightarrow \infty} x P^n) P = \lim_{n \rightarrow \infty} x P^{n+1} = \lim_{n \rightarrow \infty} x P^n = \pi$$

よって、極限分布 π は定常分布である。



(付図O-1) 状態遷移図の例



(付図O-2) 極限分布
(付式O-5)で求めた結果が示されている。

付録P. カイニ乗分布の要約 (別途資料「いろいろな確率分布」参照)

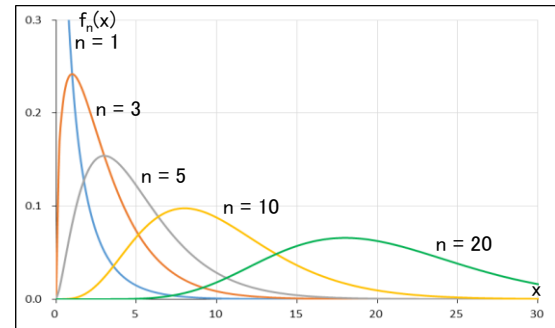
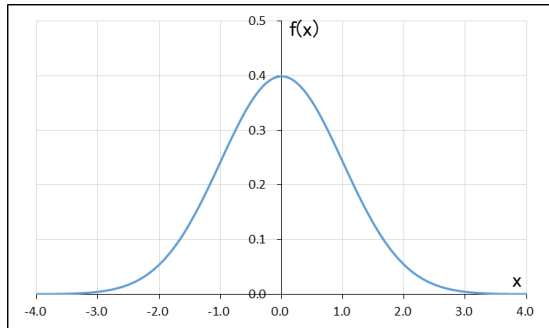
■概要

互いに独立な確率変数 $X_1, X_2, \dots, X_n$ が標準正規分布 ($\mu = 0, \sigma^2 = 1$) に従うとき、
確率変数 $X_1^2 + X_2^2 + \dots + X_n^2$ が従う確率分布を、自由度 n (n は和の数に相当) のカイニ乗分布という。

確率変数 X_i が、標準正規分布に従う

ならば、

$\sum_{i=1}^n X_i^2$ は、自由度 n のカイニ乗分布に従う



(付図P)カイニ乗分布の導出

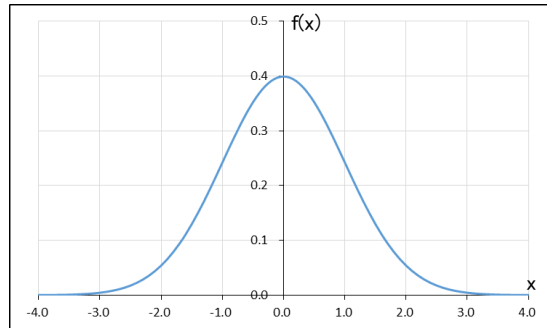
付録Q. t 分布の要約

(別途資料「いろいろな確率分布」参照)

■概要

t 分布は、標準正規分布を自由度を加味したカイニ乗分布でスケーリングした確率分布である。

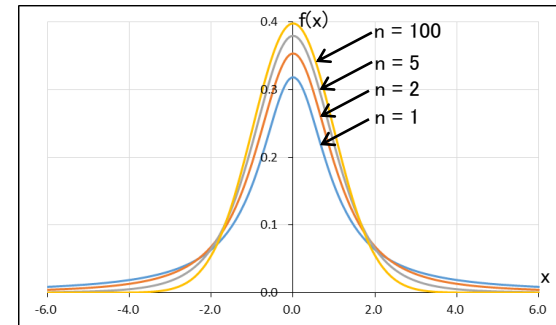
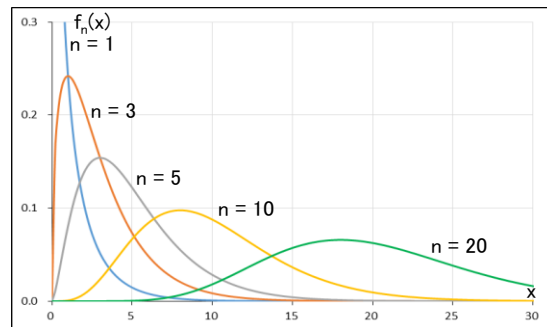
確率変数 X が、標準正規分布に従う



ならば、

$X / (Y / n)^{1/2}$ は、自由度 n の t 分布に従う

確率変数 Y が、自由度 n のカイニ乗分布に従う



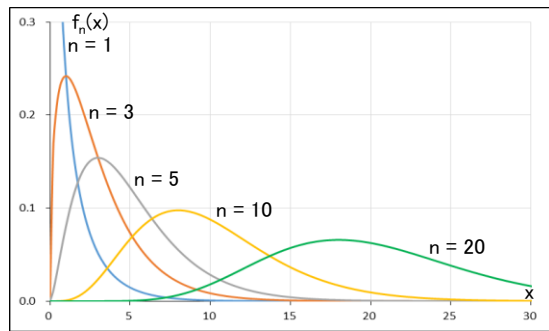
(付図Q)t 分布の導出

付録R. F 分布の要約 (別途資料「いろいろな確率分布」参照)

