TEXAS A&M UNIVERSITY

Department of Statistics

STAT626 - Methods in Time Series Analysis
Homework #6

Salih Kilicli

July 19, 2019

Problem 1: **Problem 4.1 pg 95, textbook.**

Solution: For MA(1) model: $x_t = w_t + \theta w_{t-1}$. The autocovariance and autocorrelation functions at 1 lag are given by:

$$\gamma_x(0) = var(x_t) = E[x_t^2] = E[w_t^2 + 2\theta w_t w_{t-1} + \theta^2 w_{t-1}^2] = (1+\theta^2)\sigma_w^2$$
$$\gamma_x(1) = E[x_{t+1}x_t] = E[w_{t+1}w_t + w_{t+1}\theta w_{t-1} + w_t\theta w_t + \theta^2 w_t w_{t-1}] = \theta\sigma_w^2$$

$$\rho_x(1) = \frac{\gamma_x(1)}{\gamma_x(0)} = \frac{\theta\sigma_w^2}{(1+\theta^2)\sigma_w^2} = \frac{\theta}{1+\theta^2}$$

Now, consider the equation $\rho_x(1)\theta^2 + \theta + \rho_x(1) = 0$. The type roots of the equation depends on the sign of $\nabla = \sqrt{B^2 - 4AC} = \sqrt{1 - 4\rho_x(1)^2}$. We have real solutions only when $\nabla = \sqrt{1 - 4\rho_x(1)^2} \geq 0$. This inequality implies that:

$$\sqrt{1 - 4\rho_x(1)^2} \geq 0$$
$$1 - 4\rho_x(1)^2 \geq 0$$
$$1 \geq 4\rho_x(1)^2$$
$$|\rho_x(1)| \leq \frac{1}{2}$$

In order to find the min and max values of $\rho_x(1)$ let us find the roots of first derivative function. $\frac{d\rho_x(1)}{d\theta} = \frac{1-\theta^2}{(1+\theta^2)^2} = 0$ implies the roots are $\theta = \pm 1$. Now, plugging the roots in the function we see that $max\{\rho_x(1)\} = \frac{1}{2}$ when $\theta = 1$ and $min\{\rho_x(1)\} = -\frac{1}{2}$ when $\theta = -1$. Therefore, $\rho_x(1)$ attains its maximum at $\theta = 1$ whereas it attains its minimum at $\theta = -1$.

Problem 2: **Problem 4.3 pg 95-96, textbook.**

Solution: Consider the models:

(i) $x_t = 0.80x_{t-1} - 0.15x_{t-2} + w_t - 0.30w_{t-1}$

(ii) $x_t = x_{t-1} - 0.50x_{t-2} + w_t - w_{t-1}$

a) First model can be reduced as $x_t = 0.5x_{t-1} + w_t$ since:

$$\begin{aligned}(1 - 0.8B + 0.15B^2)x_t &= (1 - 0.3B)w_t \\ \cancel{(1 - 0.3B)}(1 - 0.5B)x_t &= \cancel{(1 - 0.3B)}w_t \\ (1 - 0.5B)x_t &= w_t\end{aligned}$$

Second model cannot be reduced because the roots of autoregregressive operator $\phi(B)$ are $B = 1 \pm i$ and the only root of moving average operator $\theta(B)$ is $B = 1$.

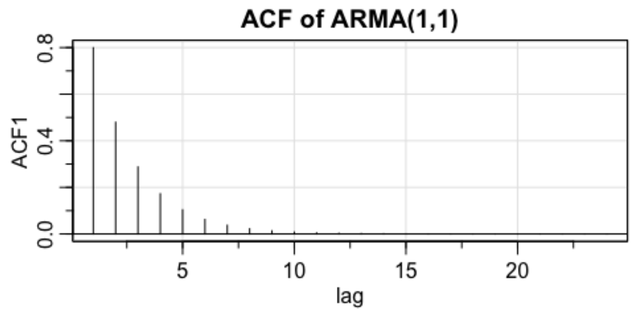
b) Notice, the root of AR term $\phi(B)$ in (i), $B = 2 > 1$ so (i) is causal. Since the reduced form is an AR(1) model (i) is also invertible with $\pi_0 = 1, \pi_1 = -0.5$. The roots of AR term $\phi(B)$ in (ii) $B = 1 \pm i$ with $|1 \pm i| = \sqrt{2} > 1$. Hence, (ii) is causal. However, the root of MA term $\theta(B)$ is $B = 1 \geq 1$; therefore, it is <u>not</u> invertible.

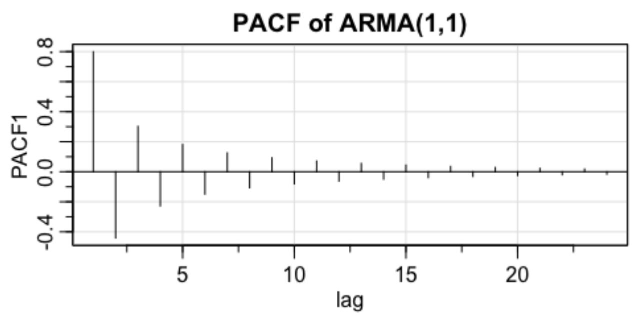
c) Notice, (i) is causal and invertible since coefficients are approaching to 0 in either case. Similarly, (ii) is causal for same reason. However, (ii) is **not** invertible since the coefficient (π_j's) are converging to a non-zero value 0.5. R output giving the same results above is given below.

```
> round(ARMAtoMA(ar=0.5,ma=0,50),2) #causal       (since coeffs \psi are converging to 0)
 [1] 0.50 0.25 0.12 0.06 0.03 0.02 0.01 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
[24] 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
[47] 0.00 0.00 0.00 0.00
> round(ARMAtoAR(ar=0.5,ma=0,50),2) #invertible   (since coeffs \pi are converging to 0)
 [1] -0.5  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
[24]  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
[47]  0.0  0.0  0.0  0.0
> round(ARMAtoMA(ar=c(1,-0.5),ma=-1,50),2) #causal (since coeffs \psi are converging to 0)
 [1]  0.00 -0.50 -0.50 -0.25  0.00  0.12  0.12  0.06  0.00 -0.03 -0.03 -0.02  0.00  0.01  0.01  0.00  0.00  0.00  0.00
[20]  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
[39]  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
> round(ARMAtoAR(ar=c(1,-0.5),ma=-1,50),2) #not invertible (since coeffs converge to 0.5)
 [1] 0.0 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5
[30] 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5
```

