DTU02417, Time Series Analysis, Assignment 2
ARMA Processes and Seasonal Processes

Edward J. Xu (Jie Xu), s181238, DTU Management, edxu96@outlook.com

April 4th, 2019

1 Question 1, Stability

Let the ARMA(2,1) process X_t be given by:

$$X_t - 0.8X_{t-1} - 0.1X_{t-2} = \varepsilon_t + 0.8\varepsilon_{t-1} \tag{1}$$

where ε_t is a white noise process with $\sigma = 0.2$.

It can be written in standard form:

$$\phi(\mathrm{B})Y_t = \theta(\mathrm{B})\varepsilon_t \tag{2}$$

$$\phi(\mathrm{B}) = 1 - 0.8B - 0.1B^2 \tag{3}$$

$$\theta(\mathrm{B}) = 1 + 0.8B \tag{4}$$

$$H(z) = \frac{\theta\left(z^{-1}\right)}{\phi\left(z^{-1}\right)} = \frac{1+0.8z^{-1}}{1-0.8z^{-1}-0.1z^{-2}} \tag{5}$$

1.1 Is the process stationary and invertible?

The roots of $\phi\left(z^{-1}\right)$ is $\frac{1}{10}(4-\sqrt{26})$ and $\frac{1}{10}(-4+\sqrt{26})$, which are within the unit circle, thus this process is stationary. The root of $\theta\left(z^{-1}\right)$ is -0.8, which is within the unit circle, so this process is invertible.

1.2 Simulate 10 realisations with 200 observations from the process and plot them (Preferably in the same plot)

The randomness seed is set to 1.

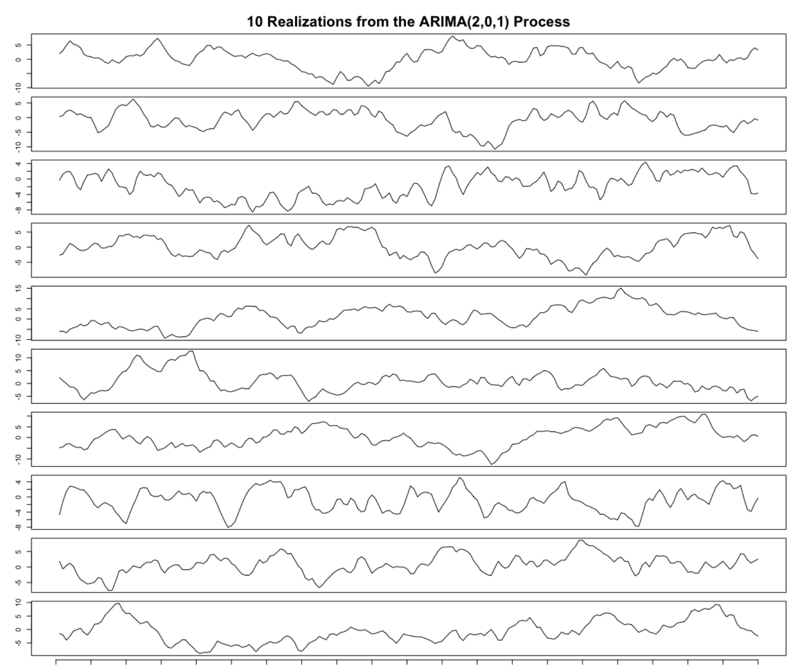

Figure 1. 10 Realizations from the ARIMA(2,0,1) Process

It seems to be very different between 10 realizations.

1.3 The second order moment representation of the process.

We get the following equations from eq(5.97) and eq(5.99) from the book[1]:

$$\gamma_{\varepsilon X}(m) = \begin{cases} = 0 & \text{for } m < 0 \\ \neq 0 & \text{for } m \geq 0 \end{cases} \tag{6}$$

$$\gamma_{\varepsilon X}(k) - 0.8\gamma_{\varepsilon X}(k-1) - 0.1\gamma_{\varepsilon X}(k-2) = \theta_k \sigma^2 \qquad k = 0,1 \tag{7}$$

$$\gamma_{XX}(k) - 0.8\gamma_{XX}(k-1) - 0.1\gamma_{XX}(k-2) = \theta_k \gamma_{\varepsilon X}(0) + \ldots + \theta_1 \gamma_{\varepsilon X}(1-k) \qquad k = 0,1 \tag{8}$$

$$\gamma_{XX}(k) - 0.8\gamma_{XX}(k-1) - 0.1\gamma_{XX}(k-2) = 0 \qquad k = 2,3,4,\ldots \tag{9}$$

Then:

$$\gamma_{\varepsilon X}(0) = \sigma^2 = 0.04 \qquad \text{when } k = 0 \tag{10}$$

$$\gamma_{\varepsilon X}(1) - 0.8\gamma_{\varepsilon X}(0) = 0.8\sigma^2 \qquad \text{when } k = 1 \tag{11}$$

$$\gamma_{\varepsilon X}(1) = 1.6\sigma^2 = 0.064 \tag{12}$$

Then:

$$\gamma_{XX}(0) - 0.8\gamma_{XX}(1) - 0.1\gamma_{XX}(2) = \gamma_{\varepsilon X}(0) + 0.8\gamma_{\varepsilon X}(1) \tag{13}$$

$$= 0.0912 \qquad \text{when } k = 0 \tag{14}$$

$$\gamma_{XX}(1) - 0.8\gamma_{XX}(0) - 0.1\gamma_{XX}(1) = 0.8\gamma_{\varepsilon X}(0) \tag{15}$$

$$= 0.032 \qquad \text{when } k = 1 \tag{16}$$

$$\gamma_{XX}(2) - 0.8\gamma_{XX}(1) - 0.1\gamma_{XX}(0) = 0 \qquad \text{when } k = 2 \tag{17}$$

$$\gamma_{XX}(k) - 0.8\gamma_{XX}(k-1) - 0.1\gamma_{XX}(k-2) = 0 \qquad \text{when } k > 2 \tag{18}$$

$$\tag{19}$$

Finally, we get:

$$\gamma_{XX}(0) = 0.5896 \tag{20}$$

$$\gamma_{XX}(1) = 0.5596 \tag{21}$$

$$\gamma_{XX}(2) = 0.5066 \tag{22}$$

$$\gamma_{XX}(k) = 0.8\gamma_{XX}(k-1) + 0.1\gamma_{XX}(k-2) \tag{23}$$

$$\tag{24}$$

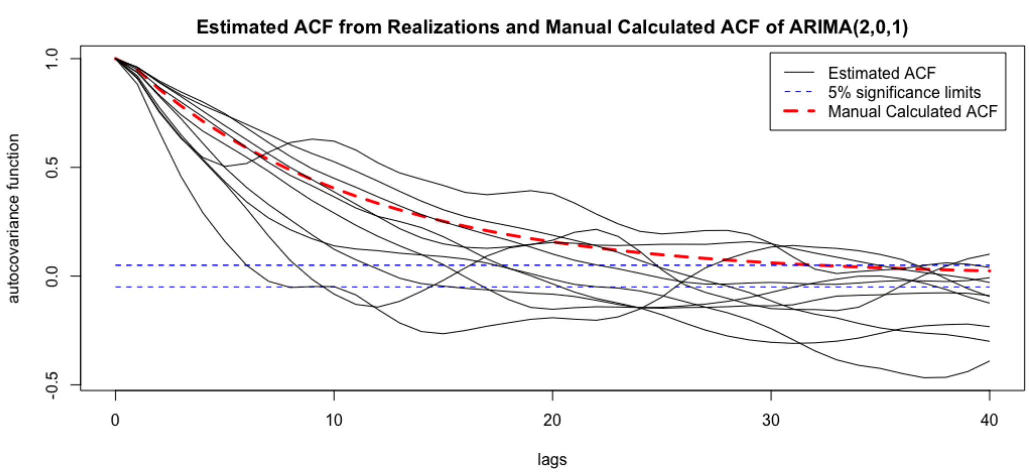

Figure 2. Estimated ACF from Realizations and Manual Calculated ACF of ARIMA(2,0,1)

Most of the ACF-plot seems to be consistent with the calculation result, while some don't, according to fig.2. However, the overall trend is right. The PACF-plot can be seen in fig.3, which seems to be fit with theoretical result quite nicely in the first 4 lags. The deviations in the following lags are acceptable.

