Modeling and Forecasting with ARMA models

1 Method of Moments Estimator: Yule-Walker Estimation

AR 모형에서 많이 사용

(1) AR models

- consider AR(p) process:

* $EX_tZ_t = E(\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t)Z_t = \sigma^2$

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + Z_t$$

- Then, we know

$$\begin{aligned}
Cov(X_t, X_t)\ \gamma(0) &= \phi_1\gamma(1) + \phi_2\gamma(2)\cdots + \phi_p\gamma(p) + \underline{\sigma_w^2} \quad // \ EX_tZ_t \leftarrow \gamma(0)\text{일때만 살음.} \\
Cov(X_t, X_{t+1})\ \gamma(1) &= \phi_1\gamma(0) + \phi_2\gamma(-1)\cdots + \phi_p\gamma(1-p) \quad EX_tZ_{t-1}\text{ 사라짐.} \\
\vdots &= \vdots \qquad \vdots \qquad \vdots \\
Cov(X_t, X_{t+p})\ \gamma(p) &= \phi_1\gamma(p-1) + \phi_2\gamma(p-2) + \cdots + \phi_p\gamma(-1)
\end{aligned}$$

- Goal: estimation of $\phi_1, \phi_2, \ldots, \phi_p$.
- Replace population ACF $\gamma(h)$ to sample ACF $\widehat{\gamma}(h)$

* $\gamma(k) = EX_tX_{t-k}$

$\downarrow$

$\hat{\gamma}(k) = \frac{1}{T}\sum_{t=k+1} X_t X_{t-k}$

$$\begin{aligned}
\widehat{\gamma}(0) &= \phi_1\widehat{\gamma}(1) + \phi_2\widehat{\gamma}(2)\cdots + \phi_p\widehat{\gamma}(p) + \sigma_w^2 \\
\widehat{\gamma}(1) &= \phi_1\widehat{\gamma}(0) + \phi_2\widehat{\gamma}(-1)\cdots + \phi_p\widehat{\gamma}(1-p) \\
\vdots &= \vdots \qquad \vdots \qquad \vdots \\
\widehat{\gamma}(p) &= \phi_1\widehat{\gamma}(p-1) + \phi_2\widehat{\gamma}(p-2) + \cdots + \phi_p\widehat{\gamma}(-1)
\end{aligned}$$

- Example with R code: AR (2)

* AR(2)

$\hat{\gamma}(0) = 8.434$

$\hat{\rho}(1) = 0.834$

$\hat{\rho}(2) = 0.476$

) 주어졌을 때 σ^2 추정.

$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + Z_t$

$\gamma(1) = \phi_1\gamma(0) + \phi_2\gamma(1)$

$\gamma(2) = \phi_1\gamma(1) + \phi_2\gamma(0)$

$\begin{pmatrix} \gamma(0) & \gamma(1) \\ \gamma(1) & \gamma(0) \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \gamma(2) \end{pmatrix}$

$= \gamma(0)\begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \gamma(1) \\ \gamma(2) \end{pmatrix}$

$\rightarrow \begin{pmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} \rho(1) \\ \rho(2) \end{pmatrix}$

$\begin{pmatrix} 1 & \hat{\rho}(1) \\ \hat{\rho}(1) & 1 \end{pmatrix}\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{pmatrix} = \begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \end{pmatrix}$

$\begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \end{pmatrix} = \frac{1}{1-\hat{\rho}(1)^2}\begin{pmatrix} 1 & -\hat{\rho}(1) \\ -\hat{\rho}(1) & 1 \end{pmatrix}\begin{pmatrix} \hat{\rho}(1) \\ \hat{\rho}(2) \end{pmatrix} = \begin{pmatrix} 1.439 \\ -0.725 \end{pmatrix}$

$\therefore \hat{\sigma}_w = \hat{\gamma}(0) - \hat{\phi}_1\hat{\gamma}(1) - \hat{\phi}_2\hat{\gamma}(2)$

$= \hat{\gamma}(0)\left(1 - \hat{\phi}_1\hat{\rho}(1) - \hat{\phi}_2\hat{\rho}(2)\right) = 1.215$

- Large Sample Distribution of Yule-Walker Estimator
  For a large sample from AR (p) process,

$$\widehat{\phi} \approx \text{Normal}\left(\phi, \quad \frac{1}{n}\sigma^2\Gamma_p^{-1}\right)$$

