**ECON 504 Discussion Note**

TA information:

Name: Xu Wang, PhD student in Economics

Email: [xwang222@binghamton.edu](mailto:xwang222@binghamton.edu)

OH: M, noon – 1 pm, Wed, 6-7 pm

Discussion I: 10:50 – 11:50 AM

Discussion II: 12:00 – 1:00 PM

Three components:

1. Concepts review
2. Demonstrate codes from lecture notes
3. Talk about the homework (hint, demonstrate)

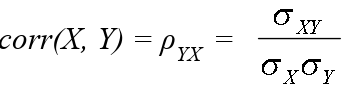
**Discussion session I (Feb 19, 2021)**

**Concepts review:**

1. Stochastic Process
2. Moments of probability distribution (univariate distribution)
   1. First moment: mean
   2. Second moment: variance
   3. Skewness
3. Moments of a joint probability distribution
   1. Covariance



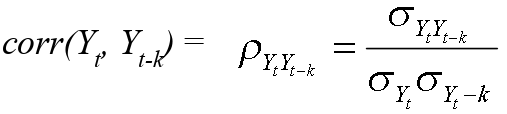
* 1. Correlation



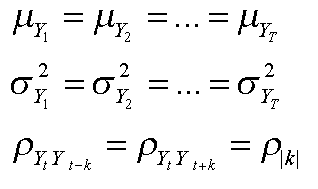
1. In the context of time series
   1. Autocovariance



* 1. Autocorrelation



1. Weakly/covariance-stationary vs Strict Stationary



1. Non-stationary process – first difference
2. Basics of R

# starts from a clean slate

# sets the default directory

# load library

# read in data

# use command from library, read documentation when needed

# save plots

# useful commands:

* Cbind() ## combine column
* par(mfrow=c(2,2)) ## set figure layout

Lab code：HW1.R

**Discussion session II**

**Concepts review:**

1. Concepts regarding forecasting
2. Minimize the expected loss (different types of loss functions)
3. If Y is i.i.d, then unconditional mean is the best forecast (Y\_bar: mean)
4. If Y is correlated to Y\_t-1 (or X), then conditional mean is optimal (Y\_t-1 and x should provide additional information)
5. Use knowledge of regression to generate optimal forecasts (conditional mean of Y)

(when we do regression models, we are implicitly model the conditional mean of Y as a function of X)

1. Difference between mean and conditional mean
2. P(A|B) = P(AB)/P(B)
3. If A and B are independent, then P(AB)=P(A)\*P(B)

Plug 2) into 1), we have:

P(A|B) = P(AB)/P(B)= P(A)\*P(B)/P(B)=P(A)

If P(A|B) = P(A), A and B are independent.

1. ACF vs PACF
2. White noise

Atime series process with zero mean and a constant finite variance whose autocorrelations and partial autocorrelations are all zero:

* Independent White Noise
* Gaussian/Normal White Noise

1. White noise test
2. Bartlett bands
3. LBQ test
4. MA(1) process

* Mean
* Variance
* Covariance
* Correlation
* ACF (for identification)
* PACF (for identification)
* Stationary
* Invertible:
  + y\_t has an autoregressive representation.
  + 🡪 can be written as a weighted sum of past Y + current shock
  + Model is identified (|theta|<1)

1. MA(2) process (see slides)
2. Calculation of autocorrelation

,

To calculation autocorrelation , you need: 1) variance ; 2) covariance

For variance,

We know that

Therefore,

We know that , therefore,

For covariance,

We know that

1. Forecasting (Lecture Slide 8)

* Point estimate
* Confidence interval

1. Hypothesis testing
2. How to state the hypothesis
3. What can we infer from the test?

H\_0: x=0

H\_1: x>1

P-value<0.05

Reject H\_0

P\_value =0.65 > 0.05

Can say: can’t reject H\_0

Can’t say: accept H\_0

**Discussion session 3 (Mar 12, 2021)**

1. Forecasting with MA(1) process:

,

Suppose

1. What do we want to know at period t?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Name** | **Notation** | **K=1** | **K=2** | **K=3** |
| True value of : |  | (1.1) | (1.2) | (1.3) |
| Unconditional mean of : |  | 8 | 8 | 8 |
| Unconditional variance of |  |  |  |  |
| Conditional mean of |  | 7.6 | 8 | 8 |
| Conditional variance of |  | 0.25 |  |  |
| Forecast interval of |  | (7.11, 8.09) |  |  |
| Point forecast: |  |  |  |  |
| Forecasting error: |  | (9.1) | (9.2) | (9.3) |
| Forecasting variance: |  | 0.25 |  |  |