The calculation of the PACF is through the following code.

```r
vec_pacf <- ARMAacf(c(0.8, 0.1), 0.8, lag.max = 20, pacf = TRUE)
```

Code 1. Functional Calculation of PACF

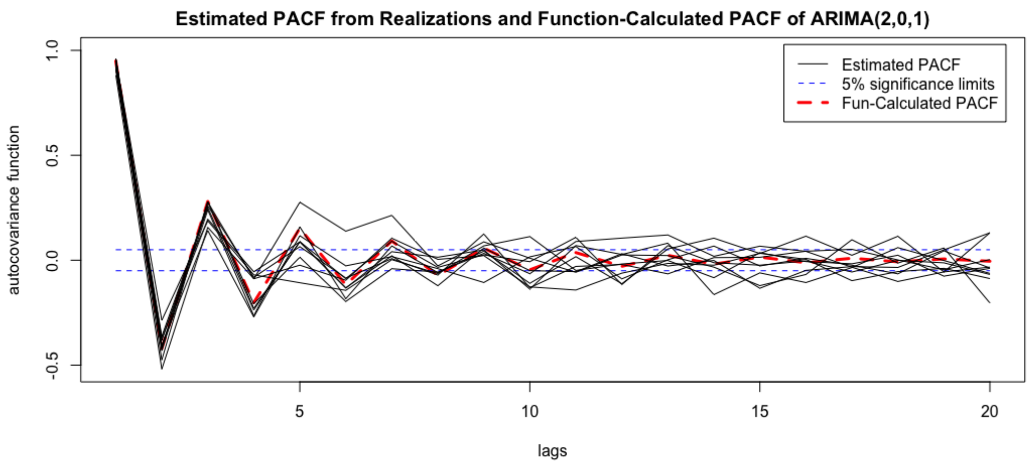

Figure 3. Estimated PACF from Realizations and Function-Calculated PACF of ARIMA(2,0,1)

1.4 Question 2, Predicting the Carbon-Dioxide Level in an Office

logarithm of the CO2 level in parts per million in an office, measured every three hours. Based on historical data the following ARIMA(1,8,0) model has been identified:

$$(1-0.547B)\left(1-0.860B^8\right)(\log(Y_t)-\mu)=\varepsilon_t \tag{25}$$

with $\phi_1 = -0.547$.

We assume a model with $X_t = \log(Y_t)-\mu$ to simplify the expression:

$$\begin{aligned}
\varepsilon_t &= (1-0.547B)\left(1-0.86B^8\right)X_t && (26)\\
&= \left(1-0.86B^8\right)X_t - 0.547\left(1-0.86B^8\right)X_{t-1} && (27)\\
&= X_t - 0.86X_{t-8} - 0.547X_{t-1} + 0.47042X_{t-9} && (28)\\
X_t &= 0.547X_{t-1} + 0.86X_{t-8} - 0.47042X_{t-9} + \varepsilon_t && (29)\\
X_{t+k} &= 0.547X_{t+k-1} + 0.86X_{t+k-8} - 0.47042X_{t+k-9} + \varepsilon_{t+k} && (30)
\end{aligned}$$

The optimal prediction is:

$$\begin{aligned}
\hat{X}_{t+1|t} &= \mathrm{E}\left[X_{t+1}|X_t, X_{t-1}, X_{t-2}, ..., X_{t-15}\right] && (31)\\
&= \mathrm{E}\left[0.547X_t + 0.86X_{t-7} - 0.47042X_{t-8} + \varepsilon_{t+1}|X_t, X_{t-1}, X_{t-2}, ..., X_{t-15}\right] && (32)\\
&= 0.547X_t + 0.86X_{t-7} - 0.47042X_{t-8} && (33)\\
\hat{X}_{t+2|t} &= \mathrm{E}\left[X_{t+2}|X_t, X_{t-1}, X_{t-2}, ..., X_{t-15}\right] && (34)\\
&= \mathrm{E}\left[0.547X_{t+1} + 0.86X_{t-6} - 0.47042X_{t-7} + \varepsilon_{t+2}|X_t, X_{t-1}, X_{t-2}, ..., X_{t-15}\right] && (35)\\
&= 0.547\mathrm{E}\left[X_{t+1}|X_t, X_{t-1}, X_{t-2}, ..., X_{t-15}\right] + 0.86X_{t-5} - 0.47042X_{t-6} && (36)\\
&= 0.299209X_t + 0.86X_{t-5} - 0.47042X_{t-6} + 0.47042X_{t-7} - 0.25731974X_{t-8} && (37)
\end{aligned}$$

The prediction error is:

$$\begin{aligned}
e_{t+1|t} &= X_{t+1} - \hat{X}_{t+1|t} && (38)\\
&= 0.547X_t + 0.86X_{t-7} - 0.47042X_{t-8} + \varepsilon_{t+1} - 0.547X_t - 0.86X_{t-7} + 0.47042X_{t-8} && (39)
\end{aligned}$$

$$
\begin{aligned}
&= \varepsilon_{t+1} && (40)\\
e_{t+2|t} &= X_{t+2} - \hat{X}_{t+2|t} && (41)\\
&= 0.547X_{t+1} + 0.86X_{t-6} - 0.47042X_{t-7} + \varepsilon_{t+2} - 0.299209X_t - 0.86X_{t-5} + 0.47042X_{t-6} - 0.47042X_{t-7} + 0.25731974X_{t-8} && (42)\\
&= 0.547X_{t+1} - 0.299209X_t - 0.86X_{t-5} + 1.33042X_{t-6} - 0.94084X_{t-7} + 0.25731974X_{t-8} + \varepsilon_{t+2} && (43)\\
&= -0.86X_{t-5} + 1.33042X_{t-6} - 0.47042X_{t-7} + 0.547\varepsilon_{t+1} + \varepsilon_{t+2} && (44)
\end{aligned}
$$

The variance of the prediction error is:

$$
\begin{aligned}
\operatorname{Var}\left[e_{t+1|t}\right] &= \operatorname{Var}\left[\varepsilon_{t+1}\right] && (45)\\
&= 4.225E-3 && (46)\\
\operatorname{Var}\left[e_{t+2|t}\right] &= \operatorname{Var}\left[-0.86X_{t-5} + 1.33042X_{t-6} - 0.47042X_{t-7} + 0.547\varepsilon_{t+1} + \varepsilon_{t+2}\right] && (47)\\
&= 6.536075E-3 && (48)
\end{aligned}
$$