Fact로 받아들이기

$$\hat{\phi} \sim N\left(\phi, \frac{1}{n}\sigma^2 \begin{pmatrix} \gamma(0) & \cdots & \gamma(n-1) \\ \gamma(1) & \ddots & \gamma(n-2) \\ & & \vdots \\ & & \gamma(0) \end{pmatrix}^{-1}\right)$$

ex) AR(1)

$$\hat{\phi} \sim N\left(\phi, \frac{1}{n}\sigma^2 \gamma(0)^{-1}\right)$$

(2) MA models

- Consider the MA(1) model $X_t = \theta X_{t-1} + Z_t$ where $|\theta| < 1$.
- Yule-walker equations are

$$\begin{aligned} \gamma(0) &= \sigma^2\left(1+\theta^2\right) \\ \gamma(1) &= \sigma^2\theta \end{aligned}$$

- The method of moments is an effective procedure for fitting auto-regressive models, it does not perform as well for ARMA models with $q > 0$. From a computational point of view, it requires as much computing time as the more efficient estimators based on either the innovations algorithm or the Hannan-Rissannen procedure and is therefore rarely used expect when $q = 0$.

- The Hannan-Rissanen Algorithm: MA 모형.
  - Assume ARMA(p,q) model
  - X_t is regressed both on $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$ and $Z_{t-1}, Z_{t-2}, \ldots, Z_{t-q}$.
  - Since $\{Z_t\}$ are unobserved, we replace them with estimated residuals from

$$\widehat{Z}_t = X_t - \widehat{\phi}_{m1}X_{t-1} - \widehat{\phi}_{m2}X_{t-2} - \cdots - \widehat{\phi}_{mm}X_{t-m}$$

    for m large enough.
  - One the estimated residuals $\{\widehat{Z}_t\}$ are computed, the vector of parameters $(\beta', \phi')'$ is estimated by least squares linear regression of X_t onto $(X_{t-1}, \ldots, X_{t-p}, \widehat{Z}_{t-1}, \ldots, \widehat{Z}_{t-q})$, $t = m+1+q, \ldots, n$ by minimizing the sum of squares

$$\mathbf{S}(\beta) = \sum_{t=m+1+q}^{n} \left(X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} - \theta_1 \widehat{Z}_{t-1} - \cdots - \theta_q \widehat{Z}_{t-q}\right)^2$$

    시거 ~ MA 차수

    with respect to β.
  - Finally

$$\widehat{\sigma^2} = \frac{\mathbf{S}(\widehat{\beta})}{n-m-q}$$

$\rightarrow \hat{\gamma}(0) = \phi_1\hat{\gamma}(1) + \cdots \phi_p\hat{\gamma}(p) + \hat{\sigma}_w$ 가능

$X_t = \underbrace{\phi_1 X_{t-1} + \cdots + \phi_p X_{t-p}}_{\text{관측가능}} + \underbrace{\theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} + Z_t}_{\text{관측 불가}}$

$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + Z_t \quad \sum_{j=1}^{\infty} |\pi_j| < \infty$; j가 충분히 커지면 0에 아주 가까움.

$X_t \approx \phi_{m1}X_{t-1} + \cdots + \phi_{mn}X_{t-n} + Z_t$ for sufficiently large m.

$\rightarrow \hat{\phi}_{m1}, \ldots, \hat{\phi}_{mn}$ 추정가능. (Yule walker or regression)

$\rightarrow \hat{Z}_t \approx X_t - \hat{\phi}_{m1}X_{t-1} - \cdots - \hat{\phi}_{mn}X_{t-m}$ $\rightarrow$ $S(\beta)$ 적용

이건 Yule walker 과정과 동일!!
Yule walker로 ϕ 추정한후
과정만 추가된 것!!

ex) MA(1)

$X_t = Z_t + \theta Z_{t-1}$

$Z_t = X_t - \theta Z_{t-1}$

$S(\beta) = \sum_{t=m+2}^{m} (X_t - \theta Z_{t-1})^2$ $\rightarrow$ θ에 대한 편미분 (구하는 걸 미분)

$\hat{\sigma}_2 = \frac{-2Z_{t-1}(X_t - \theta Z_{t-1})}{n-m-1}$

휘갈 주겠지

Model Selection by Information Criteria

$$\text{Information Criteria } = -2\log-\text{likelihood value } + \text{penalty}$$

Let k = number of parameters 변수개수

(a) Akaike's Information Criteria (AIC)

모형큰거

$$-2\log-\text{likelihood value } + \left\{2k\right\}$$

(b) Bias corrected AIC

$$-2\log-\text{likelihood value } + 2\left\{\frac{n\left(n+k\right)}{n-k-2}\right\}$$

(c) Bayesian Information Criteria (BIC)

작은거

$$-2\log-\text{likelihood value } + \left\{k\log n\right\}$$ n: sample size