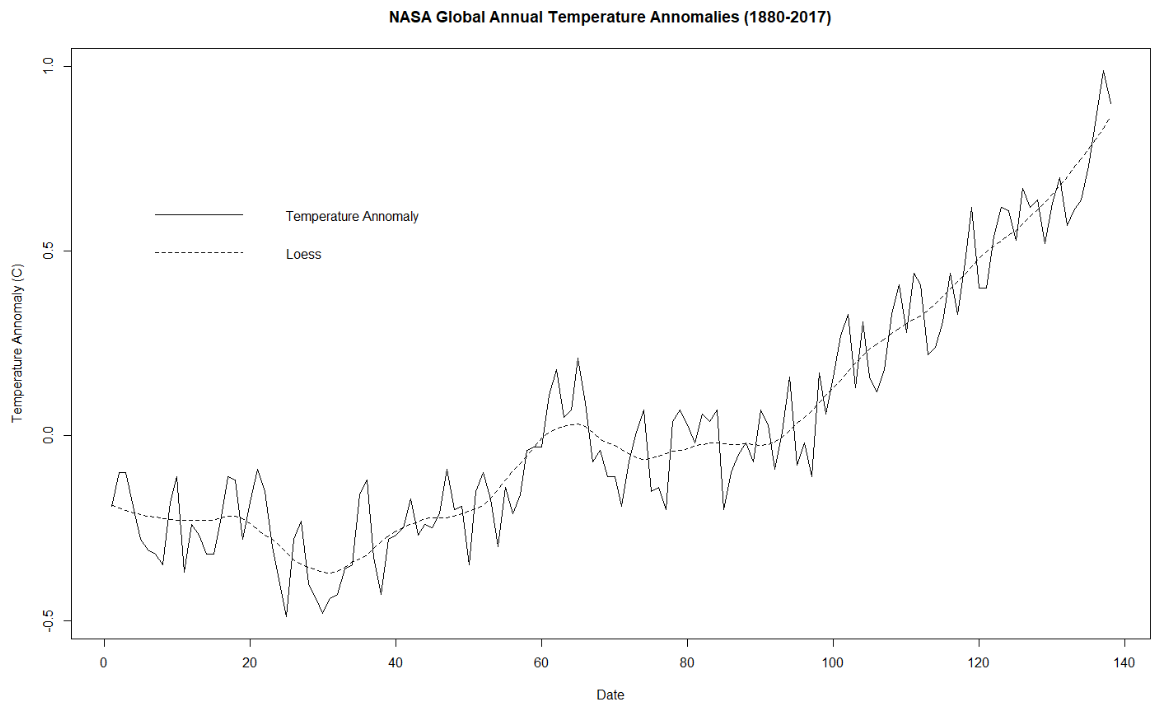
STA 137

Homework 4

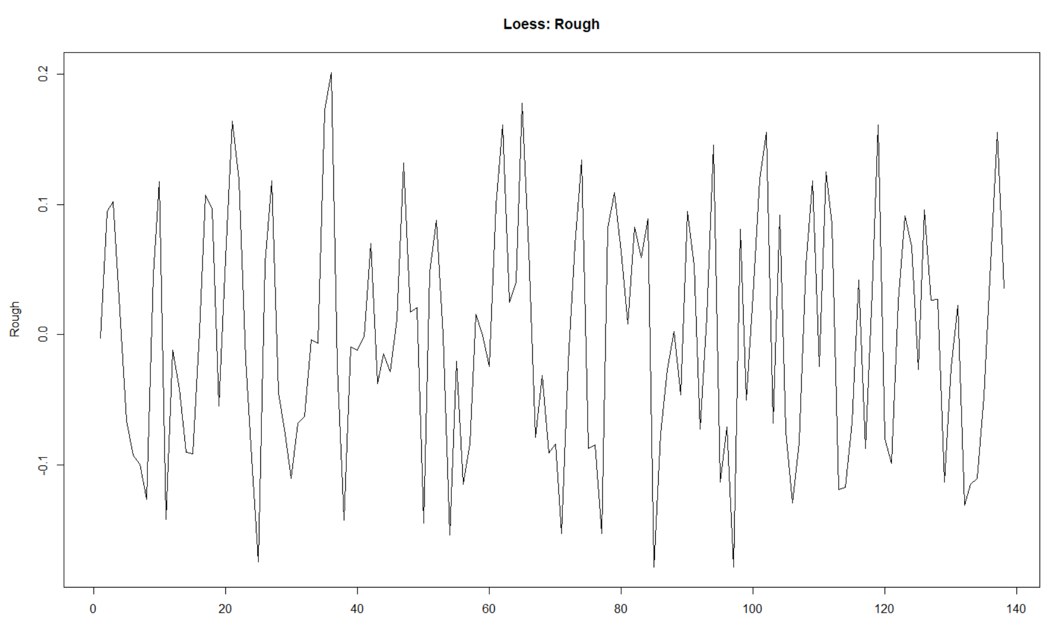
Prof. Burman

Jared Yu (914640091), Danli Zheng (915011728)
2-13-2019

1. A) A Loess model was fitted to the data using a span = 0.25. Below is a plot of the data along with the fitted loess line:

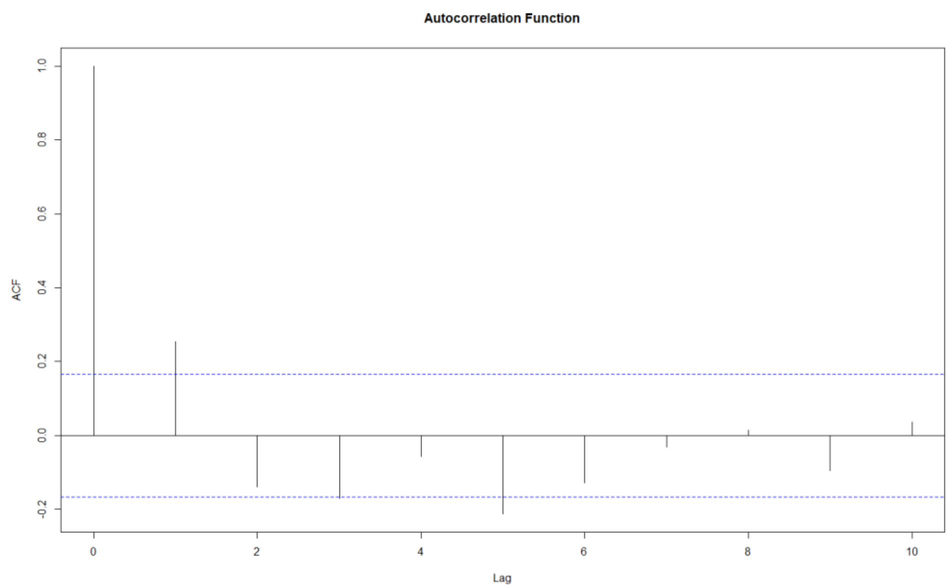


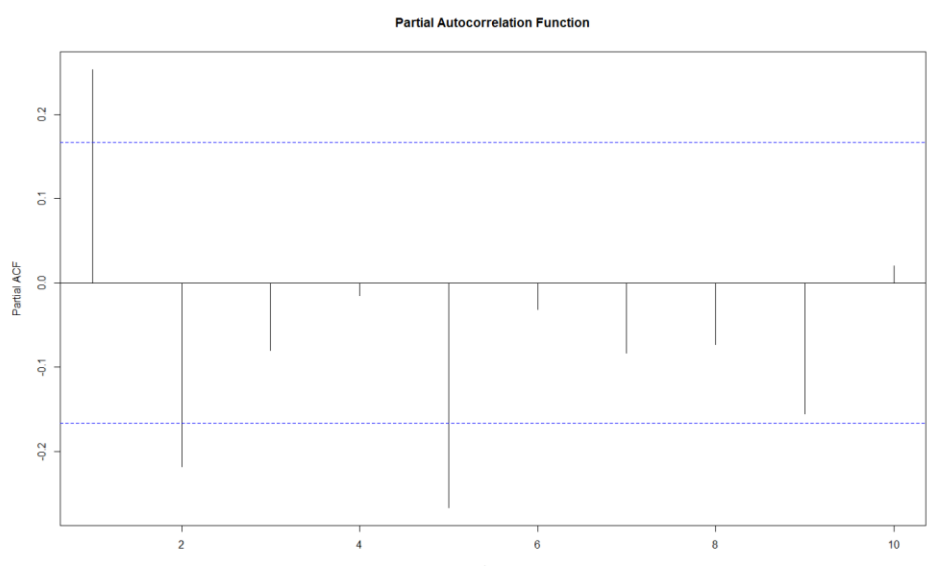
1. B) Below is an estimate of the rough part X_t plotted against time:



The data is modeled as $Y_t = m_t + X_t$, where the trend m_t is modeled with the Loess, and to find the rough, we subtract the trend from the observations Y_t. We then have an estimate of the rough $\hat{X}_t$. The rough looks quite balanced, there seems to be some possibly higher variance in the beginning, which narrows towards the end. The data seems to stay within -0.2 and 0.2. It is difficult to say whether any seasonality exists in the rough, although this component has not been considered in the modeling process. Also, the rough exhibits a typical pattern of going up and down through a range, but it is also difficult to say whether it is an AR or MA process just by looking at it.

1. C) Below is a plot of the ACF and PACF of the rough.





The two ACF and PACF plots don't show an extremely obvious type of model that has been seen before when plotting theoretical models. Here, there seems to be the possibility of a mix of AR and MA within the data. The ACF plot shows that there may be some significant correlation at lag 1 and possibly also at lag 5. This gives the idea that the data possibility has some $MA(1)$ or even $MA(5)$ connection, but the latter seems even less likely. The value at lag 1 is 0.253, while the value at lag 5 is -0.213. The PACF plot shows that there may be significant correlation at lag $1, 2$, and 5. The spike in the correlation at lag 5 is surprisingly large, suggesting that this lag is not insignificant. The value at lag 1 is 0.253, the value at lag 2 is -0.218, and the value at lag 5 is -0.267. The plot shows that the data may exhibit $AR(1), AR(2)$, and maybe $AR(5)$ patterns. To account for the combination of the patterns in both the ACF and PACF plots, it is possible that the data is $ARMA(2,1)$ or $ARMA(5,1)$.

1. D) The $AR(p)$ models for $p = 0, \cdots, 5$ were fitted. Their AICC values were also determined using the function given in the files on Canvas. Below is a table of the corresponding AICC values for each of the different models.

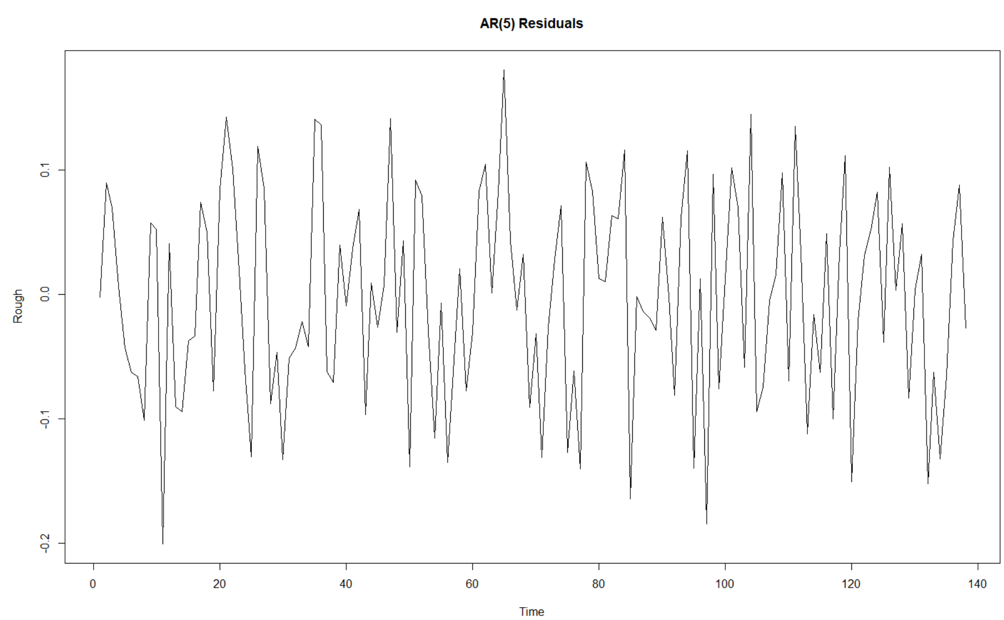
$AR(p)$	$AICC$
$AR(0)$	-268.1488
$AR(1)$	-275.2553
$AR(2)$	-279.9226
$AR(3)$	-278.7798
$AR(4)$	-276.6892
$AR(5)$	-284.5314

The value with the smallest AICC value is $AR(5)$, therefore, using the AICC criterion, we would choose $AR(5)$ out of the given $AR(p)$ models.