The $(1-\alpha)$-prediction interval is:

$$
\begin{aligned}
&\hat{X}_{t+1|t} \pm u_{\alpha/2}\sqrt{\operatorname{Var}\left[e_{t+1|t}\right]} && (49)\\
&\hat{X}_{t+2|t} \pm u_{\alpha/2}\sqrt{\operatorname{Var}\left[e_{t+2|t}\right]} && (50)
\end{aligned}
$$

Then, we use $\log(Y_t) = X_t + \mu$ to calculate the logarithm of CO2 level.
The result is shown in fig.4

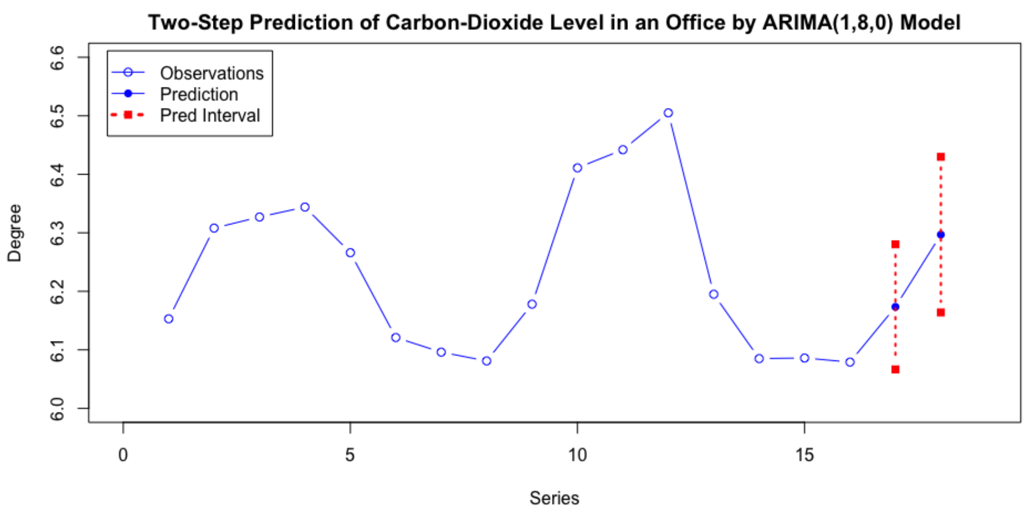

Figure 4. Two-Step Prediction of Carbon-Dioxide Level in an Office by ARIMA(1,8,0) Model

Why is a seasonality of 8 steps reasonable in this model?
According to the argument in page 152 from book[1],

> In cases where it is possible to apply physical knowledge, one should most certainly not disregard such knowledge but combine it with information from data.

The time unit of the data is 3 hours, so the seasonality of 8 steps means that it is based on the fact that there are 24 hours in every day. It's more reasonable to model the process using the seasonality of 8 steps then others.

2 Question 3, Simulating Seasonal Processes

The process Y_t of $(2,0,1)\times(1,0,1)_{12}$ model can be expressed as:

$$(1+\phi_1 B+\phi_2 B^2)(1+\Phi_1 B^{12})Y_t = (1+\theta_1 B)(1+\Theta_1 B^{12})\varepsilon_t \tag{51}$$

$$(1+\phi_1 B+\phi_2 B^2+\phi_1 B^{12}+\phi_1\Phi_1 B^{13}+\phi_2\Phi_1 B^{14})Y_t=(1+\theta_1 B+\Theta_1 B^{12}+\theta_1\Theta_1 B^{13})\varepsilon_t \tag{52}$$

The process is stationary if the roots of the following equation are within the unit circle:

$$z^{14}+\phi_1 z^{13}+\phi_2 z^{12}+\phi_1 z^2+\phi_1\phi_1 z+\phi_2\phi_1=0 \tag{53}$$

Whether it's stationary or not can be judged by the following function:

```r
jud_stationary_arima <- function(vec_psi){
    # Different definitions of ARMA models have different signs for the AR and/or MA coefficients.
    # How to decide the sign of phi is that phi and past observations are in left-hand-side
    # 190402
    vec_psi <- c(1, vec_psi)
    vec.rev_psi <- rev(vec_psi)
    vec_root <- polyroot(vec.rev_psi)
    vec_root.abs <- abs(vec_root)
    vec_root.abs.unit <- abs(vec_root) <= rep(1, length(vec_root))
    return(all(vec_root.abs.unit))
}
```

Code 2. R function to judge whether the given ARIMA model is stationary

Because there is no function for seasonal arima model in R, we have to do the following conversion:

```r
mat_phi <- matrix(0, nrow = 6, ncol = 14)
mat_theta <- matrix(0, nrow = 6, ncol = 13)
list_stationary <- logical(length = 6)
for (i in seq(6)){
    phi_1 <- list_phi_1[i]
    phi_2 <- list_phi_2[i]
    phi.cap_1 <- list_phi.cap_1[i]
    theta_1 <- list_theta_1[i]
    theta.cap_1 <- list_theta.cap_1[i]
    mat_phi[i,] <- c(phi_1, phi_2, rep(0, 9), phi.cap_1, phi_1 * phi.cap_1, phi_2 * phi.cap_1)
    mat_theta[i,] <- c(theta_1, rep(0, 10), theta.cap_1, theta_1 * theta.cap_1)
    list_stationary[i] <- jud_stationary_arima(mat_phi[i,])
}
```

Code 3. R function to judge whether the given ARIMA model is stationary

The "mat_phi" "mat_theta" is shown below:

$$\text{mat\_phi}=\begin{bmatrix} 0.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.85 & 0 & 0 \\ -0.7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.7 & -0.56 & 0 \\ 0.6 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9 & 0.54 & 0.45 \\ 0.0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{54}$$

$$\text{mat\_theta}=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.7 & 0.63 \end{bmatrix} \tag{55}$$

All the models are stationary according to the result calculated by the above function.
If the $(P,D,Q)_s$ are not all 0, the model is seasonal. So model 1 is not seasonal, while model 2, 3, 4, 5, 6 are 12 seasonal.

2.1 Simulation and Visualization of ACF and PACF

Simulate the following models (where monthly data are assumed). Plot the simulations and the associated autocorrelation functions (ACF and PACF). The simulation seed is set to 99.

Model 1: A $(1,0,0)\times(0,0,0)_{12}$ model with the parameter $\phi_1=0.8$: $(1+0.8B)Y_t=\varepsilon_t$. We can see that the ACF shaped up to be damped sine functions, while PACF becomes insignificant when lags are larger than 1. So it's very likely a AR(1) model.