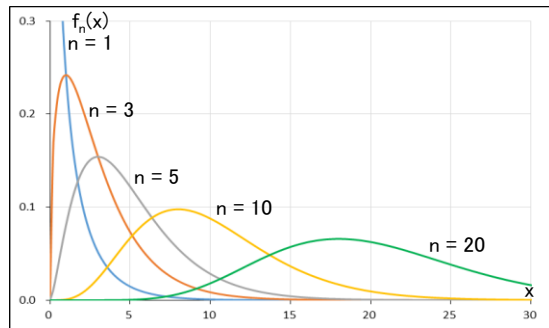
■概要

自由度 m のカイニ乗分布に従う確率変数と自由度 n のカイニ乗分布に従う確率変数が互いに独立であるとき、自由度を加味した確率変数の比が従う確率分布を、パラメータとして自由度 (m, n) の F 分布という。

確率変数 X が、自由度 m のカイニ乗分布に従う

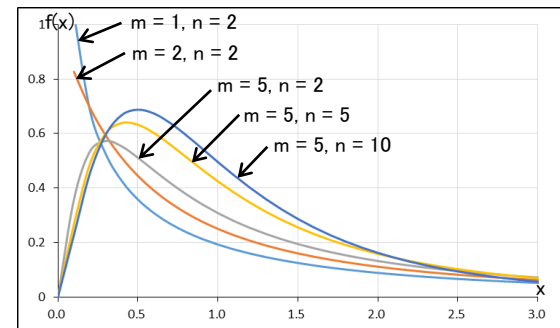


確率変数 Y が、自由度 n のカイニ乗分布に従う



ならば、

$(X / m) / (Y / n)$ は、自由度 (m, n) の F 分布に従う



参考にした資料

阿部真人:「統計学入門、仮説検定から統計モデリングまで重要トピックを完全網羅」、ソシム

石村貞夫、石村友二郎:「入門はじめての時系列分析」、東京図書

沖本竜義:「経済・ファイナンスデータの計量時系列分析」、朝倉書店

Peter J. Brockwell (Colorado State University), Richard A. Davis (Columbia University):「Introduction to Time Series and Forecasting」、[Introduction to Time Series and Forecasting | SpringerLink](#)

足立修一(慶応大学理工学部):「第十二回 Z変換 1、2」、[慶應大学講義 理工学部 物理情報数学C 第十二回 Z変換 1 - YouTube](#)

あつまれ統計の森:「偏相関係数の定義と導出方法を把握する」、[偏相関\(Partial Correlation\)係数の定義と導出方法を把握する - あつまれ統計の森 \(hello-statisticians.com\)](#)

あつまれ統計の森:「偏自己相関係数の定義や計算の流れを確認する」、[偏自己相関\(Partial AutoCorrelation\)の定義や計算の流れを確認する - あつまれ統計の森 \(hello-statisticians.com\)](#)

谷崎久志(大阪大学経済学部):「時系列解析(非定常時系列)」、[coint.pdf \(osaka-u.ac.jp\)](#)

School of Mathematics and Statistics (The University of Sydney):「Invertibility of MA Processes」、[NOTES4.pdf \(usyd.edu.au\)](#)

Richard Lockhart (Simon Fraser University, Canada):「Identifiability, Invertibility」、[web.dvi \(sfu.ca\)](#)

長倉大輔(慶応義塾大学経済学部):「時系列分析の基礎」、[ファイナンスのための数学基礎 第1回 オリエンテーション、ベクトル \(keio.ac.jp\)](#)

hKawabata:「Ljung-Box 検定」、[Ljung-Box 検定 | technical-note \(hkawabata.github.io\)](#)

北川源四郎(東京大学数理・情報センター):「時系列解析(1)〜(13)」、[Microsoft PowerPoint - 時系列解析\(1\) 2019修正版.pptx \(u-tokyo.ac.jp\)](#)

講義資料(京都大学):「補足解説1 Levinson-Durbin のアルゴリズム」、[untitled \(kyoto-u.ac.jp\)](#)

Gidon Eshel (University of Pennsylvania):「The Yule Equations for the AR Coefficients」、[YW-Eshel.pdf \(upenn.edu\)](#)

Lesse Engbo Christiansen (University of Denmark):「Estimating parameters in ARMA models」、[02417 Lecture 7 part A: Estimating parameters in ARMA models - YouTube](#)

ウサギさんの統計学サロン:「自由度が無限大のときのF分布 カイニ乗分布への収束」、[【統計学】自由度が無限大のときのF分布 カイニ乗分布への収束 \(multivariate-statistics.com\)](#)

理数アラカルト:「行列の合同とは? ~具体例と性質~」、[行列の合同とは? - 理数アラカルト - \(risalc.info\)](#)

理数アラカルト:「LU分解を解説~具体例と必要十分条件~」、[LU分解を解説 ~具体例と必要十分条件~ - 理数アラカルト - \(risalc.info\)](#)

理数アラカルト:「コレスキー分解とは?」、[コレスキー分解とは? ~解説・具体例~ - 理数アラカルト - \(risalc.info\)](#)