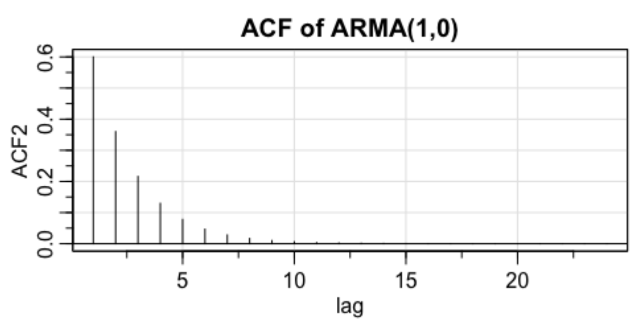
Problem 3: **Problem 4.3 pg 95, textbook.**

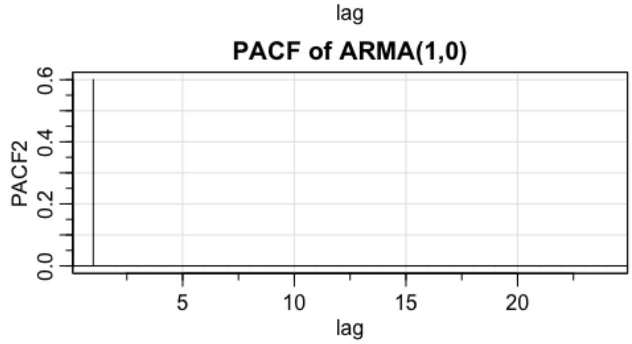
Solution: a) We see that both ACF and PACF of $ARMA(1,1)$ tails off, whereas ACF is tailing off with positive values, and PACF is tailing of with oscillating around 0. Moreover, it is not very clear **solely** from ACF and PACF how to pick p and q values for optimal $ARMA(p,q)$ model.



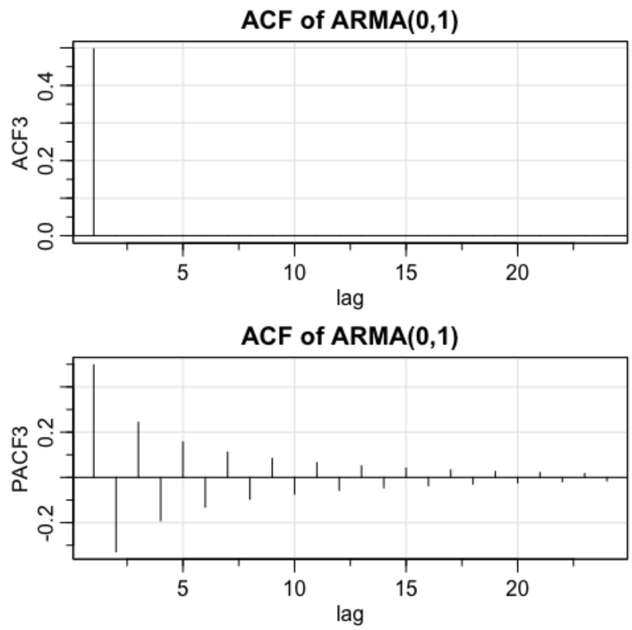


Similarly, ACF of $ARMA(1,0)$ is tailing of with positive values and PACF is cutting off at 1. Therefore, finding the order is simple and it is an order 1 AR model in this case.

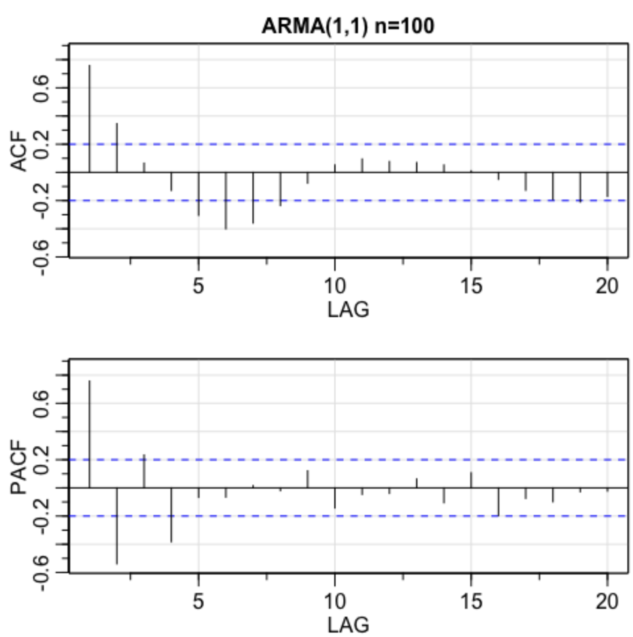




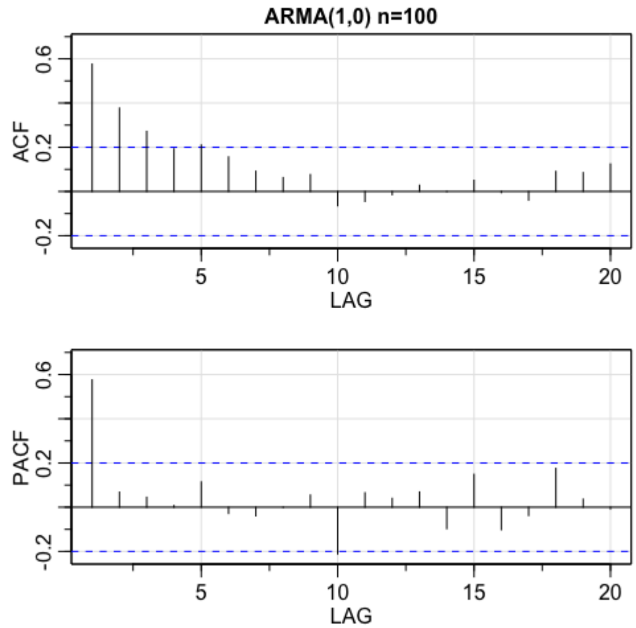
Lastly, PACF of $ARMA(0,1)$ is tailing of with oscillating values and ACF is cutting off at 1. Hence, finding the order is easy and it is an order 1 MA model.



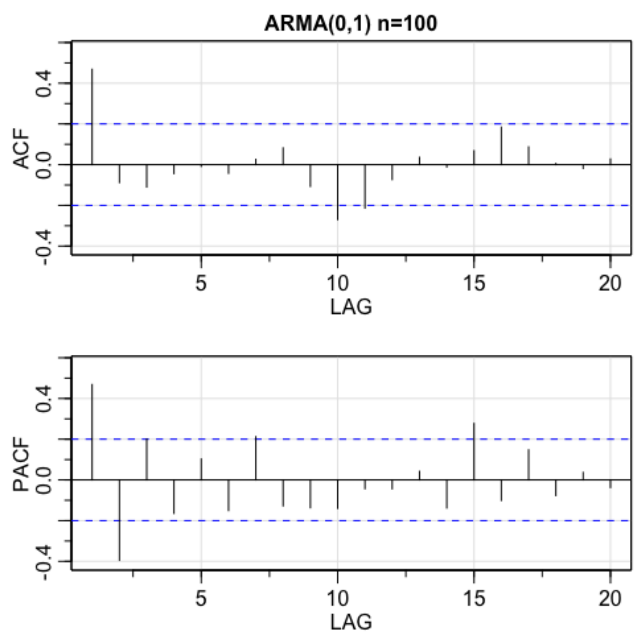
b) Similar to information given at **Table 4.1** in the book, we see that $ARMA(1,1)$ tails off in both sample ACF and PACF. The tailing off behaviour is harder to see compared to theoretical ACF and PACF. Again, it is not very clear **solely** from sample ACF and PACF how to pick p and q values for optimal $ARMA(p,q)$ model.