1. Where do I get actual

Actual comes from the data generating process (estimated):

,

,

,

Why is an equation instead of a concrete number?

Because now you are in period t, you can’t observe . However, you do have all information up to period t (ie: , for period t, t-1, t-2, …)

1. Unconditional mean of

,

,

,

For (2.1), because we are finding unconditional mean, so we pretend that we don’t observe . Therefore, we have:

1. Unconditional variance of

,

1. Conditional mean of

When k=1:

When k=2:

When k=3:

1. Conditional variance of

When k=1:

Remember that:

Also remember that:

1. is white noise, so [Note: if is not a white noise process, this equation doesn’t hold.]
2. [This comes from the assumption on distribution]

When k=2:

When k=3:

1. Forecasting error:

the difference between actual and predicted (ie: conditional mean) of

1. Find the forecast error

When k=1

Remember that you are now in period t, so you can observe (ie: is a constan, instead of a random variable). Therefore, , where C is a constant.

What is ?

Remember , this means:

1. is white noise, so [Note: if is not a white noise process, this equation doesn’t hold.]
2. [This comes from the assumption on distribution]

When k=2,

When k=3,

1. Now I know that , what else do I need to know? Isn’t that enough?

Well, is itself a random variable. So, if you don’t tell me the distribution of it, you basically didn’t tell me anything useful. You need to at least tell me the conditional mean and conditional variance of it.

1. Ok. Conditional mean and variance are easy.

10.1 Definition of optimal forecast (conditional mean=0)

10.2 Conditional variance

When k=1

is white noise, so . , so:

When k=2

When k=3

1. Forecast interval of

Forecast interval for =

1. Derive 60th percentile of N(2,6)

Let the 60th percentile be K\*

K\*=?

1. What is the probability of K>4.5? K~ N(2,6)
2. Forecasting with MA(2) process:

,

|  |  |
| --- | --- |
|  | 1.4 |
|  | 0.8 |
|  | -1 |
|  | 0.6 |

1. Derive the point forecast and forecasting interval.

Point forecast 🡪conditional mean of

Forecasting interval 🡪 conditional mean of

Standard error of forecasting (SEF)🡪sqrt of conditional variance of forecasting error

Forecasting error🡪

1. Conditional mean of

h=1:

h=2:

h=3:

1. Forecasting error

h=1:

h=2:

h=3:

**Discussion session 4 (Mar 19, 2021)**

1. AR(1) process
2. Stationary =🡺🡺 and future shocks are uncorrelated [15:54 of lecture recording]
3. Mean
4. ACF

(1)-(2):

(3)\*

(3)\*

1. : squaring eq(3) and take expectation [57:27]
2. AR(2) process
3. Stationarity condition
4. Mean
5. ACF

(4)-(5)

(6)\*

E

(6)\*

E

(6)\*

E

**Discussion session 5 (Mar 26, 2021)**

1. Forecasting with AR(1)

h=1

h=2

1. Forecasting with AR(2)

h=1

h=2

h=3

1. ARMA(p,q)🡪ARMA(1,1)
2. Stationarity + invertibility
3. Mean of [10:49]
4. Point forecast (with and given)

h=1

h=2

1. Forecast error

h=1

h=2

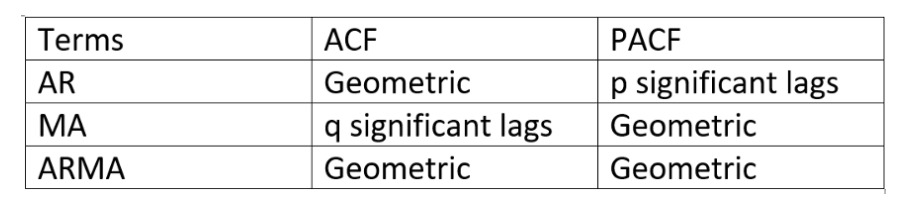
1. Forecasting with R (with given model + R results)