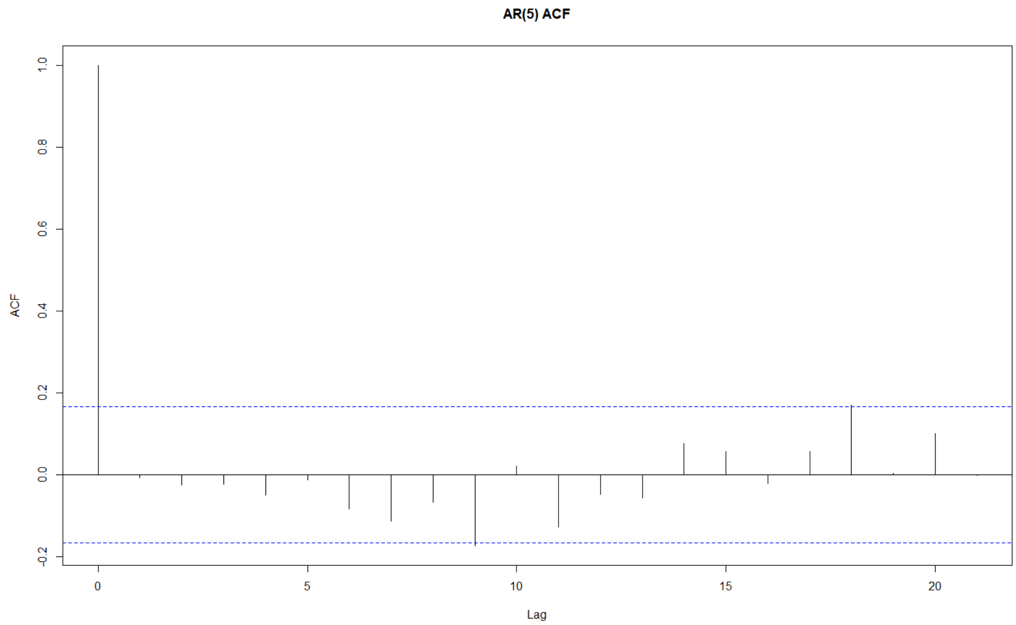
1. E) Below are the parameter estimates and their standard errors for the $AR(5)$ model:

	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\phi}_5$
$Coefficients$	0.2839	-0.2155	-0.1285	0.0583	-0.2652
$Standard\ error$	0.0817	0.0853	0.0866	0.0853	0.0821

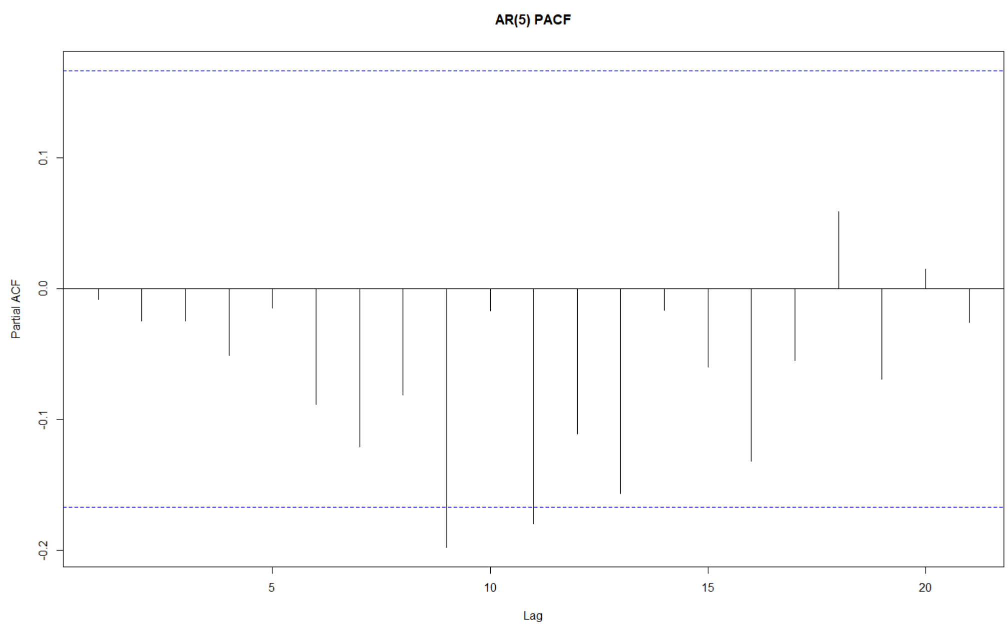
Below is a plot of the residuals for the $AR(5)$ model:



Below is a plot of the $AR(5)$ ACF:

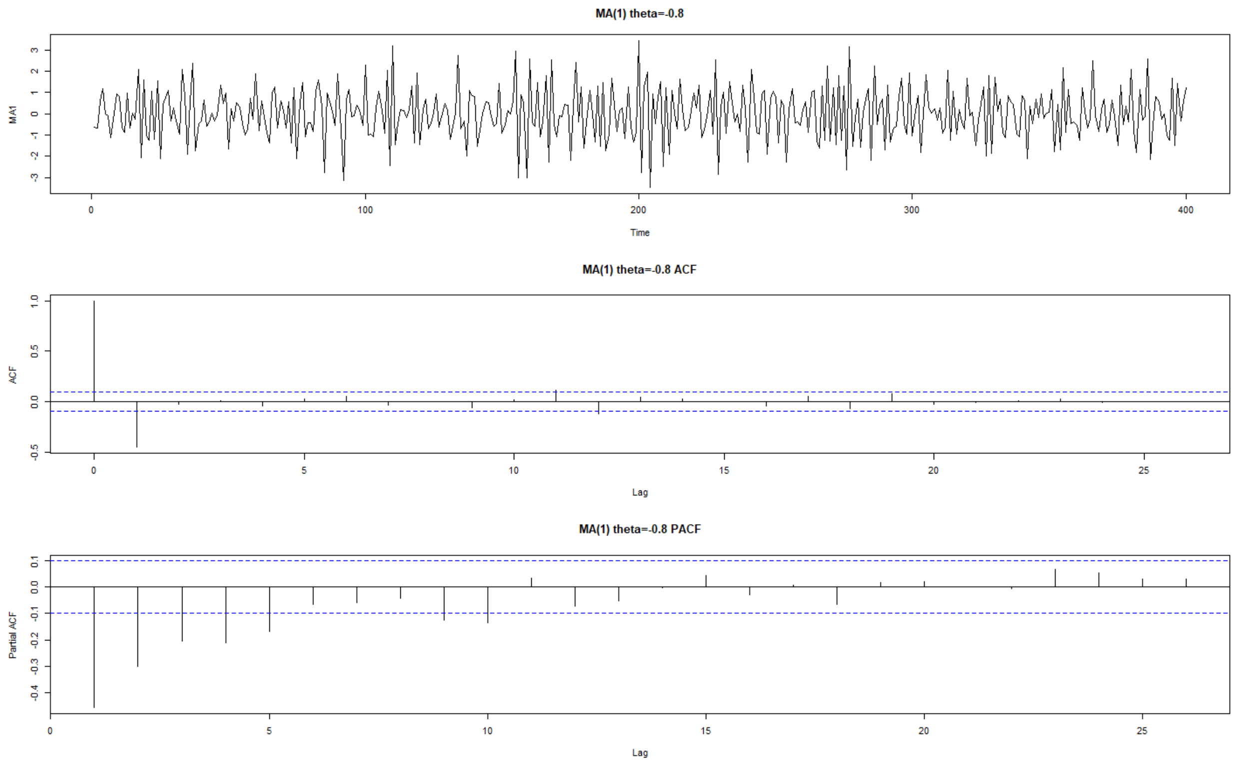


Below is a plot of the $AR(5)$ PACF:



The plot of the rough shows that it moves up and down within a range of -0.2 to 0.2. Looking at the ACF and PACF, there doesn't seem to be any significant lags. So, the residuals don't show any correlation, which is ideal. That means the residuals are like white noise, which means that the model is doing a good job.

2. A) The following models have mean zero, $n = 400$, and $\sigma^2 = 1$. They are plotted along with ACF and PACF models. Here, the first model is $MA(1), \theta = -0.8$

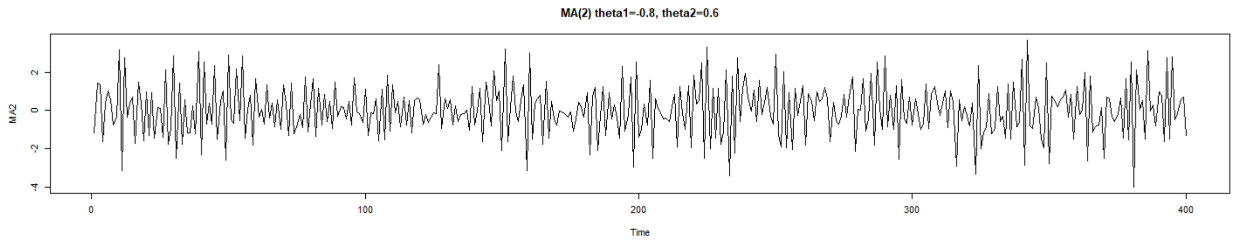


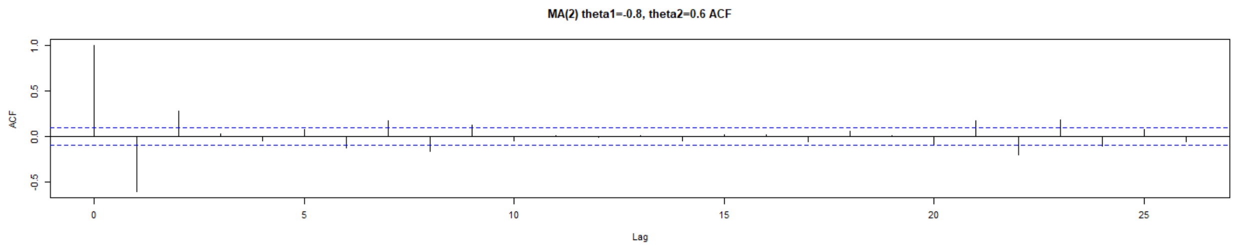
The first plot shows the $MA(1), \theta = -0.8$ data, which itself is difficult to analyze. The data looks quite scrunched together, which makes sense since it is only $MA(1)$, so the neighboring correlation between points may only last a short while. Also, the coefficient is negative, so points tend to move opposite of each other rather quickly.

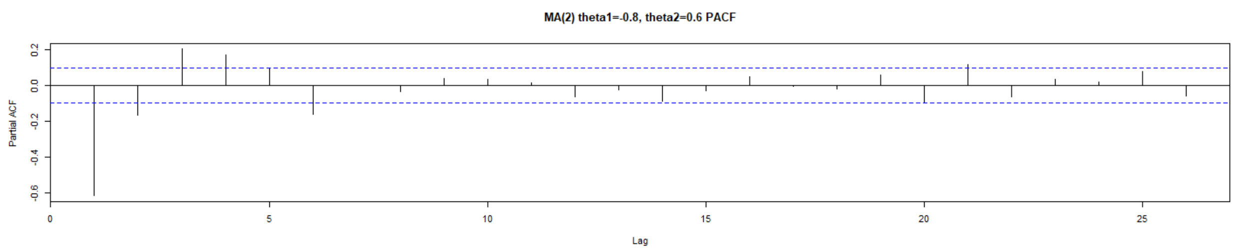
The second plot shows this pattern in the ACF, where at lag 1, there is a significant negative correlation, but afterwards there is nothing significant. The value at lag 1 is -0.456, while at other lags the value is much closer to 0.