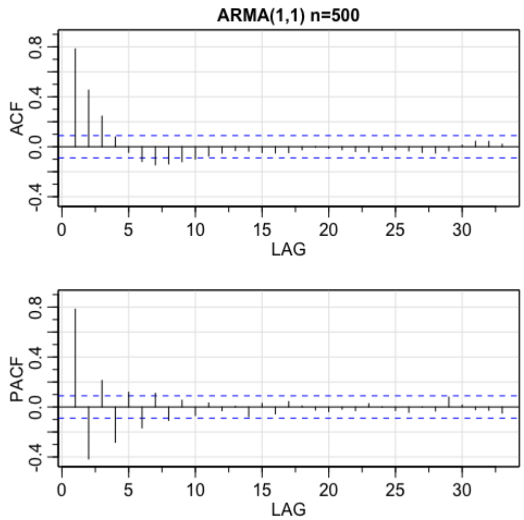
Similarly, we see that tailing off sample ACF and PACF cutting off at 1 tells us the model is an order 1 AR model, i.e, $ARMA(1,0)$.

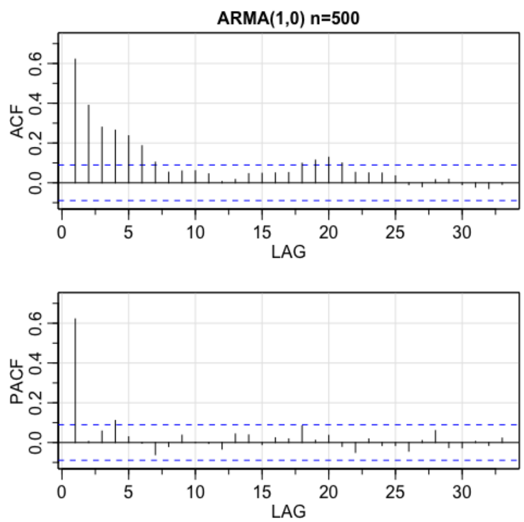


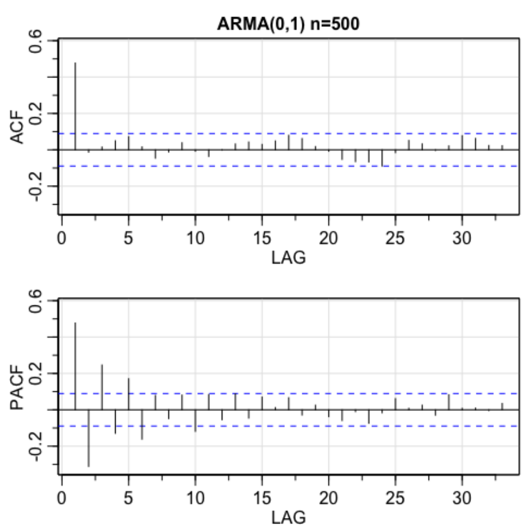
Lastly, tailing off PACF and ACF cutting off at 1 implies the model is an order 1 MA model, i.e, $ARMA(1,0)$.



c) The behaviour of sample ACF and sample PACF gets easier to read in the case of large sample size $n = 500$, in other words, sample ACF and PACF gets closer to theoretical ones as $n \to \infty$.


ARMA(1,1) n=500
ACF
LAG
PACF
LAG


ARMA(1,0) n=500
ACF
LAG
PACF
LAG


ARMA(0,1) n=500
ACF
LAG
PACF
LAG

Problem 4: ***BONUS*: Problem 4.10 pg 97, textbook.**

Solution: Let's assume an infinite story, $\{x_n, x_{n-1}, \ldots, x_1, x_0, x_{-1}, \ldots\}$, is available. Then, since

$$\psi_j\; w^n_{n-j} = w_{n-j} \text{ (known past)}, \quad w^n_{n+j+1} = E[w_{n+j+1}] = 0 \text{ for } j = 0, 1, 2, \ldots$$

we find that:

$$x^n_{n+m} = \sum_{j=0}^{\infty} \psi_j\; w^n_{m+n-j} = \psi_0\; w^n_{m+n} + \psi_1\; w^n_{m+n-1} + \cdots + \psi_m\; w^n_n + \psi_{m+1}\; w^n_{n-1} + \ldots$$

$$= \cancelto{0}{\psi_0\; w^n_{m+n}} + \cancelto{0}{\psi_1\; w^n_{m+n-1}} + \cdots + \psi_m\; w_n + \psi_{m+1}\; w_{n-1} + \cdots = \sum_{j=m}^{\infty} \psi_j\; w_{m+n-j}$$

Now, using this result we can find the mean squared prediction error as:

$$P^n_{n+m} = E[x_{n+m} - x^n_{n+m}]^2 = E\left[\sum_{j=0}^{\infty} \psi_j\; w_{m+n-j} - \sum_{j=m}^{\infty} \psi_j\; w_{m+n-j}\right]^2$$

$$= E\left[\sum_{j=0}^{m-1} \psi_j\; w_{m+n-j}\right]^2 = \sum_{j=0}^{m-1} \psi_j^2 E\left[\; w^2_{m+n-j}\right] = \sigma_w^2 \sum_{j=0}^{m-1} \psi_j^2$$

Problem 5: **Problem 5.2 pg 126, textbook.**

Solution: Plots of the GDP data and transformed GDP data are given below.