谷崎久志(大阪大学経済研究科):「グレンジャー因果性テスト」、「インパルス応答関数」、[20210701\[325\].pdf](#)

川畑康治(神戸大学経済学部):「分布ラグモデル」、[5 モデルの関数型と特殊な変数 \(kobe-u.ac.jp\)](#)

Alcia Solid Project:「VAR過程と因果について【Granger因果検定】」、[【時系列分析④】VAR過程と因果について【Granger 因果検定】#VRアカデミア #036 \(youtube.com\)](#)

Alcia Solid Project:「インパルス応答関数を用いて時系列データの関係を定量的に求める」、[【時系列分析⑤】インパルス応答関数を用いて時系列データの関係を定量的に求める #VRアカデミア #038 - YouTube](#)

AVILEN AI Trend:「時系列分析の単位根過程、ランダムウォークとは?」、[時系列分析の単位根過程、ランダムウォークとは? | AVILEN AI Trend \(ai-trend.jp\)](#)

Real STATISTICS USING EXCEL:「Dickey-Fuller Test」、[Dickey-Fuller Test | Real Statistics Using Excel \(real-statistics.com\)](#)

Ritvik Kharkar (Senior Data Scientist):「Time Series Talk : Augmented Dickey Fuller Test + Code」、[Time Series Talk : Augmented Dickey Fuller Test + Code - YouTube](#)

Qiita:「ベクトルで微分・行列で微分 公式のまとめ」、[「ベクトルで微分・行列で微分」公式まとめ #機械学習 - Qiita](#)

Future fisheries:「二次形式の最大と最小」、[Microsoft Word - V-2-6. ææâ § ä...ææä°, \(future-fisheries.jp\)](#)

桂田祐史(明治大学総合数理学部):「一般化固有値問題」、[generalized-eigenvalue-problem.pdf \(meiji.ac.jp\)](#)

田中勝人(一橋大学大学院経済研究科):「計量経済学のテキストを書き換えた「共和分」と「単位根」の概念」、[0100904001.pdf \(hit-u.ac.jp\)](#)

川崎能典(東京大学大学院経済学研究科):「Johansen の共和分検定について」、[Johansenの共和分検定について \(boj.or.jp\)](#)

Ragnar Nymoen (University of Oslo):「Cointegration, estimation and testing」、[Lect12v11.pdf \(uio.no\)](#)

Chingnun Lee (Sun Yat-sen University, Taiwan):「Johansen's MLE for Cointegration」、[Chapter24 Johansen-quot-sMLEforointegration.pdf \(nysyu.edu.tw\)](#)

しまつま総研:「繰り返し期待値の法則について/回帰モデル」、[繰り返し期待値の法則について / 回帰モデル | しまつま総研 \(info-zebra.com\)](#)

長倉大輔(慶応大学経済学部):「ARCH, GARCHモデルとその性質1、2」、[ARCH, GARCHモデルとその性質1 \(keio.ac.jp\)](#) [ARCH, GARCHモデルとその性質 \(keio.ac.jp\)](#)

長倉大輔(慶応大学経済学部):「多変量GARCHモデル」、[multivariate GARCH.pdf \(keio.ac.jp\)](#)

村澤康友(甲南大学経済学部):「第13回ARCH・GARCH過程」、[ec2-ln13.pdf \(fc2.com\)](#)

Hatena Blog:「時系列解析(AR, ARCH, GARCH) 実装編」、[時系列解析\(AR, ARCH, GARCH\)-実装編 - どうでもしか勝たん \(hatenablog.com\)](#)

森平爽一郎(早稲田大学ファイナンス研究所):「コピュラ:信用リスク管理の新たな視点」、[kaidai_140301.pdf \(saa.or.jp\)](#)

長倉大輔(慶応大学経済学部):「状態変化を伴うモデル」、[状態変化を伴うモデル+ \(keio.ac.jp\)](#)

高校数学の美しい物語:「マルコフ連鎖の基本とコルモゴロフ方程式」、[マルコフ連鎖の基本とコルモゴロフ方程式 | 高校数学の美しい物語 \(manabitemes.jp\)](#)

Aptech Systems, Inc.:「Introduction to Markov-Switching Models」、[Introduction to Markov-Switching Models | Aptech](#)

R-bloggers:「Kim(1994) Smoother Algorithm in Regime Switching Model using R code」、[Kim \(1994\) Smoother Algorithm in Regime Switching Model using R code | R-bloggers](#)

沖本竜義(クロフォード公共政策大学院):「マルコフスイッチングモデルのマクロ経済・ファイナンスへの応用」、[日本統計学会誌, 第44巻, 第1号, 137頁-157頁 \(jst.go.jp\)](#)

Qiira:「EMアルゴリズム徹底解説」、[EMアルゴリズム徹底解説 #Python - Qiita](#)

石原康博(高崎経済大学)、渡部敏明(一橋大学ソーシャル・データサイエンス研究科):「景気循環の計量分析」、[keizaikenkyu06602145.pdf \(hit-u.ac.jp\)](#)