Box-Jenkins の方法

自己回帰（AR）

任意の時系列を y_t で表す。これが p 個の過去の値に依存する時、これを次数 p の自己回帰過程（AR）と呼ぶ。

$$\mathrm{AR}(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \nu_t$$

または

$$\mathrm{AR}(p): y_t = \sum_{m=1}^{p} \phi_m y_{t-m} + \nu_t$$

ここで、ν_t は時間の経過に対して不変な分布を持つ系列相関のない撹乱＝誤差項である。

期待値 0、一定の分散 σ^2 、また ν_t と ν_{t+h} は無相関である $(h \neq 0)$ と仮定される。すなわち、

$$\mathrm{E}[\nu_t] = 0$$

$$\mathrm{E}[\nu_t \nu_s] = \begin{cases} \sigma_\nu^2 & t = s \\ 0 & t \neq s \end{cases}$$

$$\mathrm{E}[\nu_t \nu_{t-m}] = 0 \qquad m = 1,2,\ldots,p$$

である。

移動平均（MA）

次数 q の移動平均（MA）は以下で定義される。

$$\mathrm{MA}\ (q): y_t = \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} \ldots - \theta_q \nu_{t-q}$$

θ の符号はプラスで表現する場合もある。ν_t の条件は AR と同様である。

自己回帰移動平均（ARMA）

AR と MA の混合である自己回帰移動平均（ARMA）は以下で定義される。

$$\mathrm{ARMA}(p,q): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots + \phi_p y_{t-p} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} \ldots - \theta_q \nu_{t-q}$$

時系列の自己相関

時系列の観測点 y_t は確率変数であり、平均 $\mathrm{E}(y_t)$ 、分散 $\mathrm{var}(y_t)$ の確率分布 $p(y_t)$ に従う。またラグ l に依存する y_t と y_{t-l} の共分散もあり、それを自己共分散 γ_l と呼び以下で表す。

$$\gamma_l \equiv \mathrm{cov}(y_t, y_{t-l})$$

特別な場合として $l = 0$ のとき、

$\gamma_0 = \mathrm{cov}(y_t, y_t) = \mathrm{var}(y_t)$

である。同様に自己相関 ρ_t は、

$$\rho_t = \frac{\mathrm{cov}(y_t, y_{t-l})}{\sqrt{\mathrm{var}(y_t) \quad \mathrm{var}(y_{t-l})}}$$

分散が時間の経過と共に変化しなければ、$\mathrm{var}(y_t) = \mathrm{var}(y_{t-l})$ であり、自己相関は以下になる。

$$\rho_t = \frac{\mathrm{cov}(y_t, y_{t-l})}{\mathrm{var}(y_t)} = \frac{\gamma_l}{\gamma_0}$$

自己相関の計算（MA）

以下の一次形を考える。

$y_t = \nu_t - \theta \nu_{t-1} (\theta \equiv \theta_1)$

各 ν_t の平均は 0 だから y_t もまた平均 0 を持つ。分散は、

$$\mathrm{var}\, y_t = \mathrm{E}(y_t^2) = \mathrm{E}(\nu_t{}^2 + \theta^2 \nu_{t-1}^2 - 2\theta \nu_t \nu_{t-1})$$

ここで $\mathrm{E}(\nu_t^2) = \mathrm{E}(\nu_{t-1}^2)$ である。また $\mathrm{E}[\nu_t \nu_{t-1}] = 0$ である。したがって、

$$\mathrm{var}\, y_t = \sigma^2 + \theta^2 \sigma^2 + 0 = (1 + \theta^2)\sigma^2$$

この y_t の分散は、時刻 t に依存しない。ラグ 1 の自己共分散は以下になる。

$$\mathrm{E}(y_t y_{t-1}) = \mathrm{E}(\nu_t - \theta \nu_{t-1})(\nu_{t-1} - \theta \nu_{t-2})$$
$$= \mathrm{E}(\nu_t \nu_{t-1} - \theta \nu_{t-1}^2 - \theta \nu_t \nu_{t-2} + \theta^2 \nu_{t-1} \nu_{t-2})$$
$$= 0 - \theta \sigma^2 + 0 + 0$$

y_t の分散とラグ 1 の自己共分散が求まったので、ラグ 1 の自己相関係数が求まる。

$$\rho_1 = \frac{-\theta \sigma^2}{(1 + \theta^2)\sigma^2} = \frac{-\theta}{1 + \theta^2}$$