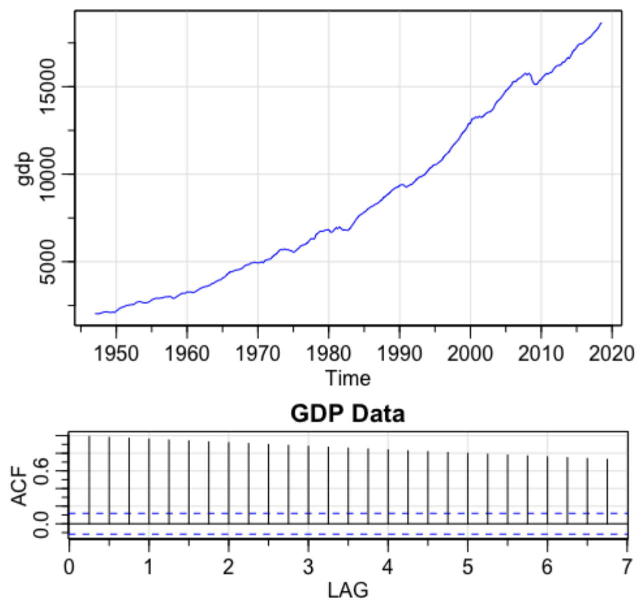


Figure 1: Original GDP Data

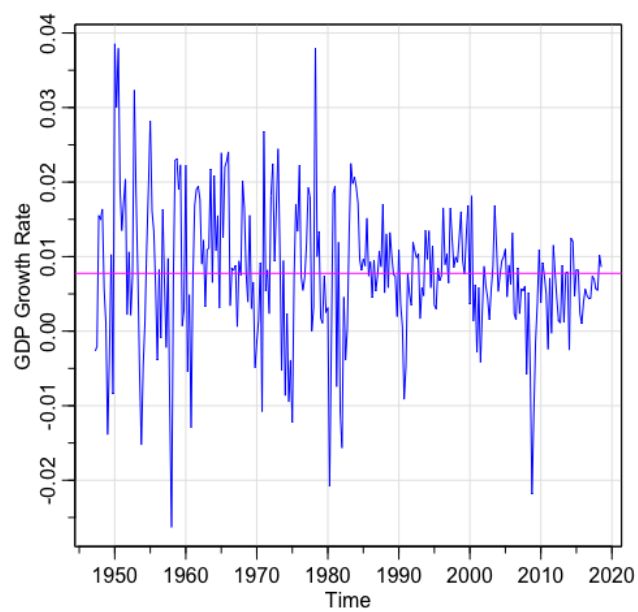


Figure 2: Transformed GDP Data

Transforming data yields, somewhat stationary model with non-zero mean. Now, let us plot sample ACF and PACFs of the transformed (diff(log(gdp))) data to come up with an appropriate model.

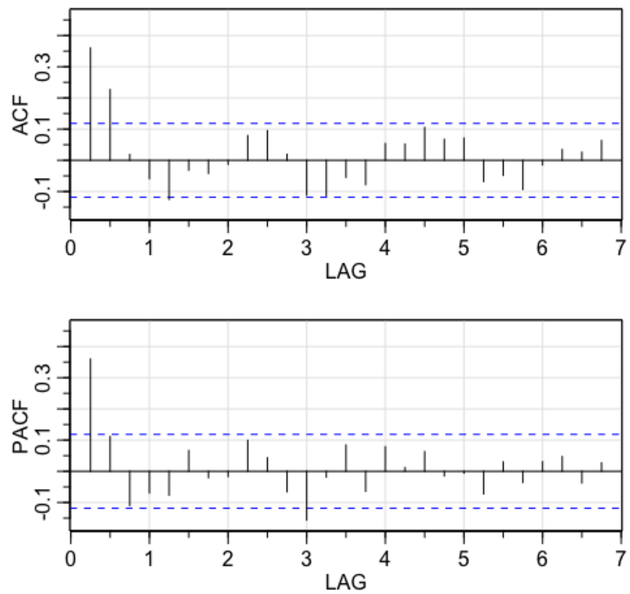


Figure 3: Transformed GDP Data

Looking at sample ACF and PACFs of diff(log(gdp)) data, I would suggest MA(2)=ARMA(0,2) (ACF is cutting off after 2) or AR(1)=ARMA(1,0) (PACF is cutting off after 1) models. Notice that, ARIMA(0,1,2) and ARIMA(1,1,0) of log(gdp) data gives the same results. Below, I presented the diagnostic plots of both of the models.

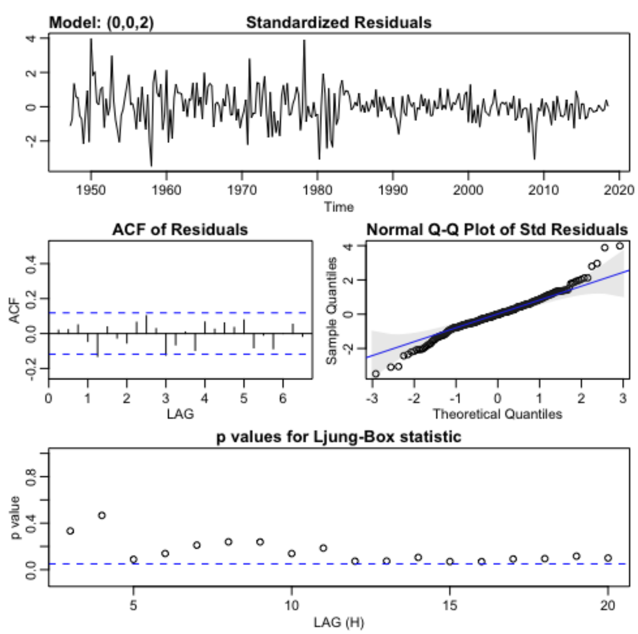


Figure 4: Diagnostics of MA(2) Model

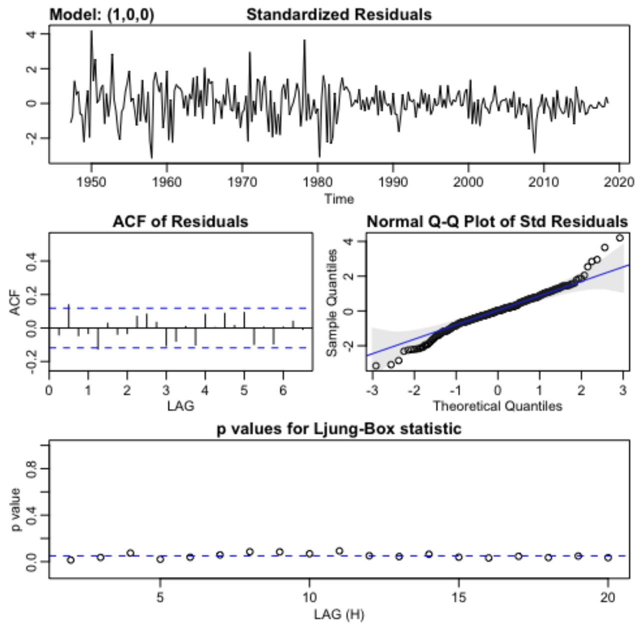


Figure 5: Diagnostics of AR(1) Model

Looking at the diagnostic plots MA(2) perform better, since most of the p-values are above the dashed line. Before getting a final model I present the AIC and BIC values of other possible fits ARMA(0,2+1),ARMA(1+1,0) and ARMA(1,2) in transformed data. Minimum values of each is given by bold color. Since adding more variables don't change the results significantly, I would pick MA(2). Notice, BIC of AR(1) is smaller; however, AIC of MA(2) is much better than AR(1).

