DS 6373 Midterm Summer 2019 Ver A Solution

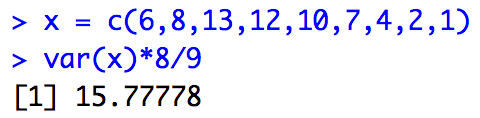
Given the data below answer the following questions:

|  |  |
| --- | --- |
|  | 6 |
|  | 8 |
|  | 13 |
|  | 12 |
|  | 10 |
|  | 7 |
|  | 4 |
|  | 2 |
|  | 1 |

Questions 1 – 9 are worth 3 points apiece.

You may use R to find the answer to 1-4. Make sure and include you R code for full credit.

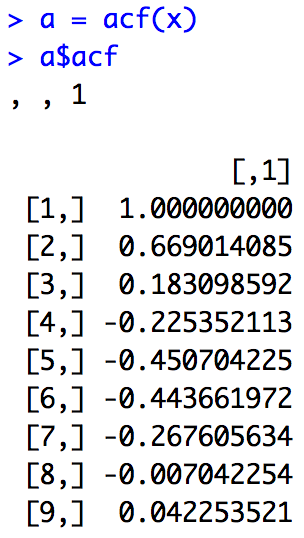
1. Calculate



var(x) =

2. Calculate

3. Calculate and how many pairs were used to find this estimate? 6



4. Which pair(s) would be used to calculate ? (1,8), (2,6)

5. Given the model: , calculate (“by hand” and show the steps)

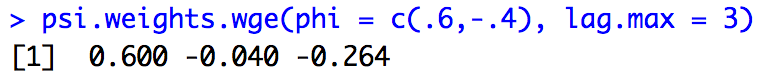
.6(1) + (-.4)(2)+7(1-.6+.4)

.6 - .8 + 7(.8)

-.2 + 5.6

5.4

6. Write an approximation of the GLP form of this model using the first three psi weights. You may use R to find the first 3 psi weights. Please include your code along with the GLP form of the model.

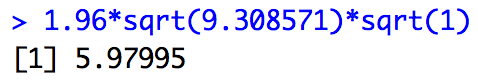


Given the way I wrote the problem, I would have also taken:

7. *Show your work and all steps* in finding an approximation of the margin of error of the prediction interval for . For simplicity, you only need to use the first 3 psi weights in the calculation.

*This admittedly was a poorly worded question and it was not my intent to ask you to find the white noise variance by hand although that is how the question is worded. In addition, you of course don’t need the first three psi weights to calculate the margin of error (MOE) (half width) of the prediction interval for .*

*Below is the calculation after using for.arma.wge() to get the estimate of the white noise variance (wnv = 9.308571). Everyone received full credit for this problem.*



5.97995

8. Say that we were later able to observe What is the value of ? Show your work and all steps (hint: there aren’t many steps.)

= 6-5.4 = .6

9. Provide a plot of the realization including the first 4 forecasts with prediction limits. Include your R code as well.



fore.arma.wge(x,phi = c(.6,-.4), n.ahead = 4)

Questions 10 – 14 are worth 3 points each.

Match the Realization with the spectral density:

|  |  |
| --- | --- |
| Realization | ACF |
| 10. |  |
| 11. |  |
| 12. |  |
| 13. |  |
| 14. |  |

Questions 15 – 20 are worth 3 points each.

Match each model on the left with an ACF OR spectral density on the right … each spectral density or ACF will be used only once and process of elimination may need to be used.

|  |  |
| --- | --- |
| 15. AR(1) |  |
| 16. AR(4) 2 sets of complex conjugate roots |  |
| 17. ARIMA(0,1,0) |  |
| 18. MA(2) with real roots |  |
| 19. AIRLINE MODEL |  |
| 20. MA(2) with complex conjugate roots |  |

Discussion ( 5 points):

21. Consider the Pennsylvania temperature data (dataset “*patemp”* in tswge package).



Is this data stationary or non-stationarity. Discuss and defend your argument addressing the 3 conditions of stationarity.

Condition 1: Mean does not depend on time: This is temperature data and it is reasonable to believe the temperature will be warmer in the summer and colder in the winter and in between in the fall and spring. For this reason this the data violate this assumption as if we were able to observe another realization we would observe the highs and lows in the same places relative to time.

Condition 2: Variance is finite and does not depend on time: This one takes some thinking as we have to ask ourselves if it is reasonable to think that temperature are more variable in different time periods (example Spring and Summer). I think it is reasonable to believe that they vary roughly the same and thus this condition may not be violated… I could see a good argument the other way as well. (Maybe the temperatures in the Spring and Fall are more consistent than the highs and lows of Summer and Winter). It is clear that the variance is finite in this case.

