Easy 4.0 STA 457 Time Series Analysis
Midterm Review – Part 2
By Sabrina Wang

Disclaimer

This complementary study package is provided by Easy 4.0 Education Inc. and its affiliated mentors. This study package seeks to support your study process and should be used as a complement, **NOT** substitute to course material, lecture notes, problem sets, past tests and other available resources.

We acknowledge that this package contains some materials provided by professors and staff of the University of Toronto, and the sources of these materials are cited in details wherever they appear.

This package is distributed for free to students participating in Easy 4.0's review seminars, and are not for sale or other commercial uses whatsoever. We kindly ask you to refrain from copying or selling in part or in whole any information provided in this package.

Thank you for choosing Easy 4.0. We sincerely wish you the best of luck in all of your exams.

Easy 4.0 Education Inc.

知识点

1. Model adequacy
2. model selection
3. Classical Decomposition
4. ARIMA Model
5. I(d) Process and Dickey-Fuller test
6. Forecasting
7. Transfer Function Noise Model
8. Box Tiao Transformation

1. model adequacy

(1) residual autocorrelation functions at lag k

Formula:

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} \hat{a}_t \hat{a}_{t+k}}{\sum_{t=1}^{n} \hat{a}_t^2}$$

$\{\hat{a}_t\}$ are residuals of a fitted ARMA(p,q) model which is given by $\hat{a}_t = X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p} - \hat{\theta}_1 a_{t-1} - \cdots - \hat{\theta}_q a_{t-q}$, hat means estimated value. Coefficients are estimated in the step 2 in Box-Jenkins approach. Residuals are approximately uncorrelated and can be regarded as white noise process because of the estimation process.

(2) portmanteau tests

Definition:
Portmanteau tests are overall tests that check an **entire group** of residual autocorrelation functions (assuming that the model is adequate)
Method:
n is the number of observations, m is number of lag. p and q are orders of ARMA model. Both methods follow Chi-squared distribution

1. Box and Pierce

$$Q_{BP} = n \cdot \sum_{1}^{m} \hat{\rho}_k^2 \sim \chi^2_{m-(p+q)}$$

2. Ljung and Box

$$Q_{LB} = \sum_{1}^{m} \frac{n \cdot (n+2)}{(n-k)} \hat{\rho}_k^2 \sim \chi^2_{m-(p+q)}$$

3. Remark
   H_0: the fitted model is adequacy
   H_1: the fitted model is NOT adequacy
   Reject H_0 if test statistics Q > $\mathcal{X}^2_{1-\alpha,\ m-(p+q)}$ $where\ a\ is\ significant\ level$

<u>Midterm practise question</u>
ARMA Model – Q5

Define two portmanteau tests taught in class for checking model adequacy.
1. define what is portmanteau test
2. list two formulas: Box and Pierce, Ljung and Box

2. model selection

<u>Midterm practise question</u>
ARMA Model – Q6
1. list two formulas: AIC and BIC
2. explain what is ML and k that appear in the formula
3. BIC tends to select a more parsimonious model

3. Classical Decomposition

Recall the definition of Classical Decomposition, we need to remove trend and seasonality component. We have several ways to do this. But in this package we only focus method of differentiation. Please do the method of regression by yourself.

3.1 Remove trend

Use Backshift operator B
$Bt = t - 1, Bc = c\ (c\ is\ constant), Bx_t = x_{t-1}$
$\nabla^d = (1 - B)^d$

<u>Example</u>

Let $Y_t = a + bt + ct^2 + X_t$ where X_t is a stationary time series.
Whether the process $(1-B)^2 X_t$ is stationary?

$$Y_t = a + bt + ct^2 + X_t$$

$$(1-B)(1-B) = 1 - 2B + B^2$$

- $(1-B)^2 a = (1-2B+B^2)a = a - 2a + a = 0$
- $(1-B)^2 bt = \nabla(1-B)bt = \nabla(bt - b(t-1))$
  $= \nabla(bt - bt + b) = (1-B)b = b - b = 0$
- $(1-B)^2 ct^2 = \nabla(1-B)ct^2$
  $= \nabla(ct^2 - c(t-1)^2)$
  $= \nabla(ct^2 - ct^2 + 2ct - c) = \nabla(2ct - c)$
  $= \nabla(2ct) - \cancel{\nabla(c)} \to 0 = 2ct - 2c(t-1)$
  $= 2c.$