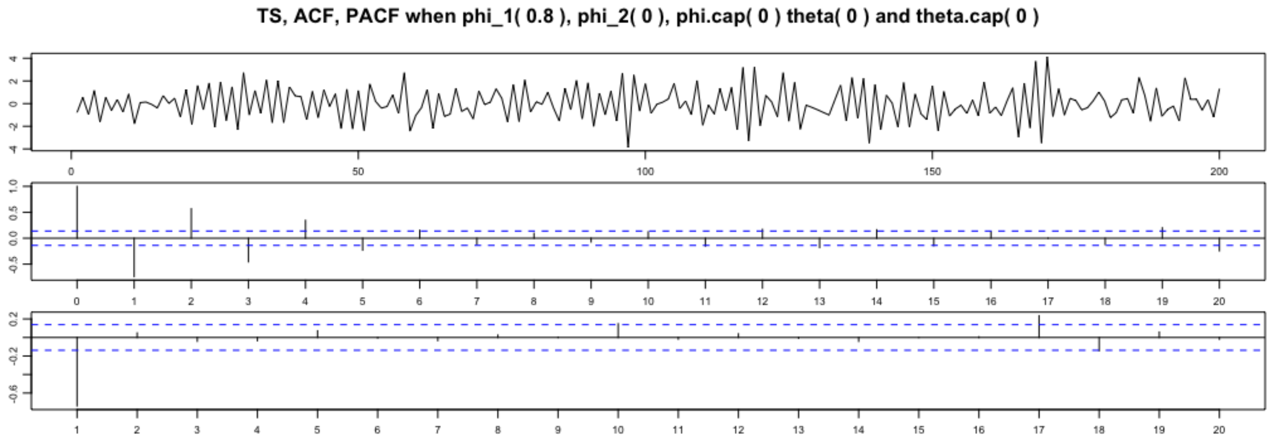

Figure 5. Time Series and its ACF, PACF from ARIMA with Parameters(0.8 0 0 0 0)

Model 2: A $(0,0,0)\times(0,0,1)_{12}$ model with the parameter $\phi_1 = -0.85$: $(1-0.85B^{12})Y_t = \varepsilon_t$. There is obvious sudden change in lages 12 in ACF and PACF, so it's a seasonal model with seasonal factor 12. The PACF becomes 0 from the start, so the MA factor is 0.

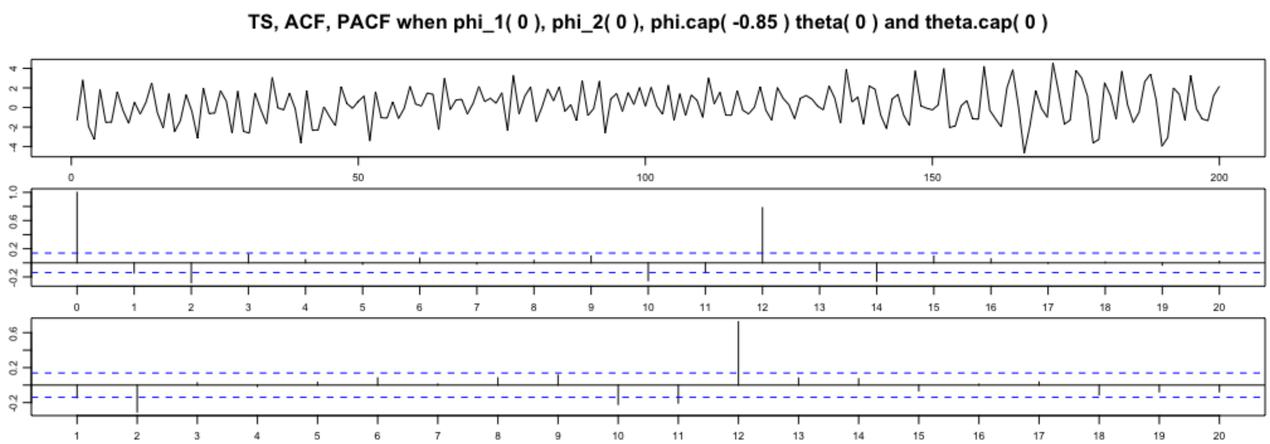

Figure 6. Time Series and its ACF, PACF from ARIMA with Parameters(0 0 -0.85 0 0)

Model 3: A $(1,0,0)\times(0,0,1)_{12}$ model with the parameter $\phi_1 = -0.7$ and $\Theta_1 = 0.8$: $(1-0.7B)Y_t = (1+0.8B^{12})\varepsilon_t$. The ACF shows some sign of exponential decreasing. Besides, the PACF becomes 0 after lag 1, except that in lag 13. So we can say it's a AR(1) model. Furthermore, There is obvious sudden change around lages 12 and 13 in ACF and PACF. According to the following simplified version of eq.52, it's very likely a seasonal AR(1) model with seasonal factor 12, nevertheless, we can not tell whether AR or/and MA has the seasonal factor.

$$(1+\phi_1 B+\phi_1 B^{12}+\phi_1\Phi_1 B^{13})Y_t = (1+\theta_1 B+\Theta_1 B^{12}+\theta_1\Theta_1 B^{13})\varepsilon_t \tag{56}$$

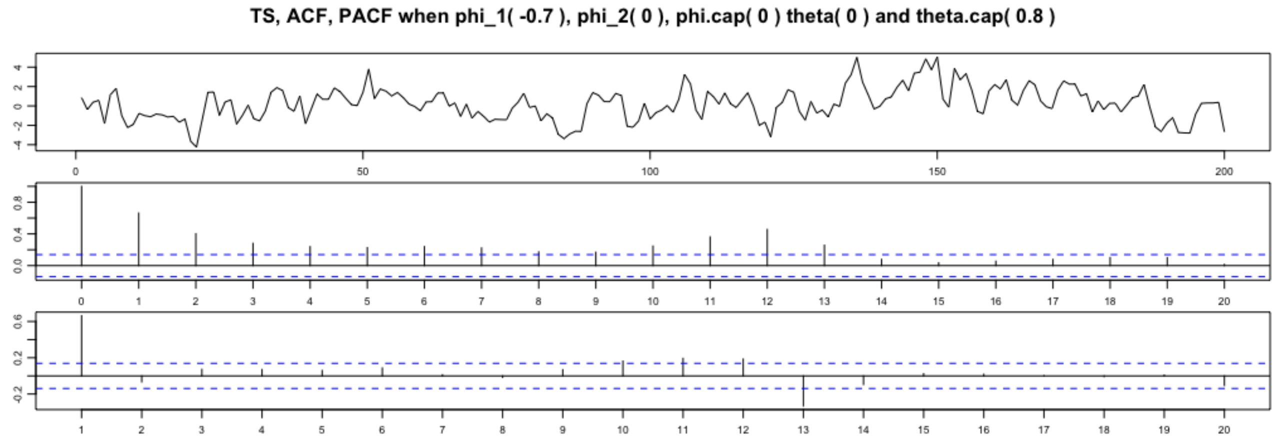

Figure 7. Time Series and its ACF, PACF from ARIMA with Parameters(-0.7 0 0 0 0)

Model 4: A $(1,0,0)\times(1,0,0)_{12}$ model with the parameter $\phi_1 = -0.8$ and $\phi_1 = 0.7$: $(1-0.8B)(1+0.7B^{12})Y_t = \varepsilon_t$. The ACF shows some sign of damped sine. Besides, the PACF becomes 0 after lag 1, except that in lag 13. So we can say it's a AR(1) model. Furthermore, There is obvious sudden change around lages 12 and 13 in ACF and PACF. According to the following simplified version of eq.52, it's very likely a seasonal AR(1) model with seasonal factor 12, nevertheless, we can not tell whether AR or/and MA has the seasonal factor.

$$(1+\phi_1 B+\phi_1 B^{12}+\phi_1\Phi_1 B^{13})Y_t = (1+\theta_1 B+\Theta_1 B^{12}+\theta_1\Theta_1 B^{13})\varepsilon_t \tag{57}$$