Condition 3: Covariance depends only on how far apart in time not where in time: The realization looks to behave very consistently from start to finish thus there is very little evidence against constant covariance across time. However, as an initial check, we will looks that the autocovariance structure for both the first and second half of the data:

|  |  |
| --- | --- |
| acf(patemp[1:90]) | acf(patemp[91:180]) |
|  |  |

It is easy to see that the acfs are nearly identical which is consistent with constant autocovariance.

Condition 3:

Multiple Choice:

Questions 22 – 25 are worth 3 points each.

22. Consider the following statements:

i. An AR(2) with and with a system frequency greater than 0 but less than .5 must have a negative .

ii. An autoregressive model can be written as an infinite order GLP.

iii. When time is recorded at the integers, the Nyquist frequency is 0.5.

Pick the correct answer:

A. Only i and ii are true.

B. Only ii and iii are true

C. Only i and iii are true

D. All are true

E. None are true.

23. True / False: After differencing the data or taking out other non-stationary factors, we should have white noise. (We could have stationary data with structure (example ARMA(p,q) with p ≠ 0 and/or q ≠ 0.)

24. What type of models will produce forecasts that will converge directly to the sample mean in a fashion similar to what is displayed below?



A. AR(1) positive phi

B. AR(1) negative phi

C. AR(2) complex conjugate roots.

d. AR(4) with two sets of complex conjugate roots

e. airline models

f. ARIMA(0,1,0) models

g. signal + noise models

25. What type of models will continue the seasonality and will continue the trend set by the last two observed values?

A. AR(1) positive phi

B. AR(1) negative phi

C. AR(2) complex conjugate roots.

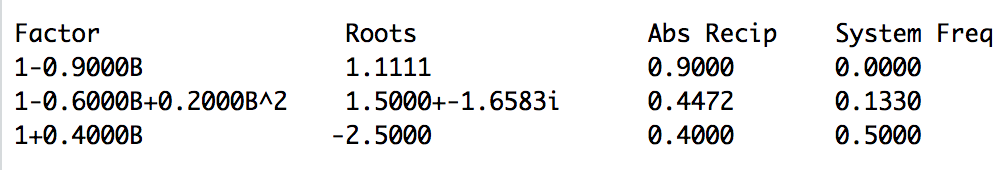
d. AR(4) with two sets of complex conjugate roots

e. airline models

f. ARIMA(0,1,0) models

g. signal + noise models

For Question 26, consider the following factor table:



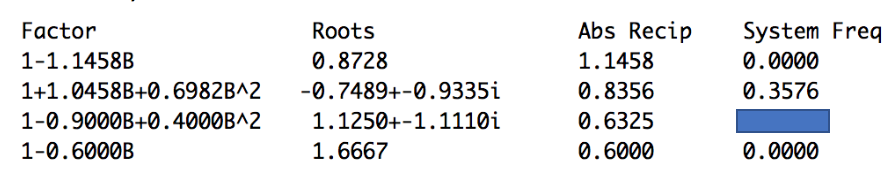
26. (3pts) Which behavior will dominate realizations from this model?

A. Wandering …. Phi = .9 and abs recip = .9 (closer to 1 than other factors… by a lot).

B. Rapid oscillating behavior

C. Periodic behavior with period of about 7.5  
D. All behavior will be exhibited roughly equally.

For question 27 and 28, consider the factor table below:



27. (3pts) Yes / No Is the factor table above associated with a stationary or non-stationary model? Non-Stationary There is a root inside the unit circle (first one).

28. (2 pt) What is the value of the missing System Frequency in the factor table above?

29. (15 pts) For Question 29 we will be using the last portion of the global temperature data from the tswge package. You can access the data with the following lines of code assuming you already have the tswge package loaded:

data("global.temp")

gt = global.temp[100:length(global.temp)]

Which model do you think is most appropriate/useful for forecasting global temperature data above?

Calculate the ASE for both models from forecasts of the last 10 values of the dataset. Provide at least 2 arguments (one can be the ASEs) as to why the model you selected is more useful than the other two in predicting the next 10 years of global temperatures. You can think of me as the client. Provide as many visual aids and explanation as is reasonable to get your point across clearly. Provide all R code used in your solution / analysis.

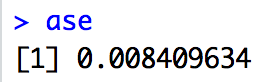
Model 1

f.gt.Model1 = fore.sigplusnoise.wge(gt,n.ahead = 10,last = TRUE)

ase = mean((f.gt.Model1$f - gt[(length(gt) - 9):length(gt)])^2)

ase





This is a model for non-stationary data with constant (linear) trend.