また、$\rho_2 = \rho_3 = \ldots = 0$ である。この二つの式は一次の移動平均に対する自己相関係数の全体を定めており、時間に依存しないのでこの移動平均は弱定常であるという。

一般に次数 q の移動移動平均は弱定常であり、自己相関係数が以下で与えられる。

$\mathrm{MA}(q)$ に対して、

$$\rho_l = \begin{cases} \dfrac{-\theta_l + \theta_l \theta_{l+1} + \theta_2 \theta_{l+2} + \ldots + \theta_{q-l} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2} & l = 1, 2, \ldots, q \\ 0 & l > q \end{cases}$$

ラグが移動平均の次数を超えれば自己相関は 0 になる。

自己相関の計算（AR）

一般的な 1 次の自己回帰を考える。

$y_t = \nu_t + \phi y_{t-1} (\phi \equiv \phi_1)$　　（1）

t のかわりに $t-1$ を代入する。

$y_{t-1} = \nu_{t-1} + \phi y_{t-2}$　　（2）

（2）を（1）に代入する。

$y_t = \nu_t + \phi[\nu_{t-1} + \phi y_{t-2}]$

さらに代入を繰り返していくと、

$y_t = \nu_t + \phi[\nu_{t-1} + \phi(\nu_{t-2} + \phi y_{t-3})]$

$y_t = \nu_t + \phi\nu_{t-1} + \phi^2\nu_{t-2} + ...$

となる。

各 ν_t の平均は 0 だから y_t もまた平均 0 を持つ。また ν_t は無相関であるので、

$$\mathrm{var}(y_t) = \mathrm{var}\,\nu_t + \phi^2\,\mathrm{var}\,\nu_{t-1} + \phi^4\,\mathrm{var}\,\nu_{t-2} + ...$$
$$= \sigma^2 + \phi^2\sigma^2 + \phi^4\sigma^2 + ...$$

という無限幾何級数になり、係数 ϕ が $|\phi| < 1$ を満たすとき、y_t の分散は以下に収束する。

$$\mathrm{var}(y_t) = \frac{\sigma^2}{1-\phi^2}$$

$|\phi| \geq 1$ の時は y_t は定常でなく、爆発的に増大する系列である。

また自己共分散は、y_t の平均は 0 だから、

$\gamma_l \equiv \mathrm{cov}(y_t, y_{t-l}) \equiv \mathrm{E}(y_t y_{t-l})$

$= \mathrm{E}(\phi y_{t-1} + \nu_t)(y_{t-l})$

$= \phi\mathrm{E}(y_{t-1}y_{t-l}) + \mathrm{E}(\nu_t y_{t-l})$

ここで y_{t-l} は ν_{t-l} や ν_{t-l-1} などの関数であり、ν_t とは無相関である。従って $\mathrm{E}(\nu_t y_{t-l}) = 0$ である。すなわち、

$\gamma_l = \phi\mathrm{E}(y_{t-1}y_{t-l})$

ここで y_{t-l} は y_{t-1} より $(t-1)-(t-l) = l-1$ だけ後方にある。よって、

$\mathrm{E}(y_{t-1}y_{t-l}) = \gamma_{l-1}$

であり、最後の二つの式から、

$\gamma_l = \phi\gamma_{l-1}$

自己相関は、

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \frac{\phi\gamma_{l-1}}{\gamma_0} = \phi\rho_{l-1}$$

である。すなわちこれは指数関数的に減衰する等比級数で、

$\rho_l = \phi^l$ である。

次数 p の一般的な自己回帰に関しても類似した関係が示せる。

$$\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2} + \dots \phi_p \rho_{l-p} \quad l > 0$$

これは p 次の差分方程式である。

自己相関の計算（ARMA）

$\mathrm{ARMA}(p,q)$ 過程、

$$\mathrm{ARMA}(p,q): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} \dots - \theta_q \nu_{t-q}$$