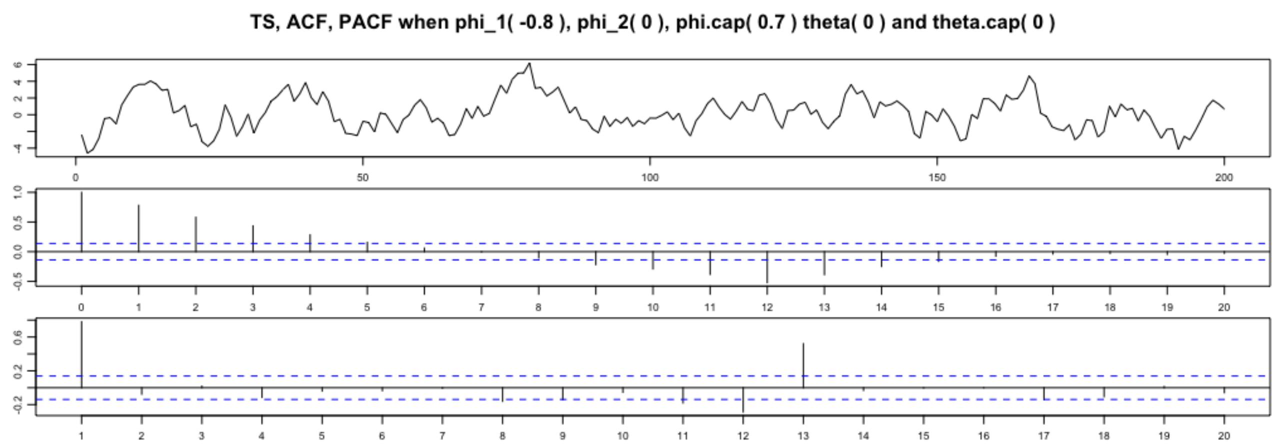

Figure 8. Time Series and its ACF, PACF from ARIMA with Parameters(0.8 0 0.7 0 0)

Model 5: A $(2,0,0)\times(1,0,0)_{12}$ model with the parameter $\phi_1 = 0.6$, $\phi_2 = 0.5$ and $\phi_1 = 0.9$: $(1+0.6B+0.5B^2)(1+0.9B^{12})Y_t = \varepsilon_t$. The ACF shows some sign of damped sine. Besides, the PACF becomes 0 after lag 2, except that in lag 12, 13 and 14. So we can say it's a AR(2) model. Furthermore, There is obvious sudden change around lages 12, 13 and 14 in ACF and PACF. According to the eq.52, it's very likely a seasonal AR(2) model with seasonal factor 12, nevertheless, we can not tell whether AR or/and MA has the seasonal factor.

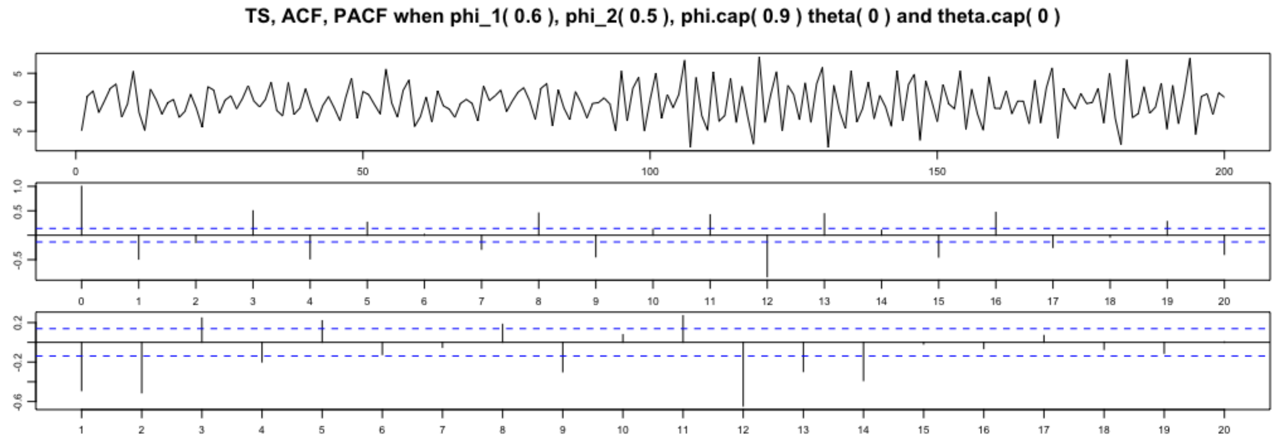

Figure 9. Time Series and its ACF, PACF from ARIMA with Parameters(0.6 0.5 0.9 0 0)

Model 6: A $(0,0,1)\times(0,0,1)_{12}$ model with the parameter $\theta_1 = -0.9$ and $\Theta_1 = 0.7$: $Y_t = (1-0.9B)(1+0.7B^{12})\varepsilon_t$. The PACF shows some sign of damped sine. Besides, the ACF becomes 0 after lag 1, except that in lag 11, 12 and 13. So we can say it's a MA(1) model. Furthermore, There is obvious sudden change around lages 11, 12 and 13 in ACF and PACF. According to eq.52, it's very likely a seasonal MA(1) model with seasonal factor 12, nevertheless, we can not tell whether AR or/and MA has the seasonal factor.

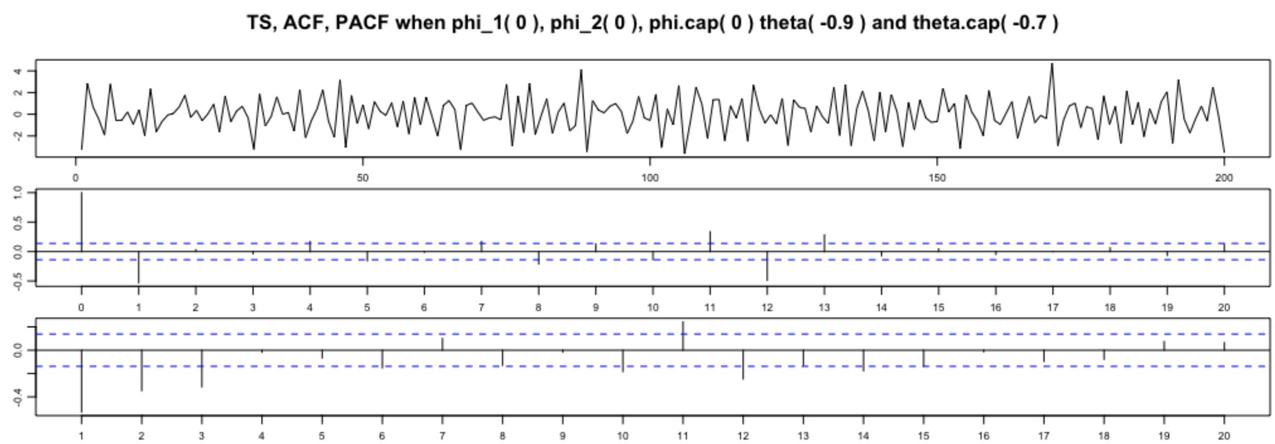

Figure 10. Time Series and its ACF, PACF from ARIMA with Parameters(0 0 0 -0.9 0.7)

2.2 Conclusion

1. Can you identify the model order based on the ACF and PACF?

We can identify whether it's AR or MA model if the ACF and PACF has regular damped exponential- or sine- shaped variation. Then, according to when PACF becomes zeros, we can identify the exact order, if it's an AR model, while it's decided according to ACF when it's a MA model.

If there is seasonal factors, thought it's obvious to see some abrupt ups and downs in the multiple lags of the season length, it's hard to identity the exact order. However, we can still tell whether it's AR or MA model with seasonal factors if it's pure AR or MA model.

If it's ARMA model with or without seasonal factor, we find it hard to decide the exact order.