Model	AIC	BIC
ARMA(0,2)	-6.636312	-6.58518
ARMA(1,0)	-6.625634	**-6.587284**
ARMA(0,3)	**-6.637412**	-6.573497
ARMA(2,0)	-6.631465	-6.580332
ARMA(1,2)	-6.63568	-6.571764

Problem 6: **Problem 5.3 pg 126, textbook.**

Solution: Plots of the Oil data and transformed oil data are given below.

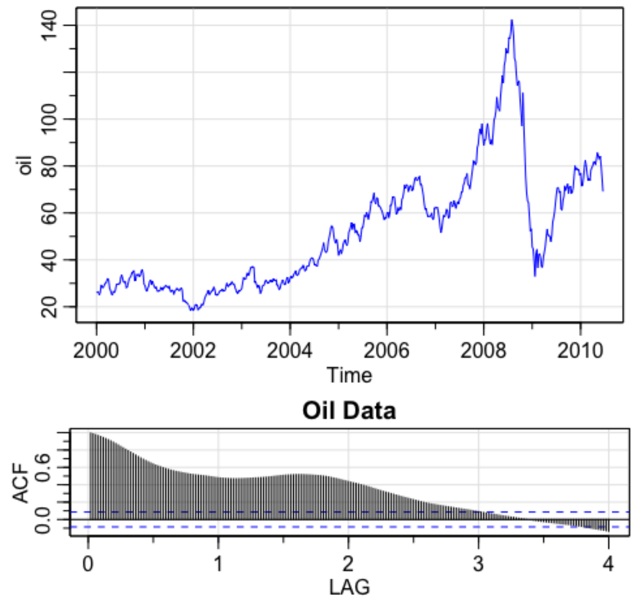


Figure 6: Original Oil Data

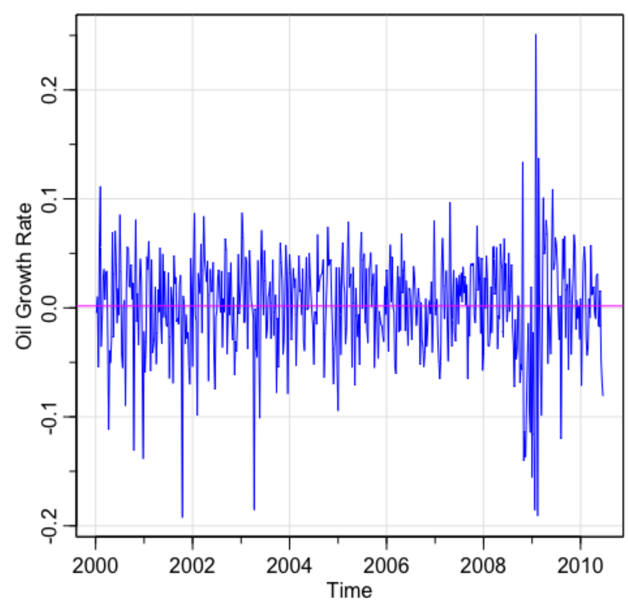


Figure 7: Transformed Oil Data

Transforming the data yields, a stationary process. Now, let us plot sample ACF and PACFs of the transformed (diff(log(oil))) data to come up with an appropriate model.

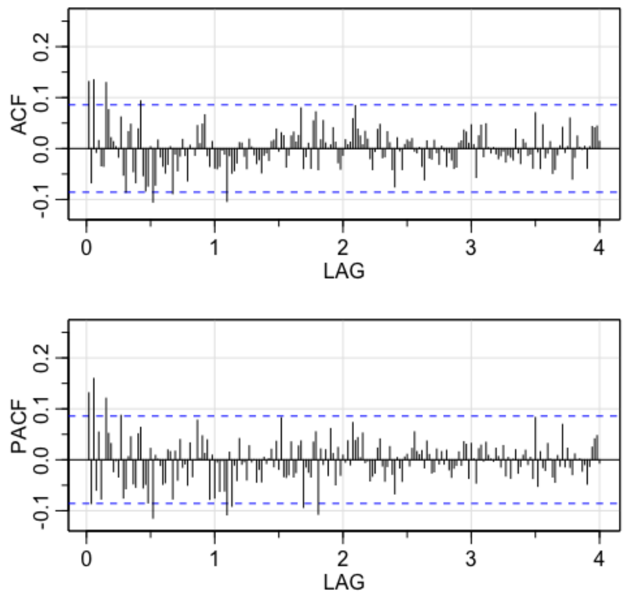


Figure 8: Transformed Oil Data

Looking at sample ACF and PACFs of diff(log(oil)) data, I would suggest MA(3)=ARMA(0,3) (ACF is cutting off after 3) or ARMA(1,1) (both ACF and PACF are tailing off) models. Notice that, ARIMA(0,1,3) and ARIMA(1,1,1) of log(oil) data gives the same results. Below, I presented the diagnostic plots of both of the models.