$\Rightarrow (1-B)^2 Y_t = 2c + (1-2B+B^2)X_t$
$= 2c + X_t - 2X_{t-1} + X_{t-2}$

Yes. given $\{X_t\}$ is stationary.

Midterm practice
Box-Jenkins approach and Unit root tests – Q1

Suppose that $x_t = \alpha + \beta t^3 + y_t$ and $y_t = a_t + \theta_1 a_{t-1}$, $a_t \sim NID(0,1)$. Show that we can make x_t a stationary time series by differencing $\{x_t\}$ three times.

3.2 Remove seasonal component

To deal with seasonality of period d by introducing the lag-d difference operator ∇_d

$\nabla_d X_t = X_t - X_{t-d} = (1 - B^d) X_t$

Applying the ∇_d operator to the classical decomposition model, $x_t = m_t + s_t + Y_t$ where s_t has period d and we have

$$\nabla_d X_t = m_t - m_{t-d} + Y_t - Y_{t-d}$$

Remark:

$m_t - m_{t-d}$ is trend component, $Y_t - Y_{t-d}$ is noise component

m_t is the trend component, s_t is the seasonal component

4. ARIMA Model

Definition:

$\{X_t\}$ is said to follow an ARIMA model of order (p,d,q) if $W_t = (1 - B)^d X_t$ is a stationary ARMA model. We have $(1 - B)^d \Phi(\mathrm{B})(X_t) = \Theta(\mathrm{B}) a_t \quad a_t \sim (0, \sigma^2)$

5. I(d) process and Dickey-Fuller test

5.1 I(d) process

Definition: a series follows a stationary ARMA model after differencing d times is said to be integrated of order d or I(d) process.

5.2 One Step DF Unit Root test

Definition:

The Dickey-Fuller test is used to test I(1) processes. Consider

$H_0 : \pi = 0 \; or \; X_t \sim I(1); \; H_1 : \pi > 0 \; or \; X_t \sim I(0)$

$X_t = \phi X_{t-1} + a_t, \quad a_t \sim NID(0, \sigma^2).$

$\Delta X_t = (\phi - 1) X_{t-1} + a_t = \pi X_{t-1} + a_t$

Remark:

$\Delta X_t = X_t - X_{t-1} = (\phi - 1) X_{t-1} + a_t$

$let \; \pi = \phi - 1$

$if \; H_0 \; is \; true \; \leftrightarrow \pi = 0 \leftrightarrow \phi = 1$

5.3 general Dickey Fuller test

Definition:

$\Delta X_t = a + \tau^T DR_t + \pi X_{t-1} + a_t,$

a is regression intercept ; DR_t deterministic independent variables; τ is the corresponding coefficient vector; $a_t \sim NID(0,1)$

Use:

The use of the Dickey-Fuller test contains two steps:
(1) remove the deterministic time trend and
(2) conduct statistical inference using the prescribed test.

Limitation:

1. considers only a single unit root
2. This two-step procedure introduce the error in variable problem for testing the presence of a unit root
3. Even using the one step procedure, the test does not consider the presence of an autocorrelated error process

Test two unit root

Please read slides P20-21 on 'ARIMA and unit root test'

5.4 Augmented Dickey Fuller (ADF) test

Definition:

$$\Delta X_t = \tau^T DR_t + \pi X_{t-1} + \sum_{j=1}^{k} \gamma_j \cdot \Delta X_{t-j} + a_t,$$

Note: k = q-1. ADF test uses the lagged terms ΔX_{t-j} to correct the impact due to the presence of the autocorrelated error terms in the one-step Dickey Fuller test.