2. Which conclusions can you draw on the general behavior of the autocorrelation function for seasonal processes?

If the order in AR or/and MA part is more than 1, the ACF and PACF of seasonal model can show significant change in the season-multiple lags and it's following lags. For example, according to eq.52, there may be order in 13 and 14 lags, when it's a 12-seasonal model.

3 Question 4

Consider the following process:

$$X_t + \phi X_{t-1} = \varepsilon_t + 0.4\varepsilon_{t-1} \tag{58}$$

with $\phi \in \{-0.9, -0.995\}$and $\sigma^2 \in \{0.1^2, 0.5^2\}$. Simulate 300 observations of each of the four processes 100 times. Subsequently, estimate the model parameters based on the simulated sequences. For each process, make a histogram plot of the estimates parameter ϕ and indicate the 95% quantiles. How do different values of ϕ affect the variance of the estimated ϕ? How do different values of σ affect the variance of the estimated ϕ?

The simulation seed is set to 99.

```
sim_ar1ma1 <- function(phi, theta, sigma, numtime_Sim){
    # Different definitions of ARMA models have different signs for the AR and/or MA coefficients.
    set.seed(99)
    lists_ts_ar1ma1 <- vector("list", 100)
    lists_mod_ar1ma1 <- vector("list", 100)
    vec_phi.est <- rep(0, 100)
    vec_theta.est <- rep(0, 100)
    for (i in seq(100)){
        lists_ts_ar1ma1[[i]] <- arima.sim(model = list(ar = - phi, ma = theta), sd = sigma, n = numtime_Sim)
        lists_mod_ar1ma1[[i]] <- arima(lists_ts_ar1ma1[[i]], order = c(1, 0, 1), method = "ML",
                                       include.mean = FALSE)
        # Remember to specify the estimation method
        vec_phi.est[i] <- lists_mod_ar1ma1[[i]]$coef[1]
        vec_theta.est[i] <- lists_mod_ar1ma1[[i]]$coef[2]
    }
    quan95 <- ApproxQuantile(hist(vec_phi.est, breaks = 50, plot = FALSE), 0.95)
    list_result <- list(quan95 = quan95, vec_phi.est = vec_phi.est, vec_theta.est = vec_theta.est)
    return(list_result)
}
```

Code 4. Function to Simulate ARMA Model and Estimate.

The estimation method is set to "maximum likelihood estimation". If it's set to "CSS", the result is similar. Nevertheless, if it's not set, there will be problem during the estimation sometimes. The reason is explained as, according to the teacher:

> The combination "CSS+ML" (the default) uses the "CSS" estimate as initial guess for the "ML" and will therefore bring you faster to the "ML" estimates - except when the "CSS" estimates are non-stationary as sometimes happen in this case.

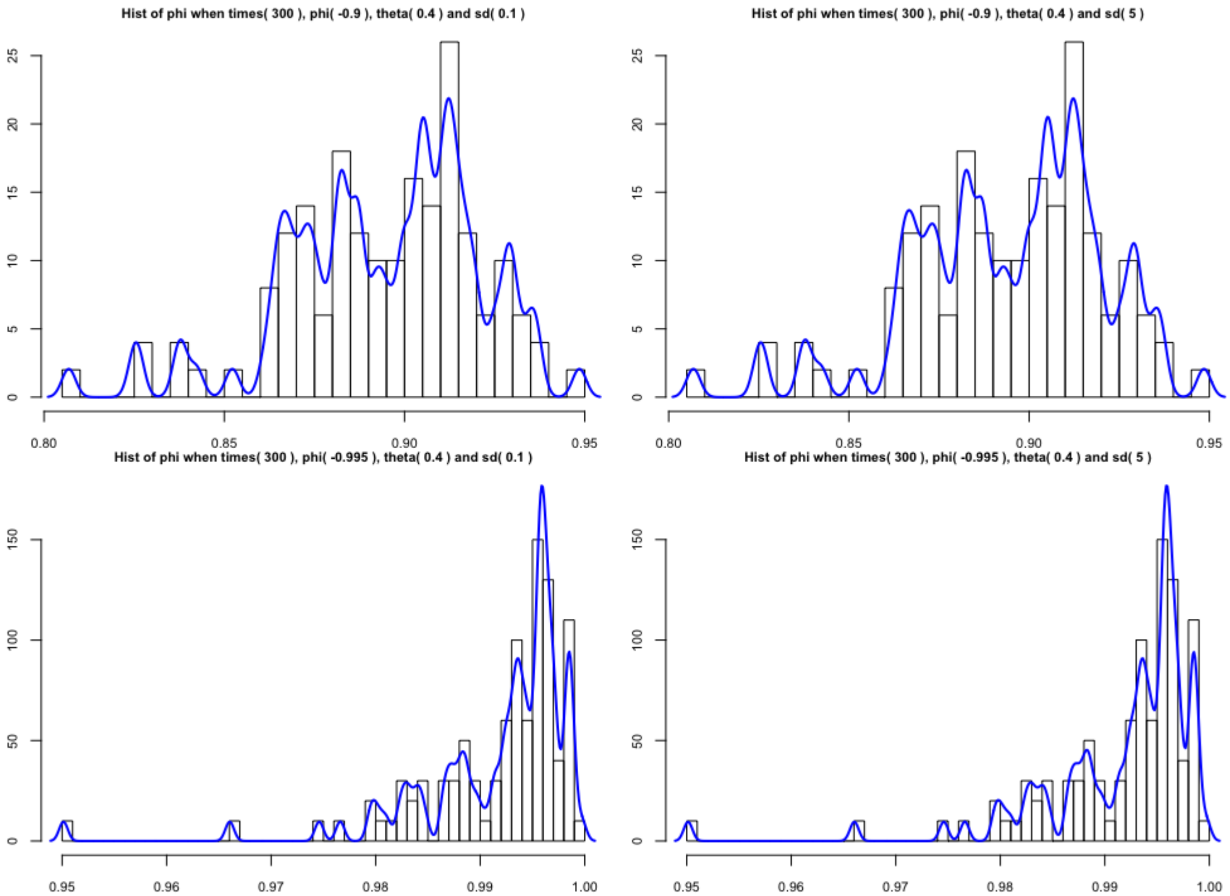

Figure 11. Histogram of Simulation and Estimation from Four ARIMA(1,0,1) Model

phi_1	sigma	95% quantile
-0.900	0.1	0.9325
-0.900	5.0	0.9325
-0.995	0.1	0.9985
-0.995	5.0	0.9985

Figure 12. %95 Quantiles of Simulation and Estimation from Four ARIMA(1,0,1) Model

Comparing the above four figures, we can say that the variance from white noise has few impact on the ARIMA estimation result, while the distribution of ARIMA estimation result varies significantly if there are some differences in ϕ.

Also, larger ϕ causes more deviations in the estimated parameters.

