MA637430 Analisis Deret Waktu **Universitas Udayana**
Semester Ganjil 2018/2019 **Program Studi Matematika**

Kuliah 7: Estimasi Parameter Model ARIMA

Koordinator Tim: I Wayan Sumarjaya (sumarjaya@unud.ac.id)
Anggota Tim Teaching I: I Gusti Ayu Made Srinadi (srinadi@unud.ac.id)
Anggota Tim Teaching II: Made Susilawati (mdsusilawati@unud.ac.id)

Capaian Pembelajaran Mata Kuliah
Mampu memilih model deret waktu yang sesuai (S5, S9, KU1, KU2, KU9, KK1, KK2, PP1)

Kemampuan Akhir yang Diharapkan
Mahasiswa mampu memilih model deret waktu ARIMA yang sesuai dengan melakukan spesifikasi dan estimasi model berdasarkan kriteria informasi dan mendemonstrasikan dengan perangkat lunak R (C5, P2, A2)

Indikator

1. Ketepatan memilih metode kemungkinan maksimum (maximum likelihood estimator)
2. Ketepatan menggunakan R untuk mengestimasi parameter menggunakan metode kemungkinan maksimum

Bahan Kajian/Materi Ajar

1. Ketepatan memilih metode kemungkinan maksimum (maximum likelihood estimator)
2. Ketepatan menggunakan R untuk mengestimasi parameter menggunakan metode kemungkinan maksimum

7.0.1 Kriteria Informasi

Kriteria informasi yang lazim digunakan dalam memilih model ARIMA adalah Akaike Information Criterion (AIC) dan Bayesian Information Criterion (BIC). AIC didefinisikan sebagai

$$\mathrm{AIC} = -2\log(\text{maximum likelihood}) + 2k \tag{7.1}$$

dengan $k = p + q + 1$ jika model berisi intersep atau suku konstan dan $k = p + q$ jika tidak. Ukuran lain adalah BIC yang didefinisikan oleh

$$\mathrm{BIC} = -2\log(\text{maximum likelihood}) + k\log n. \tag{7.2}$$

Model dengan nilai AIC dan BIC terkecil adalah kandidat model yang akan terpilih untuk digunakan pada tahap berikutnya seperti pemeriksaan diagnostik dan peramalan.

Contoh 7.0.1. Lihat kembali data Lake Huron. Kita telah melakukan *differencing* terhadap tren. Pada bagian ini kita akan melihat nilai AIC dan BIC dari Lake Huron. Ingat, kita akan menerapkan model ARIMA pada data asli.

```
> LakeHuron.ar2 <- arima(LakeHuron,order=c(2,1,0))
> LakeHuron.ar2

Call:
arima(x = LakeHuron, order = c(2, 1, 0))

Coefficients:
ar1      ar2
0.1728  -0.2233
s.e.  0.1012   0.1015

sigma^2 estimated as 0.5188:  log likelihood = -105.87,  aic = 215.74
> LakeHuron.ma1 <- arima(LakeHuron,order=c(0,1,1))
> LakeHuron.ma1

Call:
arima(x = LakeHuron, order = c(0, 1, 1))

Coefficients:
ma1
0.2003
s.e.  0.1145

sigma^2 estimated as 0.5398:  log likelihood = -107.75,  aic = 217.5
```

Dari luaran di atas diperoleh AIC untuk model AR(2) $=$ 215,74 dan AIC untuk model MA (1) $=$ 217,5. Jadi berdasarkan kriteria AIC kita akan memilih model AR(2).

7.0.2 Uji Akar Unit

Kita telah mempelajari bagaimana menggunakan *differencing* untuk mendapatkan deret waktu yang stasioner. Kita juga tahu bahwa *differencing* terhadap deret stasioner juga akan menghasilkan deret waktu stasioner. Namun, *differencing* yang berlebihan (*overdifferencing*) akan mengakibatkan korelasi yang tidak perlu pada deret dan akan memperumit proses pemodelan. Selain itu, akibat dari *overdifferencing* adalah membuat model yang tidak *invertible*. Model yang tidak *invertible* menyebabkan masalah serius pada saat kita mengestimasi parameter model.

Sebagai contoh misalkan $\{X_t\}$ adalah langkah acak (*random walk*). Melakukan *differencing* dengan bentuk

$$\nabla X_t = X_t - X_{t-1} = \varepsilon_t. \tag{7.3}$$

Namun, jika kita *differencing* lagi, dengan kata lain melakukan *overdifferencing*, kita akan mendapatkan

$$\nabla^2 X_t = \varepsilon_t - \varepsilon_{t-1} \tag{7.4}$$

yang merupakan proses MA(1), tetapi dengan $\theta = 1$. Jika kita ambil dua kali *differencing* kita akan mengestimasi nilai parameter θ yang tidak perlu. Menspesifikasikan model MA$(2, 1)$ tidaklah tepat dalam hal ini. Langkah acak (*random walk*) yang dianggap sebagai IMA$(1, 1)$ dengan $\theta = 0$ adalah model yang tepat.