The last plot shows the PACF, which shows the correlation starting high (although negative), and quickly dropping off. This makes sense, considering that the model itself is an MA model, rather than an AR model. In the case of the latter type of model, the pattern in the PACF would be more representative of the type of parameters that had been given.

2. B) The following model is $MA(2), \theta_1 = -0.8, \theta_2 = 0.6$. Below is the plot of the data along with the ACF and PACF.





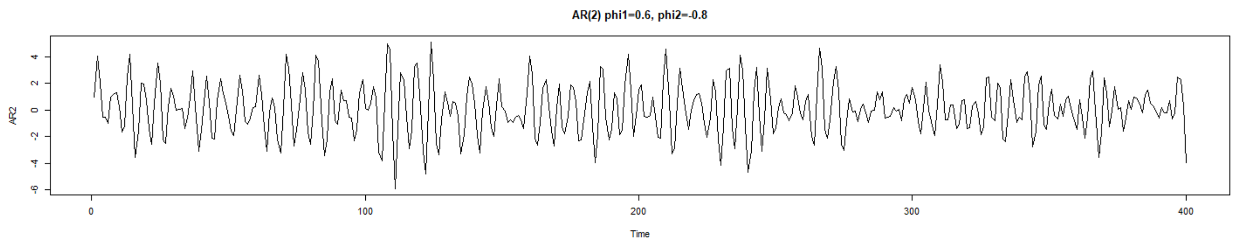


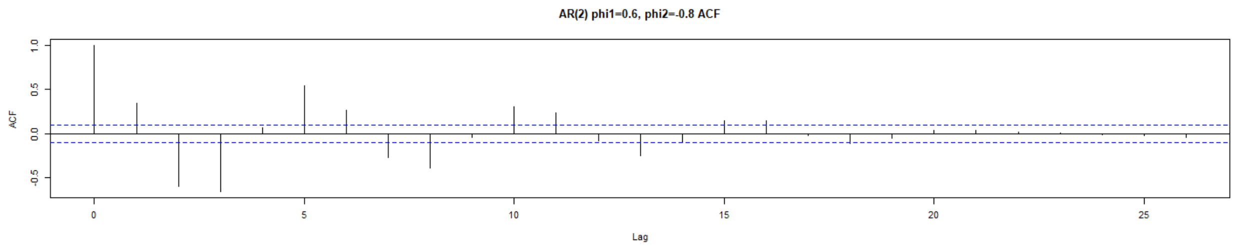
The first plot possible shows the $MA(2)$ data which appears somewhat different than the previous $MA(1)$ data. There seems to be narrower sections, and possibly even more scrunching in the pattern. This change makes sense since there is a negative correlation followed by a positive correlation. It is difficult to interpret this result in a logical way, other than noting the slight changes in the appearance of the data.

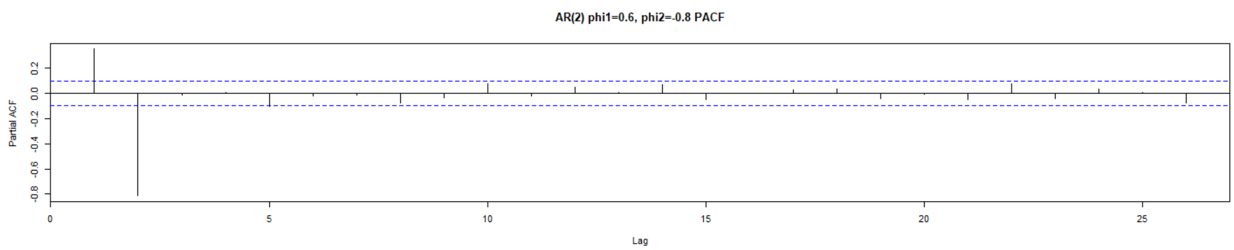
The second plot shows the ACF with significant values at lag 1 and 2. There are other spikes which extend beyond the confidence intervals, but they seem far less significant. The values at lag 1 and 2 are -0.614 and 0.272. These results make sense, considering that it is an ACF plot of an MA model, and the spikes in the correlation follow the pattern of the parameter.

The third plot shows a PACF of the data, and the pattern seems to exhibit some gyration starting at lag 1 and going until possibly lag 6. Interpreting this is difficult, since it is a PACF plot of an MA model. The pattern however of starting high and decreasing slowly is like what was seen before.

2. C) The following model is $AR(2), \phi_1 = 0.6, \phi_2 = -0.8$. Below is the plot of the data along with the ACF and PACF.





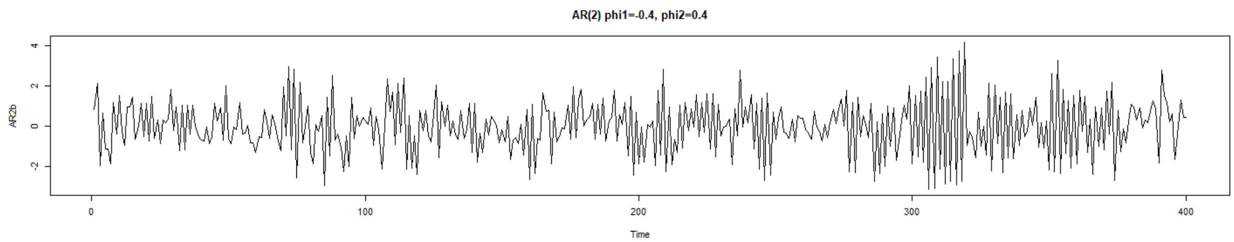


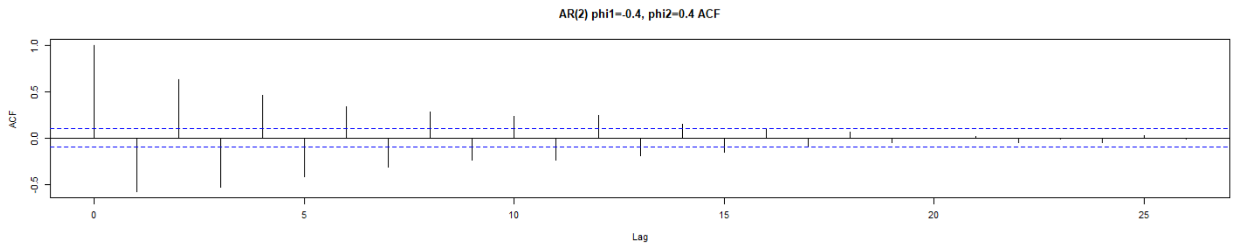
The first plot shows the plot of the $AR(2)$ data, and it exhibits a different pattern than what was seen previously with the MA models. The pattern appears much less erratic than before, and seems to cycle in a smoother pattern.