は最初の q 個の系列 $(1 \le l \le q)$ に関して複雑な自己相関係数を持つ。しかし $l > q$ に対しては MA の影響がなくなり、AR と同じ差分方程式を満足する。

$$\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2} + \dots \phi_p \rho_{l-p}$$

自己相関の要約

時系列モデル y_t	自己相関係数 ρ_l
$\mathrm{MA}\ (q): y_t = \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} \dots - \theta_q \nu_{t-q}$	$\rho_l = \begin{cases} \dfrac{-\theta_l + \theta_l \theta_{l+1} + \theta_2 \theta_{l+2} + \dots + \theta_{q-l} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & l = 1,2,\dots,q \\ 0 \quad l > q \end{cases}$
$\mathrm{AR}(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \nu_t$	$\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2} + \dots \phi_p \rho_{l-p} \quad l > 0$
$\mathrm{ARMA}(p,q): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots$ $+ \phi_p y_{t-p} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2} \dots - \theta_q \nu_{t-q}$	$\rho_l = \theta_i$とϕ_iの複雑な関係 $\quad (1 \le l \le q)$ $\rho_l = \phi_1 \rho_{l-1} + \phi_2 \rho_{l-2} + \dots \phi_p \rho_{l-p} \quad l > q$

非定常系列の差分（ARIMA）

差分を d 回取ると、非定常系列が定常な $\mathrm{ARMA}(p,q)$ になると仮定する。その時 d 回の加え合わせによって元の系列を再現できる。これを積分された AMRA 系列と呼び、それを $\mathrm{ARIMA}(p,d,q)$ と表現する。

ARIMA モデルの推定と予測

現実の時系列を ARIMA モデルで構成するためには次の 3 段階が必要である。

1. 特定化（又は識別）・・・差分の数 d 、ラグ p と q の大きさの決定

2.　パラメータ $\phi_1,...,\phi_p$ と $\theta_1,...,\theta_p$ の推定

3.　当てはめたモデルが適当かどうかを確かめるための診断

特定化

非定常系列が定常系列になるまで差分を繰り返して d を決める。ラグ p と q を決めるためには要約統計量を用いる。サンプルの自己相関係数を以下で定義する。

$$\hat{\rho}_l \equiv \frac{\sum_{t=l+1}^{n} y_t y_{t-l}}{\sum_{t=1}^{n} y_t^2}$$

これは、

$$\rho_t = \frac{\mathrm{cov}(y_t, y_{t-l})}{\sqrt{\mathrm{var}(y_t) \quad \mathrm{var}(y_{t-l})}}$$

からきている。ラグ $l = 1,2,3,...$ に関するこの値を見て特定化するが、いずれも厳密な方法ではない。

パラメータ推定

$\mathrm{ARMA}(p,q)$ 過程のパラメータ推定を行う。自己相関係数 ρ_l が、$l = 1,2,...p+q$ に対してサンプルの自己相関係数 $\hat{\rho}_l$ の値にちょうど等しくなるように $p+q$ 個のパラメータ ϕ_i と θ_i を選ぶ。これは $p+q$ 個の未知変数からなる $p+q$ 本の非線形方程式である。しかしこの方法は予測を意識していないので、以下の別の方法を取る。

$\mathrm{AR}(p)$ の場合

$$y_t = \nu_t + \phi\nu_{t-1} + \phi^2\nu_{t-2} + ...$$

の時、 y_t は撹乱項の現在の値、$\nu_t, \nu_{t-1},....$ のみに影響される。従って、 ν_t は過去の y すなわち $y_{t-1}, y_{t-2},...$ とは独立であり、

$$\mathrm{AR}(p): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \nu_t$$

に OLS を適用することによって ϕ_i の一致推定量が得られる。

$\mathrm{ARMA}(p,q)$ の場合

通常 q は小さく、q が小さければ OLS の特別な場合が推定法として利用可能である。
$p = 3, q = 2$ の場合を考える。

$\mathrm{ARMA}(3,2): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2}$
$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \phi_3 y_{t-3} = \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2}$
左辺はy_tの線形変換なので$\mathbf{T}y_t$と表す。
$\mathbf{T}y_t = \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2}$
すなわち、

$$y_t = \mathbf{T}^{-1}(\nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2})$$