Untuk menghindari *overdifferencing* sangat disarankan untuk melihat dengan hati-hati setiap beda (*difference*) dan mengingat prinsip irit (*parsimony*).

Kita juga telah tahu bahwa meluruhnya autokorelasi sampel secara linear sering merupakan indikasi bahwa deret waktu tidak stasioner dan memerlukan *differencing*. Adalah hal yang berguna jika kita juga mengetahui bukti ketidakstasioneran pada mekanisme pembangkitan data. Hal ini dapat dilakukan melalui uji hipotesis. Misalkan model

$$X_t = \alpha X_{t-1} + Y_t, \quad \text{untuk}, \quad t = 1, 2, \ldots \tag{7.5}$$

dengan $\{Y_t\}$ adalah proses stasioner. Proses $\{X_t\}$ tidak akan stasioner jika koefisien $\alpha = 1$, tetapi akan stasioner jika $|\alpha| < 1$. Misalkan $\{Y_t\}$ adalah proses AR(k) yaitu

$$Y_t = \phi_1 Y_{t-1} + \cdots + \phi_k X_{t-k} + \varepsilon_t. \tag{7.6}$$

Di bawah hipotesis nol bahwa $\alpha = 1$, $Y_t = X_t - X_{t-1}$. Misalkan $\beta = \alpha - 1$, kita akan peroleh

$$\begin{aligned} X_t - X_{t-1} &= (\alpha - 1)X_{t-1} + Y_t & (7.7)\\ &= \beta X_{t-1} + \phi_1 Y_{t-1} + \cdots + \phi_k X_{t-k} + \varepsilon_t & (7.8)\\ &= \beta X_{t-1} + \phi_1 (X_{t-1} - X_{t-2}) + \cdots + \phi_k (X_{t-k} - X_{t-k-1}) + \varepsilon_t & (7.9) \end{aligned}$$

dengan $\beta = 0$ di bawah hipotesis bahwa X_t adalah beda takstasioner (*difference nonstationary*). Dilain pihak jika $\{X_t\}$ adalah stasioner maka $-1 < \alpha < 1$, maka dapat dicek bahwa X_t masih memenuhi persamaan di atas namun dengan koefisien yang berbeda. Sebagai contoh,

$$\beta = (1 - \phi_1 - \cdots - \phi_k)(1 - \alpha) < 0. \tag{7.10}$$

Sebenarnya $\{X_t\}$ adalah proses AR$(k + 1)$ dengan persamaan karakteristik yang diberikan oleh

$$\Phi(x)(1 - \alpha x) = 0 \tag{7.11}$$

dengan $\Phi(x) = 1 - \phi_1 x - \cdots - \phi_k x^k$. Jadi dengan demikian hipotesis nol berhubungan dengan kasus bahwa polinom karakteristik AR memiliki akar unit dan hipotesis alternatif yang menyatakan bahwa tidak memiliki akar unit. Dengan demikian, pengujian terhadap *differencing* pada dasarnya menguji akar unit pada polinom karakteristik AR dari deret $\{X_t\}$.

7.1 Estimasi Parameter

Secara umum ada beberapa metode untuk menduga parameter model ARMA(p, q) seperti metode momen, metode kuadrat terkecil, metode kemungkinan maksimum, dan kuadrat terkecil tak bersyarat. Pada bagian ini kita hanya akan membahas metode kemungkinan maksimum (*maximum likelihood*).

Untuk sebarang amatan $X_1, X_2, \ldots, X_n$, deret waktu atau bukan, fungsi kemungkinan (*likelihood*) L didefinisikan sebagai densitas peluang bersama mendapatkan data yang diamati. Fungsi kemungkinan ini dianggap fungsi dari parameter yang tidak diketahui dengan amatan data dianggap tetap (*fixed*). Untuk model ARIMA fungsi kemungkinan L adalah fungsi dari ϕ, θ, μ, dan σ_ε^2 diketahui $X_1, X_2, \ldots, X_n$.

Sebagai contoh perhatikan model AR(1) dengan bentuk

$$X_t = \phi X_{t-1} + \varepsilon_t \tag{7.12}$$

diasumsikan ε_t saling bebas dan berdistribusi normal dengan nilai tengah nol dan simpangan baku σ_ε. Fungsi densitas peluang masing-masing ε_t adalah

$$(2\pi\sigma_\varepsilon^2)^{-1/2} \exp(-\varepsilon_t^2/(2\sigma_\varepsilon^2)), \quad \text{untuk } -\infty < \varepsilon_t < \infty \tag{7.13}$$

dan mengingat saling bebas, fungsi densitas peluang bersama $\varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n$ adalah

$$(2\pi\sigma_\varepsilon^2)^{(n-1)/2} \exp\left(-\frac{1}{2\sigma_\varepsilon^2} \sum_{t=2}^{n} \varepsilon_t^2\right). \tag{7.14}$$