Get length of lag:

Please read slides P23-25 on 'ARIMA and unit root test'

Midterm practice

Box – Jenkins approach and Unit root tests – Q2,3,4,5

知识点：I(d) process; unit root test; general DF test and limitation; augmented DF test; test multiple roots.

6. Forecasting

Comprehension:
The objective of forecasting is to forecast a value X_{t+l} , $l \geq 1$ when we are currently standing at time t. This forecast is said to be made at origin t for lead time l.

6.1 Forms of ARIMA Model

1) Difference equation form:

$$X_{t+l} = \varphi_1 X_{t+l-1} + \cdots + \varphi_{p+d} X_{t+l-p-d}$$
$$+a_{t+l} + \theta_1 a_{t+l-1} + \cdots + \theta_q a_{t+1-q}$$

2) **Integrated form**

$$X_{t+l} = \sum_0^\infty \psi_i a_{t+l-j} \ = a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1} + C_t(l)$$

$$C_t(l) = \sum_l^\infty \psi_j a_{t+l-j}$$

This form show a weighted sum of current and previous error terms.

3) Weighted average of previous observations

$$X_{t+l} = a_{t+l} + \sum_{j=1}^{\infty} \pi_j X_{t+l-j}$$

$$= a_{t+l} + \pi_1 X_{t+l-1} + \cdots + \pi_{l-1} X_{t+1} + \sum_{j=l}^{\infty} \pi_j X_{t+l-j}$$

6.2 Forecast

define the forecasted X_{t+l} as
$\hat{X}_t(l) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \cdots$

By minimize the mean square error of the forecast, we chose $\psi^*_{l+j} = \psi_{l+j}$.
Now it is easy to show that

$$X_{t+l} = a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1} + (\psi_l a_t + \psi_{l+1} a_{t-1} + \cdots)$$
$$= e_t(l) + \hat{X}_t(l),$$

Definitions

1)

$\hat{X}_t(l)$: minimum mean square error forecast at t, for lead time l.

$\hat{X}_t(l) = E_t(X_{t+l}) = E(X_{t+l} \mid X_t, X_{t-1} \cdots)$ i.e. given all info up to time t

conditional expectation of X_{t+l} @ time t.

$= \psi_l a_t + \psi_{l+1} a_{t-1} + \psi_{l+2} a_{t-2} + \cdots$.

$\{a_t\}$: random shocks.

need $E_t(a_{t+j}) = E(a_{t+j} \mid X_t, X_{t-1}, \dots) = 0$. $\forall j > 0$.

2)

2) $e_t(l)$ is the forecast error for lead time l.

$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1}$

- $E_t(e_t(l)) = E(e_t(l) \mid X_t, X_{t-1}, \dots) = 0$.
  i.e. unbiased forecast
- $Var(e_t(l)) = Var(a_{t+l} + \psi_1 a_{t+l-1} + \cdots + \psi_{l-1} a_{t+1})$

$a_t \sim NID(0, \sigma^2)$. $= (1 + \psi_1^2 + \cdots + \psi_{l-1}^2)\sigma^2$.

(Note: linear function of $\sum_1^L w_l \hat{X}_t(l)$ is min. mean square error forecast of function $\sum_1^L w_l X_{t+l}$)

i.e. the MMSE property maintains in linear function.

6.3 Rules for Calculating the Conditional expectation

1.3 Rules for Calculating the Conditional expectation

方框表示 conditional expectation @ time t

① $[X_{t-j}] = E_t[X_{t-j}] = X_{t-j}, \quad j = 0,1,2,\ldots$ recall $E_t(X_{t-j}) = E(X_{t-j} \mid X_t, X_{t-1}, \ldots)$

② $[X_{t+j}] = E_t[X_{t+j}] = \hat{X}_t(j), \quad j = 1,2,\ldots$

③ $[a_{t-j}] = E_t[a_{t-j}] = a_{t-j} = X_{t-j} - \hat{X}_{t-j-1}(1), \quad j = 0,1,2,\ldots$ $X_{t-j} - E_{t-j-1}(X_{t-j})$

④ $[a_{t+j}] = E_t[a_{t+j}] = 0, \quad j = 1,2,\ldots$

e.g. $j = 0$. $a_t = X_t - \hat{X}_{t-1}(1)$.