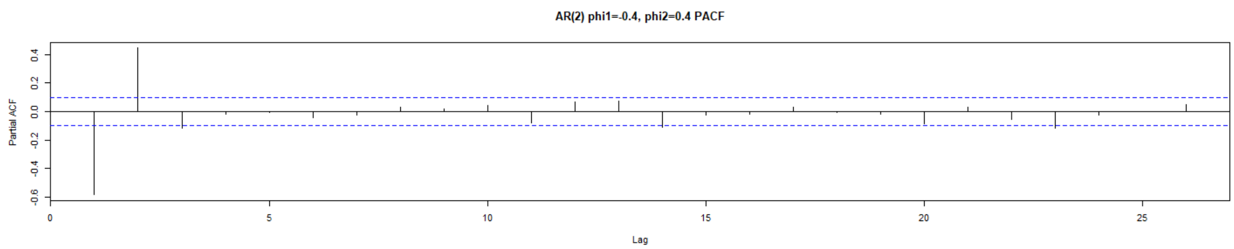
The second plot shows that the ACF plot will start high, and then gyrate around zero, before eventually diminishing until the points are all insignificant. This makes sense, given that the model itself is an AR model, and the ACF plot does not do an adequate job of describing the pattern for an AR model.

The third plot shows the PACF, which does a better job of describing the pattern of the $AR(2)$ data. It is evident in the PACF, that there is significant correlation at lag 1 and 2, which makes sense given that $p = 2$ in the $AR(p)$ model. The values at lag 1 and 2 are 0.347 and -0.593. The value at lag 2 is higher, which makes sense given that the coefficient at ϕ_2 is larger than at ϕ_1.

2. D) The following model is $AR(2), \phi_1 = -0.4, \phi_2 = 0.4$. Below is the plot of the data along with the ACF and PACF.





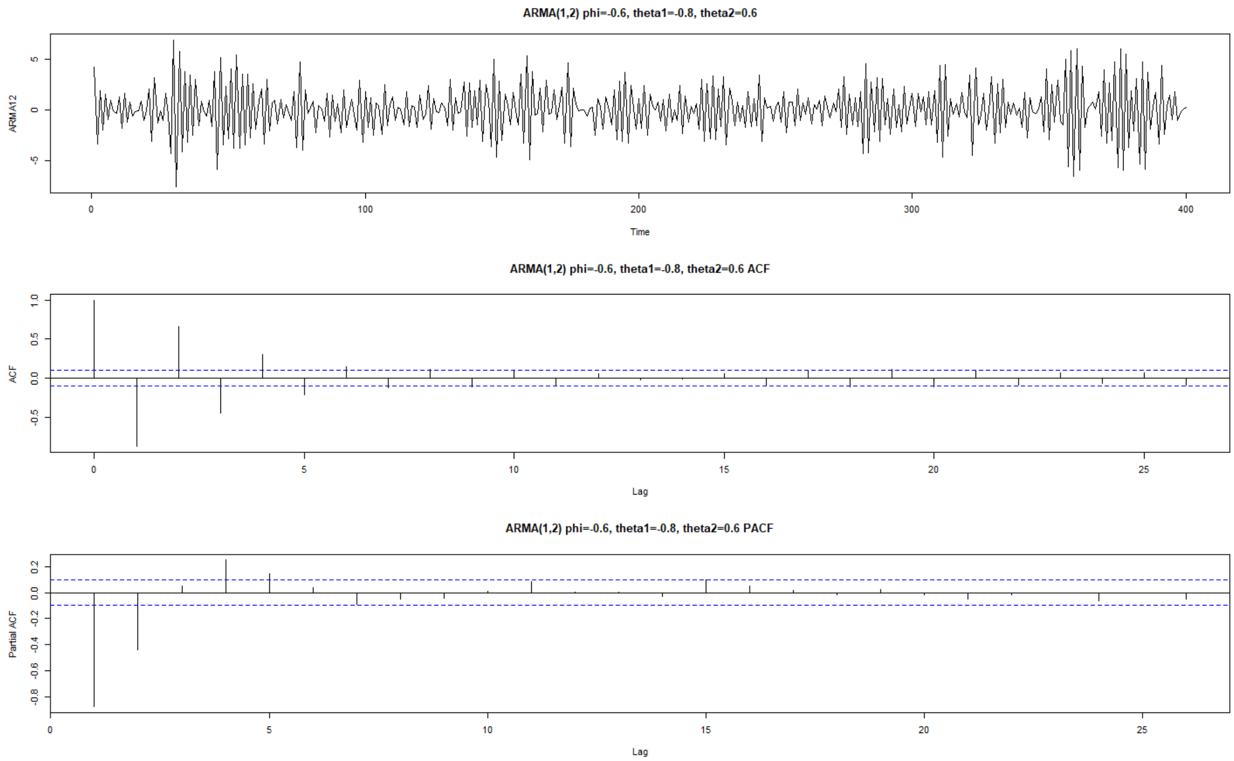


The first plot shows the $\mathrm{AR}(2)$ data, and the pattern is noticeably different from the previous $\mathrm{AR}(2)$ model. The variation of the movements is much smaller at first, and towards the end become larger temporarily. Before, there were much smoother cycles, and this is likely due to the first coefficient being negative in this case, where as previously it was positive. The negative neighboring coefficient can lead to this quick gyration with an apparently lower variance. However, it seems also possible that this variance can become quite large at times.

The second plot gyrates about zero before eventually diminishing to an insignificant level. This makes sense because it is an ACF plot, where the data itself follows an $\mathrm{AR}(2)$ model. This is like what happened in the previous model where an ACF plot was trying to describe an $\mathrm{AR}(2)$ sequence.