Coefficients:

|  |  |  |
| --- | --- | --- |
| Ar1 | Ar2 | Mean |
| 0.6 | -0.1 | 5 |

Signma^2 estimated as 25

1. What is your forecasting model?
2. What is your point forecast?
3. What is your forecasting interval?
4. R application
5. Model identification
   1. Plot the series and ensure stationarity
   2. Tentative identification based on ACF and PACF



Suppose you have 3 candidates:

1)ARMA(4,0);

2) ARMA(2:2);

3) ARMA(3,1)

* 1. Estimate the model
  2. Diagnostic
     1. Whether residuals are white noise?
* Bartlett band and Q stat. Good if can’t reject🡪 residual is WN 🡪 regression utilized all useful information.
* Degree of freedom
  + 1. AIC + BIC 🡪 smaller is better.
  1. Forecast with selected model

1. Application

**Discussion session 6 (Apr 9, 2021)**

1. Technic details regarding OLS
2. Evaluating forecast
   1. Optimality of forecast
      1. Zero mean under quadratic loss (Test1, MPE test)
      2. 1-step forecast errors are WN (Test2)
      3. H-step –Ma(h-1)
      4. Variance converge
      5. Optimality w/r/t information set (Test3)
   2. Accuracy of forecast
      1. RMSE – root mean squared error
      2. MAE – mean absolute error
      3. MAPE – mean absolute percentage error
3. Test1(MPE test)

Problem 1: How to obtain ?

1990-2021, monthly

Train: 1990-2015

Test: 2016-2021

Predicted: 2016-2021

Forecast error=test – predicted value

Answer: split data into training and test; using training data to train the model and use the actual value in test set to calculate

Problem 2: How to obtain multiple values of ?

Answer: 1) Fixed scheme; 2) rolling; 3) rolling origin

Fixed scheme: training data and test data are fixed; only train model once; use multiple points in test data to obtain multiple .[80:20]

Total: 100 data points

Train: first 80

Test: last 20

Rolling scheme: training data and test data are re-determined in each iteration; use only one point in test data to obtain .[80:20] [81:19] [82:18]…

Iteration1: 80: 20

Iteration2: 81: 19

Iteration3:82: 18

Rolling origin: training data and test data are re-determined in each iteration; use only one point in test data to obtain .[80:20] [80:19] [80:18]…

Iteration1: 80: 20

Iteration2: 80, started with index 2: 19

Iteration3:80, started with index 3: 18

|  |  |
| --- | --- |
| Index |  |
| 1 | 1.5 |
| 2 | 0.3 |
| 3 | 0.4 |
| … | … |

Problem 3: Suppose we have 20 observations of , why not simply calculating and check if it equals 0?

Answer: suppose , but we don’t know if it is statistically different from 0.

Problem 4: How to do the test?

Answer: do the following linear regression and check if the intercept is significantly different from 0

If we can’t reject the null: , then the forecast is good.

1. Test3 (information efficiency test, MZ test)

If , then your forecast error is not a function of past information, which is good.

In one word, all tests in the optimality check are to look for insignificance.

1. Compare one model with another model

Get the difference of squared error between two models, then regress the difference on 1. If the coefficient is not significant, this means the difference is not significant, meaning one model is not significantly different/better than the other.