理解. actual shock. what $t-1$ expect

The expectation of past observation given condition on future observation = value of past observation.

7. Transfer Function Noise Model

7.1 Basic Definition:

1) A TFN model is a regression equation is used to predict current values of a dependent variable based on both the current values as well as the lagged (past period) values.

2) The mathematical form of this model

$$y_t = a + v_0 x_t + v_1 x_{t-1} + v_2 x_{t-2} + \cdots + \varepsilon_t$$
$$= a + \sum_i^{inf} v_i x_{t-i} + \varepsilon_t$$

3) the coefficients $v_0, v_{1,\ldots}$ are impulse response function of the system. The impulse responses must be absolutely summable which means $\sum_{j=0}^{inf} |v_j| < \infty$. The sum of coefficients is called stead-state gain.

7.2 Model

rotational distributed lag model

$w(B) = w_0 + w_1 B + w_2 B^2 + \cdots$ can be approximated by $\frac{\delta(B)}{\tau(B)}$

$\delta(B) = \delta_0 + \delta_1 B + \delta_2 B^2 + \cdots + \delta_r B^r$; $\tau(B) = 1 - \tau_1 B - \cdots - \tau_s B^s$

$V(B) = \sum_{i=0}^{\infty} V_i B^i$

Given $y_t = \sum_{i=0}^{\infty} V_i X_{t-i} + \varepsilon_t = V(B) X_t + \varepsilon_t$

$\Rightarrow$ rewrite $y_t = \frac{\delta(B)}{\tau(B)} X_t + \varepsilon_t$

$\varepsilon_t \sim$ ARMA model. $\Phi(B)\varepsilon_t = \Theta(B) e_t$

7.3 Preliminary Identification & prewhitening

Given $y_t = \sum_{i=0}^{\infty} V^i X_{t-i} + \varepsilon_t = V(B) X_t + \varepsilon_t$ ①.

ε_t follows a stationary & invertiable ARMA (p, q) model

Suppose X_t follows a ARMA model

$\phi_x(B) X_t = \theta_x(B) a_t$. $a_t \sim NID(0, \sigma^2)$.

$\Rightarrow a_t = \frac{\phi_x(B)}{\theta_x(B)} X_t$. mutiple ① by $\frac{\phi_x(B)}{\theta_x(B)}$.

$$\underbrace{\frac{\phi_x(B)}{\theta_x(B)} y_t}_{\tau_t} = V(B) \underbrace{\frac{\phi_x(B)}{\theta_x(B)} X_t}_{a_t} + \underbrace{\frac{\phi_x(B)}{\theta_x(B)} \varepsilon_t}_{n_t}.$$

$\Rightarrow \tau_t = V(B) a_t + n_t$ ②. $\{n_t\} \perp\!\!\!\perp \{a_t\}$.

mutiply ② by a_{t-j} and take expectation

$E(a_{t-j} \tau_t) = E(V(B) a_t a_{t-j}) + E(n_t a_{t-j})$.

Given $E(a_t) = 0 \quad \forall t$.

$\Rightarrow COV(a_{t-j}, \tau_t) = COV(V(B) a_t, a_{t-j}) + \cancel{COV(n_t, a_{t-j})}^{0}$.

$= V_j \, COV(a_{t-j}, a_{t-j})$

$= V_j \, Var(a_{t-j})$.

$$\Rightarrow v_j = \frac{cov(r_t, a_{t-j})}{Var(a_{t-j})} = \frac{corr(r_t, a_{t-j}) \cdot se(r_t)\, se(a_{t-j})}{se(a_{t-j})^2}$$

$$= Corr(r_t, a_{t-j}) \cdot \frac{se(r_t)}{se(a_{t-j})}$$

$$= \rho_{ra}(j) \cdot \frac{se(r_t)}{se(a_{t-j})}$$

So it want to test whether v_j is significant, we can test whether $\rho_{ra}(j)$ is significant or not.