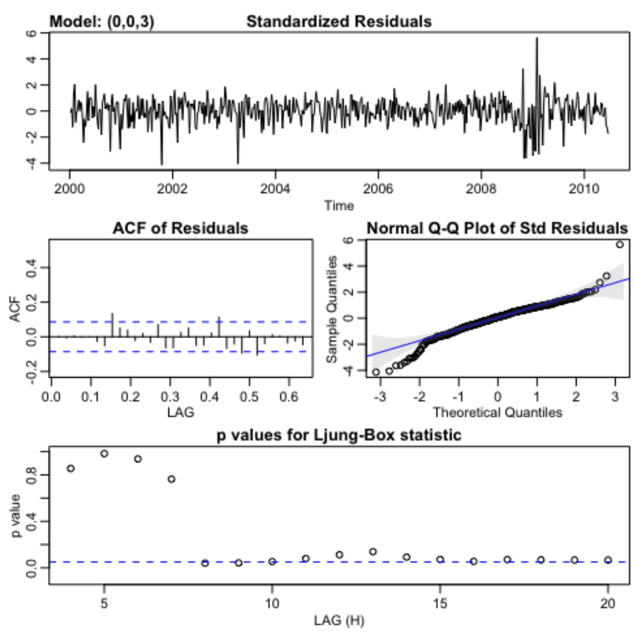


Figure 9: Diagnostics of MA(3) Model

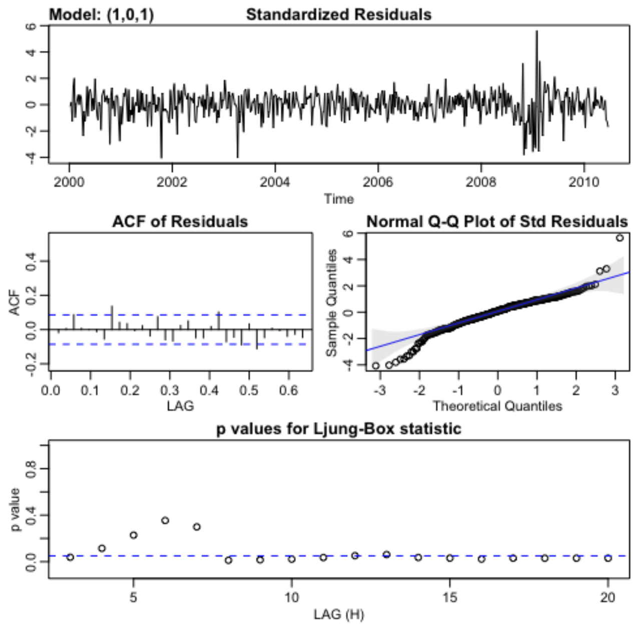


Figure 10: Diagnostics of ARMA(1,1) Model

Looking at the diagnostic plots they are pretty much similar, and most of the p-values are above the dashed line. Before getting a final model I present the AIC and BIC values of other possible fits ARMA(0,3+1),ARMA(1+1,1), ARMA(1,1+1), and ARMA(1+1,1+1) in transformed data. Minimum values of each is given by bold color. Since adding more variables don't change the results significantly, I would pick between MA(3) and ARMA(1,1). In this case, since difference between AIC is not significant and ARMA(1,1) is simpler, I would pick ARMA(1,1).

Model	AIC	BIC
ARMA(0,3)	**-3.318638**	-3.279125
ARMA(1,1)	-3.312109	**-3.280499**
ARMA(0,4)	-3.315108	-3.267693
ARMA(2,1)	-3.31002	-3.270508
ARMA(1,2)	-3.309568	-3.27005
ARMA(2,2)	-3.306659	-3.259244

Problem 7: **Problem 5.6 pg 126-127, textbook.**

Solution: Plots of the SO2 data and transformed SO2 data are given below.

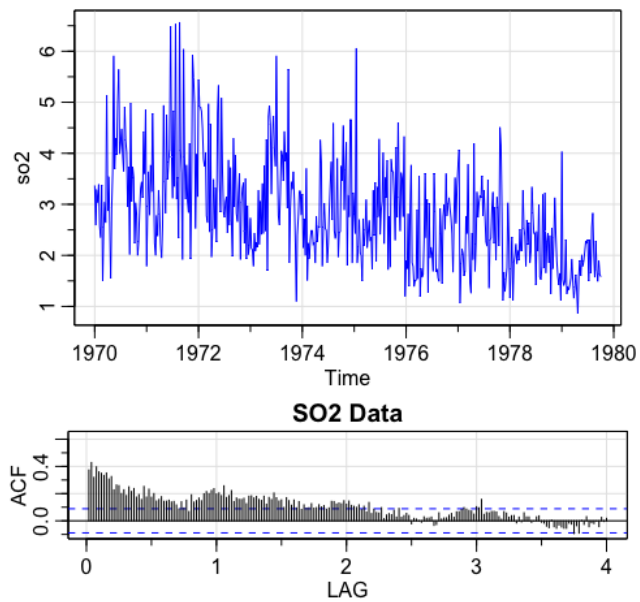


Figure 11: Original SO2 Data

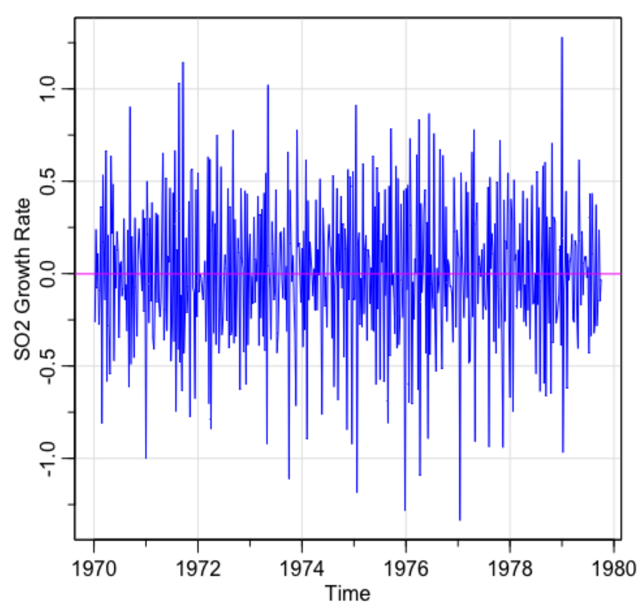


Figure 12: Transformed SO2 Data

Transforming the SO2 data yields, a stationary process with zero mean. Now, let us plot sample ACF and PACFs of the transformed (diff(log(so2))) data in order to come up with an appropriate model.

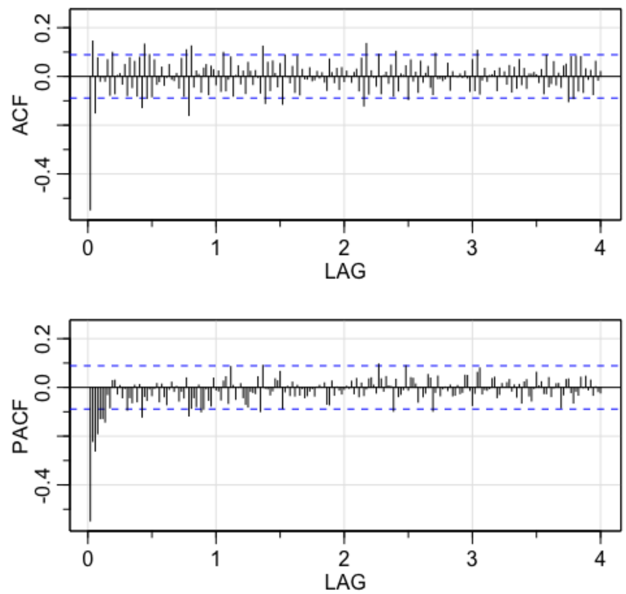


Figure 13: Transformed SO2 Data

Looking at sample ACF and PACFs of diff(log(so2)) data, I would suggest MA(1) (ACF is cutting off after 1) or ARMA(1,1) (both ACF and PACF are tailing off) models. Notice that, ARIMA(0,1,1) and ARIMA(1,1,1) of log(so2) data will give the same results. Below, I presented the diagnostic plots of both of the models.