1. R application

Step 1: import data and define as time series

g <- ts(data$G,start=c(1990,2), end=c(2021,2), freq=4)

Step 2: split into training and testing set

train <- window(g, start=c(1990,2), end=c(2020,2))

Step 3: Use training data to fit the model

fit <- Arima(train, order=c(1,0,0))

Step 4: Obtain

Forecast <- window(fitted(newfit), start=c(2020,3), end=c(2021,2))

Step 5: Obtain

Forecast error <- actual - Forecast

Step 6: Form the Table

cbind(actual, Forecast, Forecast Error)

|  |  |  |  |
| --- | --- | --- | --- |
| Actual | Forecast | Forecast Error | Constant |
| xxx | aaa | xxx-aaa | 1 |
| yyy | bbb | Yyy-bbb | 1 |
| … | … | … | .. |
|  |  |  |  |

Step 7: Test1(MPE test)

Regress Forecast Error on 1

Step 8: Test3 (IE test)

Regress Forecast Error on Forecast

**Discussion session 7 (Apr 16, 2021)**

1. Combined forecast

100 data point

Model A: high accuracy for the first 50 data points, low accuracy for last 50 data points

Model B: high accuracy for 51-100, low accuracy for 1- 50.

Ensemable model

Model A:

Model B:

* 1. Equally weighted (CF1): 0.5\*FA+0.5\*FB=0.5\*5+0.5\*5.5=
  2. Inverse MSE (CF2):
  3. Regression method (CF3):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Actual | Forecast\_A | Forecast Error\_A | Forecast\_B | Forecast Error\_B | CF\_1 | CF\_2 | CF3 | Constant |
| xxx | aaa | xxx-aaa | ccc | xxx-ccc | 0.5\*aaa+0.5\*ccc |  |  | 1 |
| yyy | bbb | Yyy-bbb | ddd | Yyy-ddd |  |  |  | 1 |
| … | … | … |  |  |  |  |  | .. |
|  |  |  |  |  |  |  |  |  |

1. Static model with time series data
   1. Residual analysis (see slides)
   2. Breusch-Godfrey test (H0: residual is WN) (see slides example)

Objective: you want to choose a model where the residual is WN🡪you correctly specified the structure of autocorrelation in the error term

* 1. Graphical test (plot residual against predictor/fitted value)
  2. Model selection --- predictor selection

1. Forecast with Static model (static in terms of the right-hand side variable)

One period forward

* 1. Type of forecast
     1. Unconditional forecast (unconditional on X)

One period forecast

Unconditional means needs to be forecasted. We don’t have the true value to condition on.

* + 1. Conditional forecast (Conditional on ) (scenario analysis)

One period forecast

Here, the forecast means: if , what is my Y?

* + 1. Lagged X

One period forward

One period forecast

1. Forecast with Dynamic model
   1. Model specification
      1. Distributed Lag Model
   2. Regression with ARMA errors (Special case of ADLM)
   3. Regression with lags in both Y and X (Autoregressive Distributed Lag Model)

**Discussion session 8 (Apr 23, 2021)**

1. Forecast with seasonality (deterministic)
   1. Model specification
   2. Deterministic: same for each year
2. Forecast with seasonality (Stochastic)
   1. Model specification
   2. Forecast
3. Combine seasonality with ARMA model
4. Nonstationary
   1. Trend stationary – the process become stationary around its time trend
   2. Trend models
      1. Linear Trend
      2. Log-linear Trend (adjustment)
      3. Piece-wise Linear
      4. Cubic Spline model
      5. Natural Spline model

**Discussion session 9 (Apr 30, 2021)**

1. Stochastic non-stationarity – Unit Root process

is called the AR lag polynomial of the process and suppose there are two roots and to . Suppose , then the process is not covariance stationary because covariance stationary requires

Because , we can rewrite the above equation as:

Now, is a stationary ARMA(1,1) process

1. Random work

Mean of the process:

Variance of the process:

Covariance of the process:

Correlation of the process:

Therefore, it is not stationary.

What to do? Take the difference. (is difference stationary).