$\mathbf{T}$が線形なので$\mathbf{T}^{-1}$も線形である。従って、

$$y_t = \mathbf{T}^{-1}\nu_t - \mathbf{T}^{-1}\theta_1 \nu_{t-1} - \mathbf{T}^{-1}\theta_2 \nu_{t-2}$$

$z_t \equiv \mathbf{T}^{-1}\nu_t$とすると、

$y_t = z_t - \theta_1 z_{t-1} - \theta_2 z_{t-2}$
$z_t = y_t + \theta_1 z_{t-1} + \theta_2 z_{t-2}$
となる。$z_0 = z_{-1} = 0$（他に情報がないため、期待値を 0 とおく）を初期値として、この式から系列z_tを再現する。その際、パラメータθ_1、θ_2の試験的な値も与える。

ここで$z_t \equiv \mathbf{T}^{-1}\nu_t$を以下に書き換える。

$\mathbf{T}z_t = \nu_t$
$\mathbf{T}$は任意の系列に適用できるので、系列z_tにも適用できる。つまり、
$z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \phi_3 z_{t-3} = \nu_t$
$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \nu_t$
先ほど作られた系列z_tを用いて、OLS によりパラメータϕ_iを推定する。
この自己回帰からの推定された残差ν_tを 2 乗し、加え合わせて以下の基準を作る。

$$S = \sum \hat{\nu}_t^2$$

つまり、試験的なパラメータ(θ_1, θ_2)はパラメータ(ϕ_1, ϕ_2, ϕ_3)の推定された組と残差二乗和Sを与える。

あらゆる(θ_1, θ_2)を試し、Sを最小にする(θ_1, θ_2)および(ϕ_1, ϕ_2, ϕ_3)を決定する。

ARMA モデルの予測

$t = 1,2,\ldots,n$に対するデータが、
$\mathrm{ARMA}(2,2): y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \nu_t - \theta_1 \nu_{t-1} - \theta_2 \nu_{t-2}$

を当てはめるために用いられていると仮定する。
一番近い将来の値 y_{n+1} を予測するためには、論理式として以下を得る。

$$y_{n+1} = \phi_1 y_n + \phi_2 y_{n-1} + \nu_{n+1} - \theta_1 \nu_n - \theta_2 \nu_{n-1}$$

ここで、推定されたパラメータ $(\hat{\phi}_1, \hat{\phi}_2, \hat{\theta}_1, \hat{\theta}_2)$ と、推定された残差 $\hat{\nu}_n, \hat{\nu}_{n-1}$ を用いる。将来の残差 ν_{n+1} は期待値が 0 であることしかわからないので、$\nu_{n+1} = 0$ とする。過去の値 y_n と y_{n-1} は知られているので、以下で予想できる。

$$y_{n+1} = \hat{\phi}_1 y_n + \hat{\phi}_2 y_{n-1} + 0 - \hat{\theta}_1 \hat{\nu}_n - \hat{\theta}_2 \hat{\nu}_{n-1}$$

その次の将来の値 y_{n+2} を予測するためには、論理式として以下を用いる。

$$y_{n+2} = \phi_1 y_{n+1} + \phi_2 y_n + \nu_{n+2} - \theta_1 \nu_{n+1} - \theta_2 \nu_n$$

同様に $\nu_{n+1} = 0$ であり、また $\nu_{n+2} = 0$ とも推定し、y_{n+1} は直前の式から推定される。したがって、

$$y_{n+2} = \hat{\phi}_1 \hat{y}_{n+1} + \hat{\phi}_2 y_n + 0 - 0 - \hat{\theta}_2 \hat{\nu}_n$$

さらにもう一つの将来の値を予測するためには、同じ議論により、

$$y_{n+3} = \hat{\phi}_1 \hat{y}_{n+2} + \hat{\phi}_2 \hat{y}_{n+1}$$

であり、ここに至って移動平均の部分がなくなって、予測は純粋に自己回帰となる。