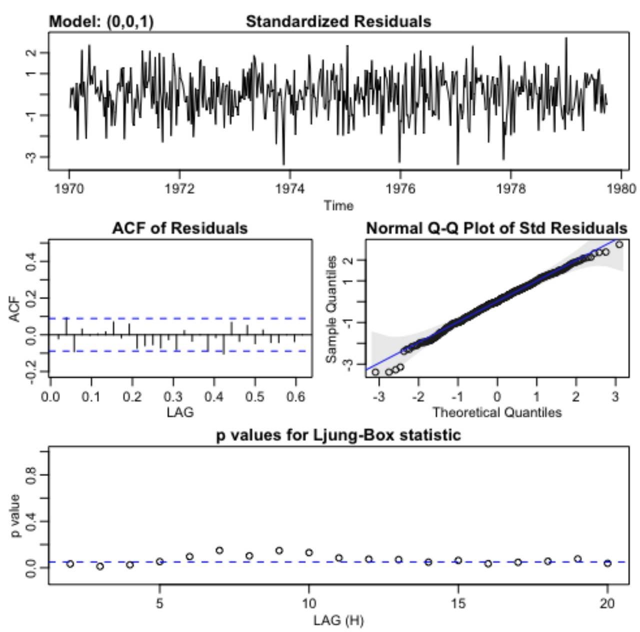


Figure 14: Diagnostics of MA(1) Model

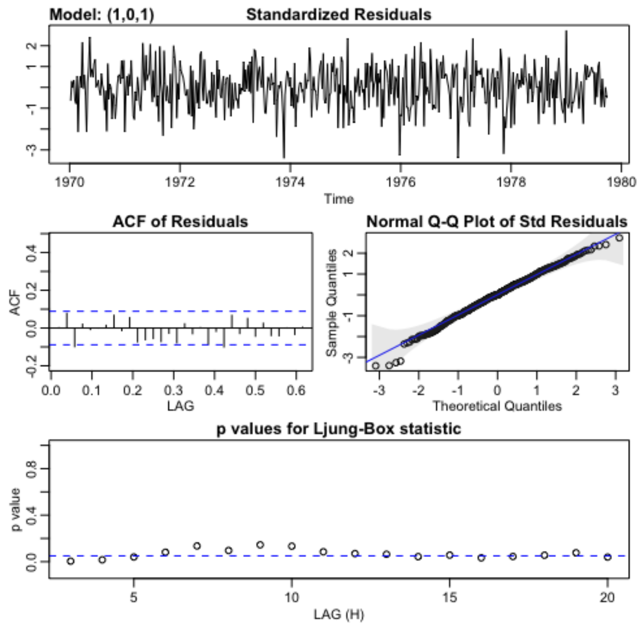


Figure 15: Diagnostics of ARMA(1,1) Model

Looking at the diagnostic plots they are I would say MA(3) is doing a better job even tough they are very close to each other. Before getting a final model I will now present the AIC and BIC values of other possible fits ARMA(0,2), and ARMA(0,3) in transformed data as well. Minimum values of each is given by bold color. Adding more variables yields slightly better in AIC of MA(3); however, BIC of MA(1) performs better. Since, MA(1) is much simpler model and there is not a significant difference between two of the models, I would pick MA(1).

Model	AIC	BIC
ARMA(0,1)	0.4989322	**0.523953**
ARMA(1,1)	0.5022039	0.5355649
ARMA(0,2)	0.5023514	0.5357125
ARMA(0,3)	**0.4949513**	0.5366526

Now, let us fit a SARIMA(0,0,1) model on diff(log(so2)) and forecast the future for four time periods. First, 95% prediction intervals,

$$x_{n+j}^{n} \pm 1.96 * se\big(x_{n+j}^{n}\big) \quad \text{for} \quad j = 1, 2, 3, 4)$$

are given by the table below:

Future	Prediction Interval
x_{n+1}^{n}	(-0.4811,0.7275)
x_{n+2}^{n}	(-0.7893,0.7873)
x_{n+3}^{n}	(-0.7893,0.7873)
x_{n+4}^{n}	(-0.7893,0.7873)

Then, I present the forecast plot of future four time periods above.

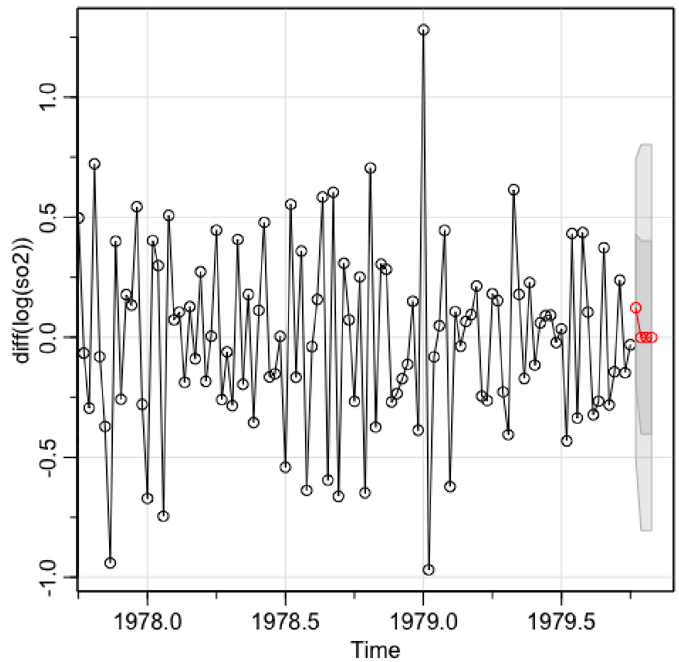


Figure 16: Four future forecast of diff(log(so2)) with MA(1) Model

Problem 8: **Problem 5.7 pg 127, textbook.**

Solution: First, plotting the AirPassengers data we see a clear increasing trend as well as an annual seasonality (12 months). On the other hand, due to increasing variances, I would first take the log of the data. Below, the plots of data and transformed $\nabla\nabla^{12}log(AirPassengers)$ data.