1. Random work with a drift

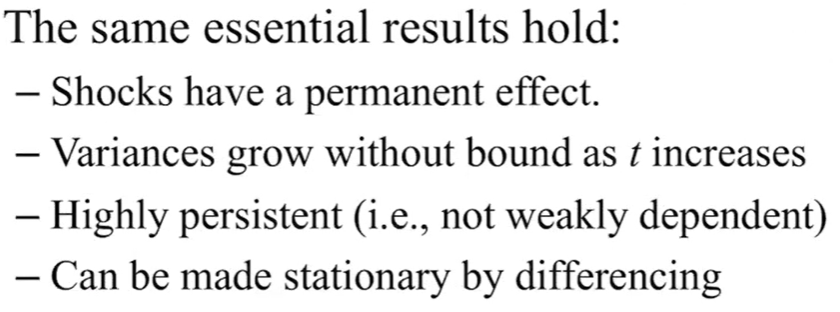
Mean of the process:

Variance of the process:

Covariance of the process:

Correlation of the process:

To summarize:



1. Seasonal Unity Root Process
2. Diagnostics for Unity Root Process
   1. Plot the series to check for trend
   2. Check ACF and PACF (slowly die-down of ACF, very high PACF)
   3. KPSS Test
      1. Regular unit root process

Step 1: check the graph to see if there is a linear trend. If yes, then specify “Tau”.

Tau: null-🡪Trend stationary

Blank: null 🡪level stationary

Step 2: check the t statistics. If very large, reject null. 🡪there is unit root🡪need to know how many times of difference you need to take.

Step 3: use ndiffs() to determine the number of differences you need to take.

Step 4: generate the differenced process, then go back to step 1. Stop until you can’t reject your null.

* + 1. Seasonal unit root process

Step 1: check the graph to see if there is a linear trend. If yes, then specify “Tau”.

Tau: null-🡪Trend stationary

Blank: null 🡪level stationary

Step 2: check the t statistics. If very large, reject null. 🡪there is unit root🡪need to know how many times of difference you need to take.

Step 3: use nsdiffs() to determine the number of differences you need to take.

Step 4: generate the differenced process, do the kpss test again, check if there is still unit root.

Step 5: then use ndiffs() to determine the number of differences you need to take.

Step 6: generate the differenced process, then go back to step 1. Stop until you can’t reject your null.

* 1. Goal is to make sure the resulting process is stationary on which our regression and forecast will be based.

1. Forecast with random work

AR process: the current information will become less and less useful as time goes.

RW process: the current information is always as important and will not diminish with the time.

Forecast error:

Variance of Forecast error:

1. Forecast with random work with drift

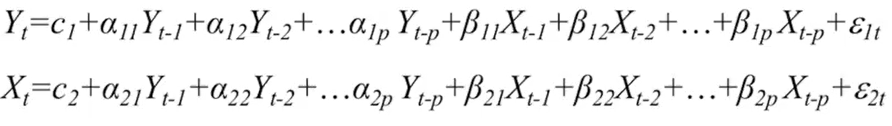
Forecast error:

Variance of Forecast error:

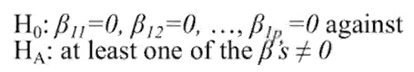
1. Cubic polynomial trend
2. Exponential trend

**Discussion session 10 (May 13, 2021)**

1. Vector Autoregression



* 1. Not coming from economic theory
  2. Treating all variable’s endogenous (impossible to learn the contemporary causality)
  3. Mainly reveals correlation between variables
  4. Can contain arbitrary number of variables
  5. Only contain lagged values
  6. Order of lags are same for all variables
  7. Decision of order of lag (information criteria, library(vars), VARselect)
  8. Forecasting with VAR (need to use another equation for period beyond t+1)
  9. Granger causality test (after controlling the lags of itself, does the history of other variables help to forecast it?)
     1. Lmtest, grangertest (F-test, for multiple lags/constraints)
     2. T-test (significance in the regression, 1-lag)



1. Cointegration
   1. Test of long-run eqm
   2. Individual I(1) but linear combination I(0)
   3. Engel-Granger Method
      1. H0: no cointegration
      2. Dickey-Fuller test on the residual (ur.df())
   4. Error Correction Model (link short-run adjustment with long-run eqm)