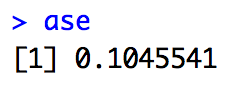
Model 2

f.gt.Model2 = fore.aruma.wge(gt, phi = c(-.62,.06, -.31), theta = c(-.17, -.75), d = 1, n.ahead = 10, lastn = TRUE)

ase = mean((f.gt.Model2$f - gt[(length(gt)- 9):length(gt)])^2)

ase





Model 1 (the signal plus noise model) is a model for non-stationary data that will forecast the upward trend to continue as can be seen in the forecasts above. The forecasts do not appear to behave like the realization although the ASE (.008) suggests that it performs better overall with respect to Model 2.

Reasons for Choosing Model 1

1. Smaller ASE

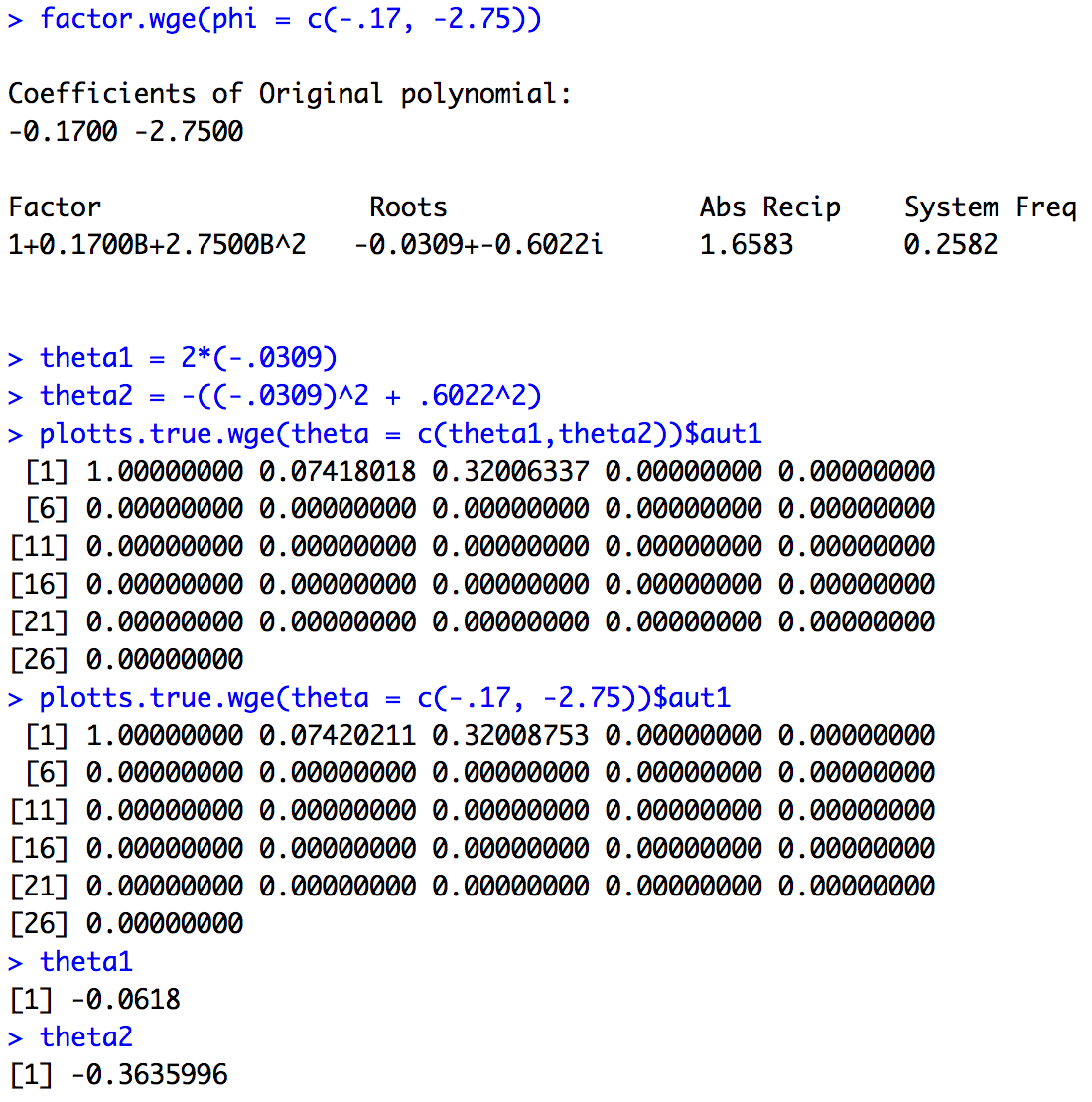
2. If someone believes there is a positive trend this would be the model that would include that belief in the forecasts.

3. While Model 2 (the ARIMA(3,1,2))has forecasts that are noticeably different from Model 1, they also behave differently that the realization. The larger ASE (.104) supports this idea as well.

**BONUS +2**

Given the non-invertible model below,

find an invertible model that has the same acf and spectral density.



+.0618