7.4 Diagnostic Checking on TFN

1. Please read the **Cross-correlation check** and **Autocorrelation check**

$$Q_0 = m(m+2)\sum_{j=0}^{K}(m-j)^{-1}\,\hat{\rho}^2_{\hat{\alpha}\hat{a}}(j) \sim \chi^2_{K+1-M}$$

$$Q_1 = m(m+2)\sum_{j=1}^{K}(m-j)^{-1}\,\hat{\rho}^2_{\hat{a}}(j).$$

2. Understand the cross-correlation checks whether the noise $\{\varepsilon_t\}$and the input series $\{x_t\}$are uncorrelated. Given model: $y_t = V(B)X_t + \varepsilon_t$ and $\Phi_x(B)X_t = \theta_x(B)a_t$
3. Understand the autocorrelation checks whether $\{e_t\}$ has no serial correlation where $\Phi(B)\varepsilon_t = \theta(B)e_t$

8. Box and Tao Transformation for Estimation

Box & Tiao Transformation.

Assume $y_t = v(B) X_t + e_t$. $\quad e_t \sim$ ARMA $\Rightarrow a_t = \frac{\phi(B)}{\theta(B)} e_t$

$\phi(B) e_t = \theta(B) a_t$

Multiply $\frac{\phi(B)}{\theta(B)}$ $\quad \underbrace{\frac{\phi(B)}{\theta(B)} y_t}_{\tilde{y}_t} = v(B) \underbrace{\frac{\phi(B)}{\theta(B)} X_t}_{\tilde{X}_t} + \underbrace{\frac{\phi(B)}{\theta(B)} e_t}_{a_t}$

$$\tilde{y}_t = v(B) \tilde{X}_t + a_t = \sum_{j=0}^{s} v_j \tilde{X}_{t-j} + a_t$$

Step1: Run the Ordinary least squares regression on

$$y_t = \sum_{j=1}^{s} v_j x_{t-j} + e_t.$$

we can get the residual $\hat{e}_t$

Step2: fit ARMA(p,q) to the residuals getting from step 1

Step3: Apply Box and Tiao transformation using the identified model (in step 2) to filter $\{x_t\}$ and $\{x_t\}$ for all t

Step4: Run the OLS regression of equation in step 1 using the transformed variables obtained in Step 3

Step5: Check whether the regression residuals on Step 4 are serially uncorrelated.

If not, repeat Step 2 to 4; If yes, the model estimation is complete.

Midterm practice
Transfer function noise model – Q1,Q2

Consider an infinite distributed lag model

$$y_t = \sum_{i=0}^{K} v_i \, x_{t-i} + a_t \qquad a_t \sim NID(0,1)$$

$$x_t = \phi x_t + e_t \qquad e_t \sim NID(0,1)$$

1. Determine K in first equation using the prewhitening procedure
2. Suppose K is infinity and $v_i = v^i$

a. Approximate the first equation using the rational distributed lag model. Specially, let

$$y_t = \frac{\delta(B)}{\vartheta(B)} x_t + \xi_t$$

where $\delta(B) = \delta_0 + \delta_1 B + \cdots + \delta_r B^r, \vartheta(B) = \vartheta_0 + \vartheta_1 B + \cdots + \vartheta_s B^s$

Write $\delta(B), \vartheta(B), \xi_t$ in terms of v_i, a_t and backward shift operater B

b. Suppose that $\delta(B) = 1$ and $\vartheta(B) = 1 - \vartheta B$. How to estimate ϑ using the Box and Tiao transformation

c. Sate how to use portmanteau test to check model adequacy

Notes

Reference

Jen-wen Lin (2018). Lecture note [PowerPoint slides]. Retrieved from University of Toronto Portal site.

Jen-wen Lin (2018). Midterm Practice [PDF]. Retrieved from University of Toronto Portal site.