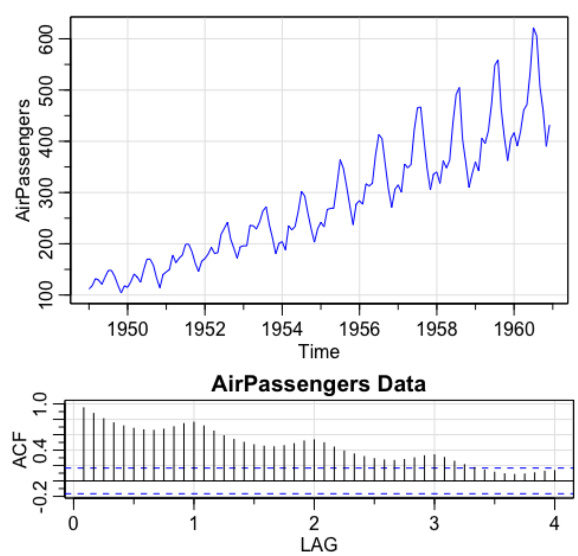


Figure 17: Original AirPassengers Data

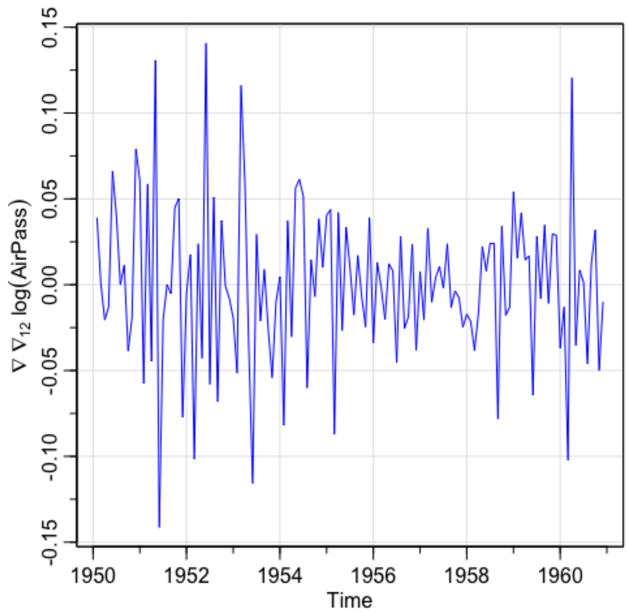


Figure 18: Transformed $\nabla\nabla^{12}log(AirPassengers)$ Data

Transforming the AirPassengers data yields a stationary process. Now, let us plot sample ACF and PACFs of the transformed $\nabla\nabla^{12}log(AirPassengers)$ data in order to come up with an appropriate model.

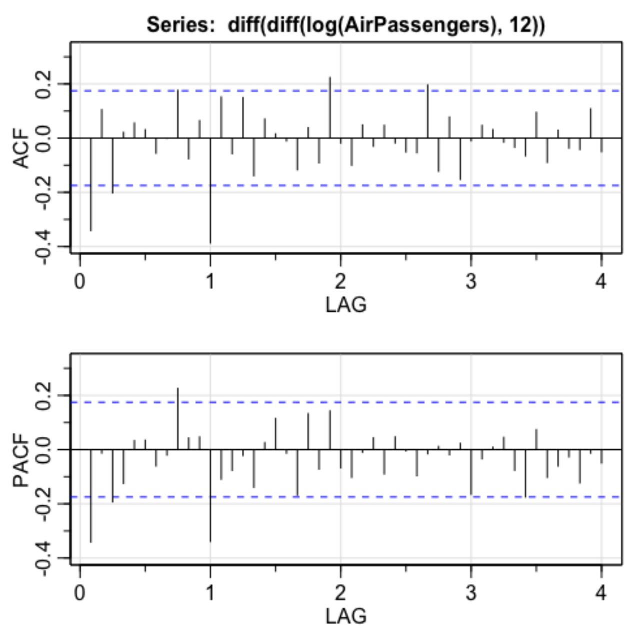


Figure 19: ACF and PACF of AirPassengers Data

Looking at sample ACF and PACFs of $\nabla\nabla^{12}log(AirPassengers)$ data, I would suggest $SARIMA(0,1,1) \times (0,1,0)_{12}$ and $SARIMA(0,1,1) \times (1,1,0)_{12}$ models. Below, I presented the diagnostic plots of both of the models.

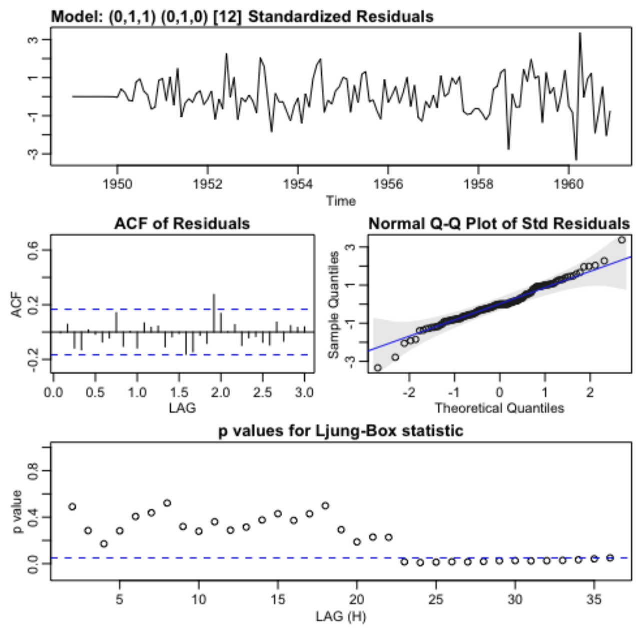


Figure 20: Diagnostics of $SARIMA(0,1,1) \times (0,1,0)_{12}$ Model

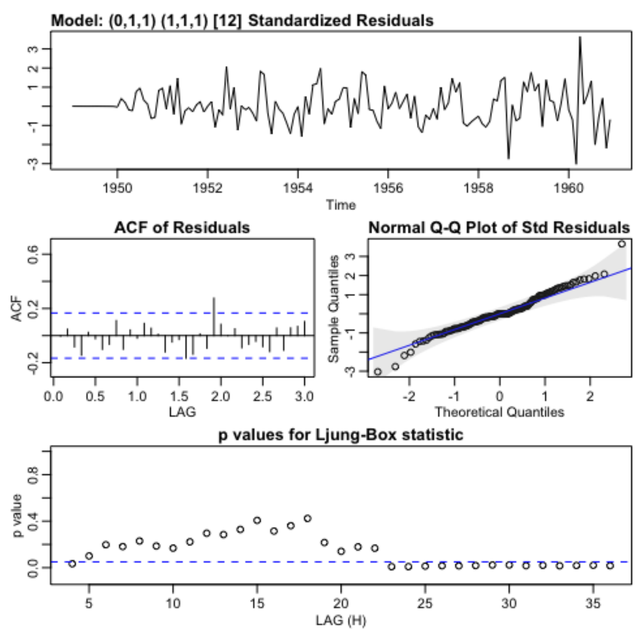


Figure 21: Diagnostics of $SARIMA(0,1,1) \times (1,1,0)_{12}$ Model

Looking at the diagnostic plots and comparing AIC, BIC values I would pick my final model as $SARIMA(0,1,1) \times (0,1,0)_{12}$ since it is much simpler model and doing better overall.

Model	AIC	BIC
$SARIMA(0,1,1) \times (0,1,0)_{12}$	7.1876	**7.228095**
$SARIMA(0,1,1) \times (1,1,0)_{12}$	**7.185411**	7.266402