The third plot shows an PACF plot of the $\mathrm{AR}(2)$ model, with significant correlation at lag 1 and 2. The values at these lags are -0.583 and 0.444 respectively. The rest of the values do not appear significant. This result is expected, considering that the PACF is good at modeling AR models. Also, the size of the values at the lags are much closer than before. This is likely due to the coefficients being much closer in size, despite being of different signs (negative vs. positive).

2. E) The following model is $ARMA(1,2), \phi = 0.6, \theta_1 = -0.8, \theta_2 = 0.6$. Below is the plot of the data along with the ACF and PACF.



The first plot shows the $\mathrm{ARMA}(1,2)$ data, which appears slightly different from what was seen in the previous MA and AR models. Here, the data seems a bit more organized and less random. Before, the data seemed somewhat unstructured, while in this case, although there is still randomness, the pattern appears much neater and less scattered. This may be a result of the model itself being more complex than before. The ARMA model includes both an AR component and an MA component, therefore it could make sense that the model would seem more structured in nature than before.

The second plot shows a gyration around zero which was seen before in similar cases where the 'incorrect' ACF was used on an MA series. The spikes will start high and become smaller until it finally tails off. This is an expected outcome from an $ARMA$ series when plotted with an ACF plot.

The third plot shows a similar tailing off as in the second plot. However, it is not as apparent in this case. It does somewhat look as if there is correlation at 1 or 2 lags, but also it occurs again at 4 lags. Yet when looking closer, the same gyration becomes apparent, as the spikes eventually tail off and become insignificant. A possible reason for this 'less-obvious' tailing off is that the series is $ARMA(1,2)$, with only one coefficient in the AR series, and two coefficients in the MA component of the series. Therefore, the tailing off is more apparent in the second plot rather than the first plot.

3. A) Given that X_t is a stationary sequence with mean μ, autocovariance γ, and that $\rho(1) = \phi$, we are first to show that the best linear predictor of X_t using X_{t-1} is $X_t^{(f)} = \mu + \phi(X_{t-1} - \mu)$:
The best linear predictor is $l(X_{t-1}) = aX_{t-1} + b$ such that $E\big(X_t - l(X_{t-1})\big)^2$ is minimized.

$$f(a,b) = E(X_t - aX_{t-1} - b)^2$$

$$\frac{\partial f}{\partial a} = E[-2X_{t-1}(X_t - aX_{t-1} - b)]$$

$$= -2\{E(X_{t-1}X_t) - aE((X_{t-1})^2) - bE(X_{t-1})\}$$

$$= -2\{Cov(X_{t-1}, X_t) + E(X_{t-1})E(X_t) - a\left(Var(X_{t-1}) + \big(E(X_{t-1})\big)^2\right) - b\mu\}$$

$$= -2\{\gamma(1) + \mu^2 - a(\gamma(0) + \mu^2) - b\mu\}$$

$$\frac{\partial f}{\partial b} = E[-2(X_t - aX_{t-1} - b)]$$

$$= -2\{E(X_t) - aE(X_{t-1}) - b\}$$

$$= -2(\mu - a\mu - b)$$

Set $\frac{\partial f}{\partial a} = 0$,

$$a(\gamma(0) + \mu^2) + b\mu = \gamma(1) + \mu^2$$

Let this be equation (1)

Set $\frac{\partial f}{\partial b} = 0$,

$$a\mu + b = \mu$$

Let this be equation (2)

Let us do, equation (1) – equation (2)$\times$ μ

$$a\gamma(0) = \gamma(1)$$

$$a = \frac{\gamma(1)}{\gamma(0)} = \rho(1)$$

Then, let us place a into equation (2),

$$b = \mu - a\mu$$

$$b = \mu - \rho(1)\mu$$

$$= \mu(1 - \rho(1))$$

So, the best linear predictor of X_t using X_{t-1} is,

$$X_t^{(f)} = aX_{t-1} + b$$

$$= \rho(1)X_{t-1} + \mu(1 - \rho(1))$$

$$= \phi X_{t-1} + \mu(1-\phi)$$
$$= \mu + \phi(X_{t-1} - \mu)$$

3. B) We would next like to show that the best linear predictor of X_{t-2} using X_{t-1} is $X_{t-2}^{(b)} = \mu + \phi(X_{t-1} - \mu)$

The best linear predictor is $l(X_{t-1}) = aX_{t-1} + b$ such that $E\left(X_{t-2} - l(X_{t-2})\right)^2$ is minimized.

$$f(a,b) = E\left(X_{t-2} - l(X_{t-2})\right)^2$$
$$= E(X_{t-2} - aX_{t-1} - b)^2$$
$$\frac{\partial f}{\partial a} = E(-2X_{t-1}(X_{t-2} - aX_{t-1} - b))$$
$$= -2E(X_{t-1}X_{t-2}) + 2aE((X_{t-1})^2) + 2bE(X_{t-1})$$
$$= -2\left(Cov(X_{t-1}, X_{t-2}) + E(X_{t-1})E(X_{t-2})\right) + 2a\left(Var(X_{t-1}) + \left(E(X_{t-1})\right)^2\right) + 2b\mu$$
$$= -2(\gamma(1) + \mu^2) + 2a(\gamma(0) + \mu^2) + 2b\mu$$
$$= -2\gamma(1) - 2\mu^2 + 2a\gamma(0) + 2a\mu^2 + 2b\mu$$
$$\frac{\partial f}{\partial b} = E\left(-2(X_{t-2} - aX_{t-1} - b)\right)$$
$$= -2E(X_{t-2}) + 2aE(X_{t-1}) + 2b$$
$$= -2\mu + 2a\mu + 2b$$

Set $\frac{\partial f}{\partial a} = 0$,

$$a\gamma(0) + a\mu^2 + b\mu = \gamma(1) + \mu^2$$
$$a(\gamma(0) + \mu^2) + b\mu = \gamma(1) + \mu^2$$

Let this be equation (1)