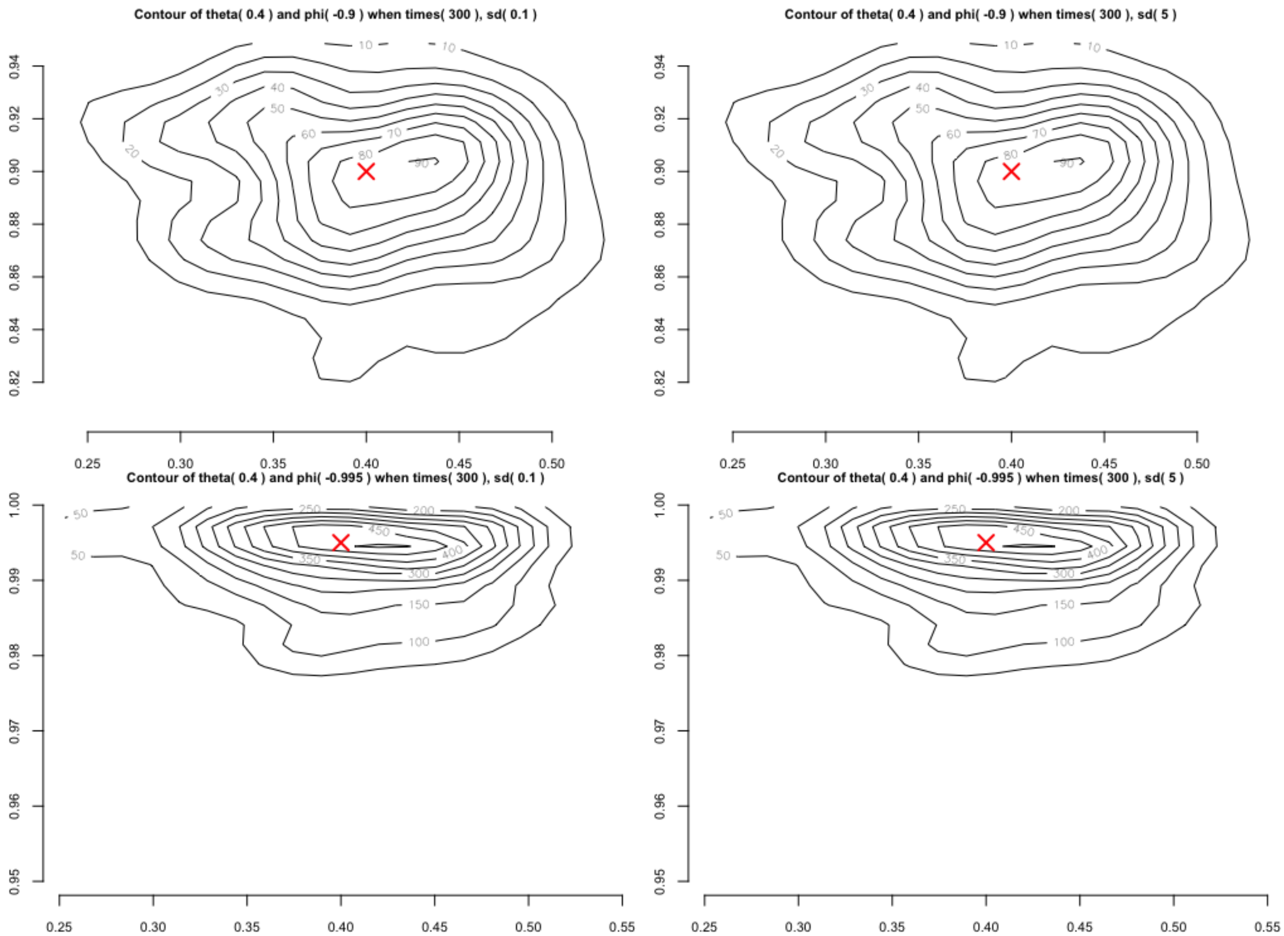

Figure 13. Contour Plot of of Simulation and Estimation from Four ARIMA(1,0,1) Model

Comparing the Q-Q plot of different ϕ, we find that larger ϕ cause the distribution of estimation to deviate from the Q-Q line, which means more simulations and estimations are needed to get a good normal distribution of the result.

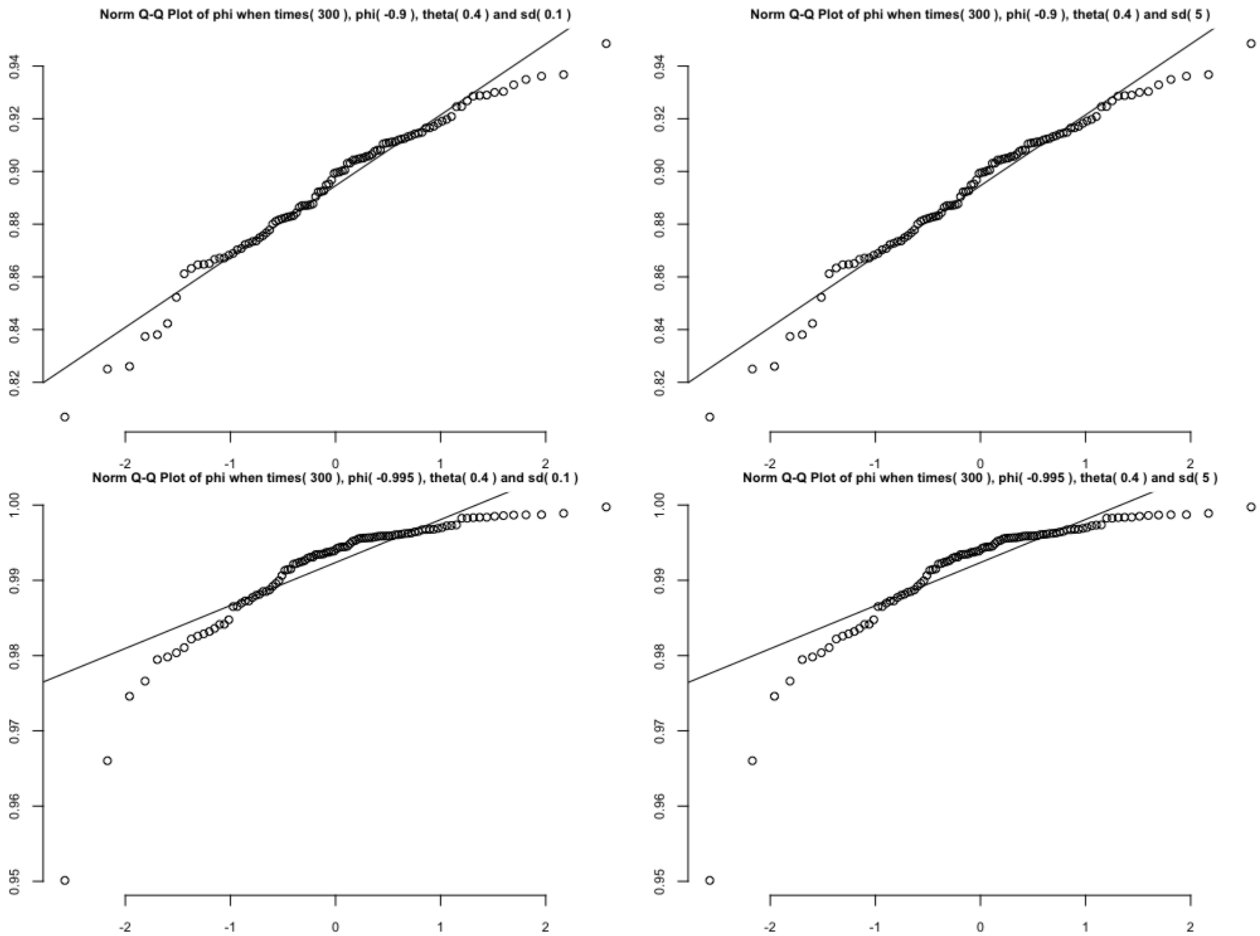

Figure 14. Normal Q-Q Plot of of Simulation and Estimation from Four ARIMA(1,0,1) Model

3.1 The Effect of Number of Simulated Observations

Try different numbers of observations (e.g. 100 and 1000) in the simulation (with $\phi = 0.9$ and $\sigma = 1$) and investigate how it affects the variance of the estimates.