Sekarang misalkan

$$\begin{aligned}
X_2 - \mu &= \phi(X_1 - \mu) + \varepsilon_2 \\
X_3 - \mu &= \phi(X_2 - \mu) + \varepsilon_3 \\
&\vdots \\
X_n - \mu &= \phi(X_{n-1} - \mu) + \varepsilon_n.
\end{aligned} \tag{7.15}$$

Selanjutnya diperoleh

$$f(x_2, x_3, \ldots, x_n|x_1) = (2\pi\sigma_\varepsilon^2)^{-(n-1)/2} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\sum_{t=2}^{n}[(x_t - \mu) - \phi(x_{t-1} - \mu)]^2\right). \tag{7.16}$$

Fungsi kemungkinan yang bersesuaian adalah

$$L(\phi, \mu, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-n/2}(1 - \phi^2)^{1/2} \exp[-S(\phi, \mu)/2\sigma_\varepsilon^2] \tag{7.17}$$

dengan

$$S(\phi, \mu) = \sum_{t=2}^{n}[(X_t - \mu) - \phi(X_{t-1} - \mu)]^2 + (1 - \phi^2)(X_1 - \mu). \tag{7.18}$$

Fungsi $S(\phi, \mu)$ disebut fungsi jumlah kuadrat tak bersyarat (*unconditional sum-of-squares function*). Fungsi log kemungkinan (*log-likelihood*) dinotasikan

$$\ell(\phi, \mu, \sigma_\varepsilon^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma_\varepsilon^2) + \frac{1}{2}\log(1 - \phi^2) - \frac{1}{2\sigma_\varepsilon^2}S(\phi, \mu). \tag{7.19}$$

Menurunkan persamaan log kemungkinan terhadap masing-masing parameter ϕ, μ, dan σ_ε^2 akan diperoleh nilai dugaan masing-masing parameter. Salah satunya adalah

$$\hat{\sigma}_\varepsilon^2 = \frac{S(\hat{\phi}, \hat{\mu})}{n}. \tag{7.20}$$

Untuk nilai ϕ dan μ, $\ell(\phi, \mu, \sigma_\varepsilon^2)$ dapat dimaksimalkan secara analitik terhadap σ_ε^2 dan kita peroleh

$$\hat{\sigma_\varepsilon^2} = \frac{S(\hat{\phi}), \hat{\mu}}{n}. \tag{7.21}$$

7.1.0.1 Sifat-sifat sampel besar

Untuk n besar, penduga akan mendekati normal dengan varians dan korelasi sebagai berikut:

AR(1) memiliki varians:

$$\text{var}(\hat{\phi}) \approx \frac{1-\phi^2}{n} \tag{7.22}$$

AR(2) memiliki varians dan korelasi:

$$\text{var}(\hat{\phi}_1) \approx \text{var}(\hat{\phi}_2) \approx \frac{1-\phi_2^2}{n} \tag{7.23}$$

dan

$$\text{cor}(\hat{\phi}_1, \hat{\phi}_2) \approx \frac{\phi_1}{1-\phi_2} = -\rho_1. \tag{7.24}$$

MA(1) memiliki varians:

$$\text{var}(\hat{\theta}) \approx \frac{1-\theta^2}{n}. \tag{7.25}$$

MA(2) memiliki varians dan korelasi

$$\text{var}(\hat{\theta}_1) \approx \text{var}(\hat{\theta}_2) \approx \frac{1-\theta_2^2}{n} \tag{7.26}$$

dan

$$\text{cor}(\hat{\theta}_1, \hat{\theta}_2) \approx -\frac{\theta_1}{1-\theta_2}. \tag{7.27}$$

ARMA(1, 1) memiliki varians dan korelasi:

$$\text{var}(\hat{\phi}) \approx \left(\frac{1-\phi^2}{n}\right)\left(\frac{1-\phi\theta}{\phi-\theta}\right)^2 \tag{7.28}$$

$$\text{var}(\hat{\theta}) \approx \left(\frac{1-\theta^2}{n}\right)\left(\frac{1-\phi\theta}{\phi-\theta}\right)^2 \tag{7.29}$$

$$\text{cor}(\hat{\phi}, \hat{\theta}) \approx \frac{\sqrt{(1-\phi^2)(1-\theta^2)}}{1-\phi\theta}. \tag{7.30}$$

Contoh 7.1.1. Lihat kembali data Lake Huron di atas. Kita peroleh estimasi untuk model AR(2) dengan estimasi parameter $\hat{\phi}_1 = 0{,}1728$ dan $\hat{\phi}_2 = -0{,}2233$.

7.2 Pengayaan

Buku-buku seperti Cryer and Chan (2008), Shumway and Stoffer (2011), Brockwell and Davis (2016), dan Box et al. (2016) dapat digunakan untuk pengayaan lebih lanjut.

Daftar Pustaka

George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Inc.,, Hoboken, New Jersey, fifth edition, 2016.

Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting*. Springer, New York, third edition, 2016.

Jonathan D Cryer and Kung-Sik Chan. *Time Series Analysis with Applications in R*. Springer, New York, second edition, 2008.

Robert H. Shumway and David S. Stoffer. *Time Series Analysis and Its Applications with R Examples*. Springer, New York, 2011.