Set $\frac{\partial f}{\partial b} = 0$,

$$a\mu + b = \mu$$
$$b = \mu(1-a)$$

Let this be equation (2), now plugging equation (2) into equation (1),

$$a(\gamma(0) + \mu^2) + \mu^2(1-a) = \gamma(1) + \mu^2$$
$$a\gamma(0) + a\mu^2 + \mu^2 - a\mu^2 = \gamma(1) + \mu^2$$
$$a\gamma(0) = \gamma(1)$$
$$a = \frac{\gamma(1)}{\gamma(0)} = \rho(1)$$

Now plugging the a into equation (2),

$$b = \mu(1 - \rho(1))$$

So, the best linear predictor of X_{t-2} using X_{t-1} is,

$$X_{t-2}^{(b)} = aX_{t-1} + b$$
$$= \rho(1)X_{t-1} + \mu\big(1 - \rho(1)\big)$$
$$= \phi X_{t-1} + \mu(1 - \phi)$$
$$= \mu + \phi(X_{t-1} - \mu)$$

3. C) We would like to show that $E\left(X_t - X_t^{(f)}\right)^2 = (1 + \phi^2)\gamma(0) - 2\phi\gamma(1)$,

   It was previously shown that $X_t^{(f)}$ was the best linear unbiased predictor, and so the following results are a consequence of that,

$$E(X_t) = \mu$$

Also,

$$E\left(X_t^{(f)}\right) = E\big(\mu + \phi(X_{t-1} - \mu)\big)$$
$$= \mu + \phi(E(X_{t-1}) - \mu) = \mu$$

Therefore,

$$E\left(X_t - X_t^{(f)}\right) = E(X_t) - E\left(X_t^{(f)}\right) = 0$$

So, using the formula $Var(X) = E(X^2) - \big(E(X)\big)^2$, we can rewrite the following as,

$$E\{\left(X_t - X_t^{(f)}\right)^2\} = Var(X_t - X_t^{(f)})$$
$$= Var(X_t) + Var\left(X_t^{(f)}\right) - 2Cov\left(X_t, X_t^{(f)}\right)$$
$$= Var(X_t) + Var\big(\mu + \phi(X_{t-1} - \mu)\big) - 2Cov\big(X_t, \mu + \phi(X_{t-1} - \mu)\big)$$
$$= Var(X_t) + \phi^2 Var(X_{t-1}) - 2\phi Cov(X_t, X_{t-1})$$
$$= \gamma(0) + \phi^2\gamma(0) - 2\phi\gamma(1)$$
$$= (1 + \phi^2)\gamma(0) - 2\phi\gamma(1)$$

3. D) We would like to show that $E\left(X_t - X_{t-2}^{(b)}\right)^2 = (1 + \phi^2)\gamma(0) - 2\phi\gamma(1)$,

   It was previously shown that $X_{t-2}^{(b)}$ was the best linear unbiased predictor, and so the following results are a consequence of that,

$$E(X_t) = \mu$$

Also,

$$E\left(X_{t-2}^{(b)}\right) = E\big(\mu + \phi(X_{t-1} - \mu)\big)$$
$$= \mu + \phi(E(X_{t-1}) - \mu) = \mu$$

Therefore,

$$E\left(X_t - X_{t-2}^{(b)}\right) = E(X_t) - E\left(X_{t-2}^{(b)}\right) = 0$$

So, using the formula $Var(X) = E(X^2) - \left(E(X)\right)^2$, we can rewrite the following as,

$$E\{\left(X_{t-2} - X_{t-2}^{(b)}\right)^2\} = Var(X_{t-2} - X_{t-2}^{(b)})$$

$$= Var\left(X_{t-2} - \mu - \phi(X_{t-1} - \mu)\right)$$

$$= Var(X_{t-2}) + \phi^2 Var(X_{t-1}) - 2\phi Cov(X_{t-2}, X_{t-1})$$

$$= \gamma(0) + \phi^2\gamma(0) - 2\phi\gamma(1)$$

$$= (1 + \phi^2)\gamma(0) - 2\phi\gamma(1)$$

3. E) We would like to show that $Cov\left(X_t - X_t^{(f)}, X_{t-2} - X_{t-2}^{(b)}\right) = \gamma(2) - 2\phi\gamma(1) + \phi^2\gamma(0)$
   We will show that below,

$$Cov\left(X_t - X_t^{(f)}, X_{t-2} - X_{t-2}^{(b)}\right) = Cov\left(X_t - \left(\mu + \phi(X_{t-1} - \mu)\right), X_{t-2} - \left(\mu + \phi(X_{t-1} - \mu)\right)\right)$$

$$= Cov(X_t - \phi X_{t-1}, X_{t-2} - \phi X_{t-1})$$

$$= Cov(X_t, X_{t-2}) - \phi Cov(X_t, X_{t-1}) - \phi Cov(X_{t-1}, X_{t-2}) + \phi^2 Cov(X_{t-1}, X_{t-1})$$

$$= \gamma(2) - 2\phi\gamma(1) + \phi^2\gamma(0)$$

3. F) We would like to show that the partial correlation between X_t and X_{t-2} given X_{t-1} is given by $\frac{[\rho(2) - 2\phi\rho(1) + \phi^2]}{[1 + \phi^2 - 2\phi\rho(1)]}$. We will use the correlation formula to do so, it is as follows,

$$Corr\left(X_t - X_t^{(f)}, X_{t-2}\right) = \frac{Cov\left(X_t - X_t^{(f)}, X_{t-2} - X_{t-2}^{(b)}\right)}{\sqrt{Var\left(X_t - X_t^{(f)}\right) Var\left(X_{t-2} - X_{t-2}^{(b)}\right)}}$$

$$= \frac{\gamma(2) - 2\phi\gamma(1) + \phi^2\gamma(0)}{\sqrt{(1 + \phi^2)\gamma(0) - 2\phi\gamma(1)}\sqrt{(1 + \phi^2)\gamma(0) - 2\phi\gamma(1)}}$$

$$= \frac{(\gamma(2) - 2\phi\gamma(1) + \phi^2\gamma(0)/\gamma(0)}{((1 + \phi^2)\gamma(0) - 2\phi\gamma(1))/\gamma(0)}$$

$$= \frac{\rho(2) - 2\phi\rho(1) + \phi^2}{1 + \phi^2 - 2\phi\rho(1)}$$