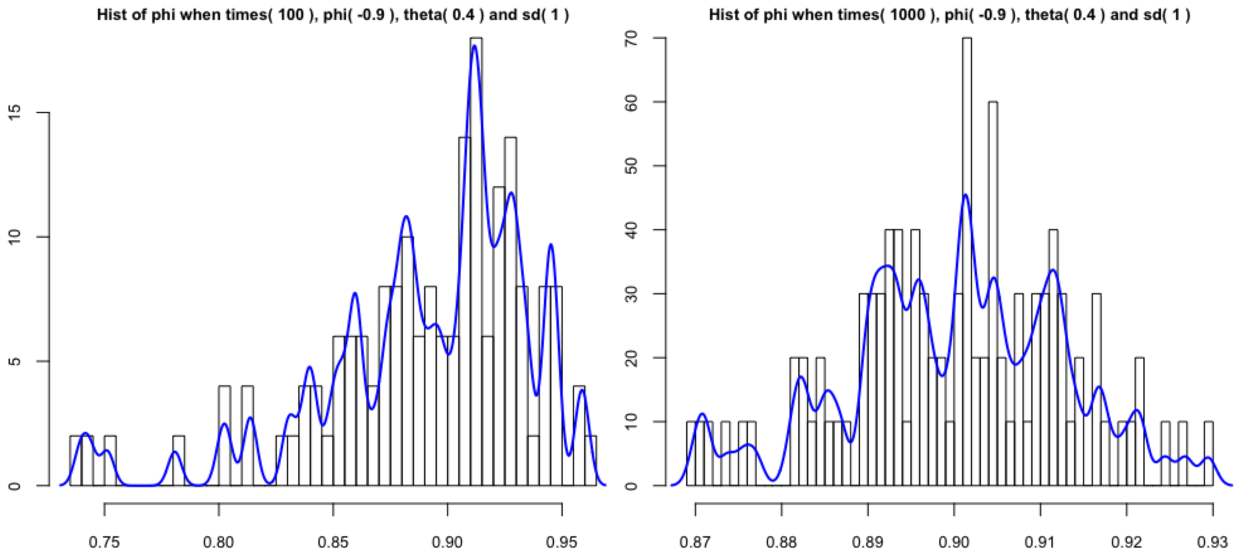

Figure 15. Histogram of of Simulation and Estimation from Two ARIMA(1,0,1) Model

According to contour plots, the distribution of the estimations are more intense around true value if there are more simulation times.

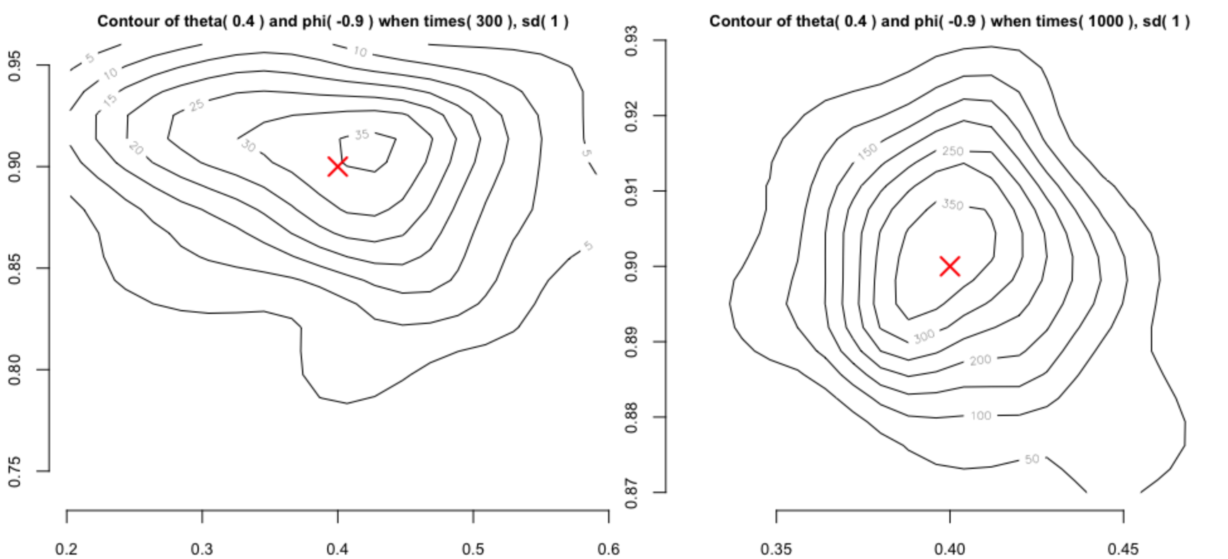

Figure 16. Contour Plot of of Simulation and Estimation from Two ARIMA(1,0,1) Model

We can see from the Q-Q plot (fig.17) that more simulations contribute the result to a further approximation to normal distribution.

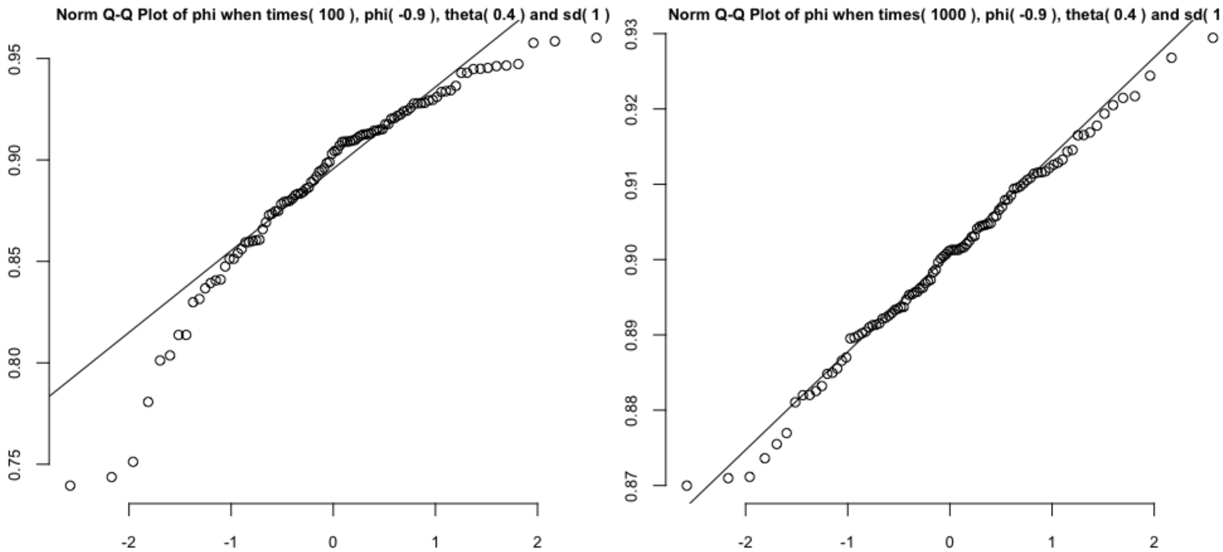

Figure 17. Normal Q-Q Plot of of Simulation and Estimation from Two ARIMA(1,0,1) Model

References

1. Madsen, H. *Time series analysis* (Chapman and Hall/CRC, 2007).