Homework 5

Brian Detweiler

March 2 , 2017

1. For each of the following, identify it as an ARIMA model. That is, find the values of p, d, and q and the values of the parameters (ϕ's and θ's). Recall that by definition ARMA(p, q) models must be stationary and invertible.

(a) $Y_t = 0.6Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$

This appears to be an ARMA(2, 2), with $\phi_1 = 0.6$ and $\phi_2 = 0.4$, $\theta_1 = -0.5$ and $\theta_2 = 0.25$.

We must verify the assumptions that it is stationary and invertible.

$$
\begin{aligned}
\phi_1 + \phi_2 &= 0.6 + 0.4 = 1.0 \not< 1.0 \\
\phi_2 - \phi_1 &= 0.4 - 0.6 = -0.2 < 1 \\
|\phi_2| &= 0.4 < 1
\end{aligned}
$$

Here the first constraint is violated, so we transform this to the first difference, $\nabla_t = Y_t - Y_{t-1}$ \

$$
\begin{aligned}
Y_t &= 0.6Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2} \\
Y_t - Y_{t-1} &= (0.6Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}) - Y_{t-1} \\
\nabla Y_t &= 0.6Y_{t-1} - Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2} \\
&= -0.4Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}
\end{aligned}
$$

Letting $W_t = \nabla Y_t$, we have

$$
W_t = -0.4W_{t-1} + e_t - 0.5e_{t-1} + 0.25e_{t-2}
$$

Now this appears to be an ARMA(1, 2) with

$$
\begin{aligned}
\phi_1 &= -0.4 \\
\theta_1 &= -0.5 \\
\theta_2 &= 0.25
\end{aligned}
$$

Checking constraints, we have

$$
\begin{aligned}
\phi_1 + \phi_2 &= -0.4 + 0 = -1.0 < 1 \\
\phi_2 - \phi_1 &= 0 - (-0.4) = 0.4 < 1 \\
|\phi_2| &= 0 < 1
\end{aligned}
$$

Likewise, checking for invertibility, we have

$$\begin{aligned}\theta_1 + \theta_2 &= -0.5 + 0.25 = -0.25 < 1 \\ \theta_2 - \theta_1 &= 0.25 - (-0.5) = 0.75 < 1 \\ |\theta_2| &= 0.25 < 1\end{aligned}$$

The constraints are satisfied, and thus the first difference W_t is ARMA(1, 2). Therefore, Y_t is ARIMA(1, 1, 2).

(b) $Y_t = 2Y_{t-1} - Y_{t-2} + e_t$

Verifying the assumptions that it is stationary and invertible,

$$\begin{aligned}\phi_1 + \phi_2 &= 2 + (-1) = 1.0 \not< 1.0 \\ \phi_2 - \phi_1 &= (-1) - 2 = -3 < 1 \\ |\phi_2| &= 1 \not< 1\end{aligned}$$

Since the assumptions are violated, this is not stationary as an AR(2) model.

We can actually rewrite this as ∇Y_t, and we have

$$\begin{aligned}Y_t &= Y_{t-1} + Y_{t-1} - Y_{t-2} + e_t \\ Y_t - Y_{t-1} &= Y_{t-1} - Y_{t-2} + e_t \\ \nabla Y_t &= Y_{t-1} - Y_{t-2} + e_t\end{aligned}$$

Verifying the assumptions for stationary, we have

$$\begin{aligned}\phi_1 + \phi_2 &= 1 + (-1) = 0 < 1.0 \\ \phi_2 - \phi_1 &= 1 - (-1) = 2 \not< 1 \\ |\phi_2| &= 1 \not< 1\end{aligned}$$

So this is still not stationary.

Now we look at the second difference,

$$\begin{aligned}\nabla^2 Y_t &= \nabla(\nabla Y_t) \\ W_t &= \nabla(Y_t - Y_{t-1}) = Y_t - 2Y_{t-1} + Y_{t-2}\end{aligned}$$

Now, since $W_t = Y_t - 2Y_{t-1} + Y_{t-2} = e_t$, the second difference is a white noise process. Thus, it is an IMA(2, 0).

(c) $Y_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$

This seems to be an ARMA(2, 1) with $\phi_1 = 0.5, \phi_2 = -0.5$ and $\theta_1 = -0.1$.

The conditions for stationary hold,

$$\begin{aligned} \phi_1 + \phi_2 &= 0.5 + (-0.5) = 0 < 1.0 \\ \phi_2 - \phi_1 &= -0.5 + 0.5 = 0 < 1 \\ |\phi_2| &= 0.5 < 1 \end{aligned}$$

And since $|\theta_1| = 0.1 < 1$, then it is also invertible.

■

2. For each ARIMA model described in Question 1, find the numerical values of $\psi_0, \psi_1, \psi_2, \psi_3, \psi_4$ and a recurrence relation for $\psi_k, k > 4$.

(a) $Y_t = 0.6Y_{t-1} + 0.4Y_{t-2} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$

We saw that this ARIMA(1, 1, 2) could be represented as an ARMA(1, 2) as

$$W_t = -0.4W_{t-1} + e_t - 0.5e_{t-1} + 0.25e_{t-2}$$

```r
AR <- c(-0.4)
MA <- c(0.5, -0.25)
m <- 4
ARMAtoMA(ar = AR, ma = MA, lag.max = m)
```

```
## [1]  0.1000 -0.2900  0.1160 -0.0464
```

$$\begin{aligned}
\psi_0 &= 1 \\
\psi_1 - \phi_1\psi_0 &= \theta_1 + \theta_2 \\
\psi_1 &= \phi_1\psi_0 + \theta_1 + \theta_2 \\
\psi_1 &= (-0.4)(1) + (-0.5) + 0.25
\end{aligned}$$

This one I'm not quite sure on...

(b) $Y_t = 2Y_{t-1} - Y_{t-2} + e_t$

This was shown to be an IMA(2, 0).

By equation 4.1.1.

$$Y_t = e_t + \psi_1 e_{t-1} + \psi_2 e_{t-2} + \dots$$

But since we only have $W_t = e_t$, then we only have $\psi_0 = 1$, and $\psi_j = 0, j = 1, 2, 3, 4$

(c) $Y_t = 0.5Y_{t-1} - 0.5Y_{t-2} + e_t - 0.1e_{t-1}$

Recall this was shown to be an ARMA(2, 1) with $\phi_1 = 0.5, \phi_2 = -0.5$ and $\theta_1 = -0.1$.

$$
\begin{aligned}
\psi_0 &= 1 \\
\psi_1 - \phi_1\psi_0 &= \theta_1 \\
\psi_1 &= \phi_1\psi_0 + \theta_1 \\
&= 0.5(1) + 0.1 = 0.6 \\
\psi_2 &= \phi_1\psi_1 + \phi_2\psi_0 \\
&= 0.5(0.6) + (-0.5)(1) = -0.2 \\
\psi_3 &= \phi_1\psi_2 + \phi_1\psi_1 \\
&= 0.5(-0.2) + (-0.5)(0.6) = (-0.1) + (-0.3) = -0.4 \\
\psi_4 &= \phi_1\psi_3 + \phi_1\psi_2 \\
&= 0.5(-0.4) + (-0.5)(-0.2) = (-0.2) + 0.1 = -0.1
\end{aligned}
$$

Checking our answers,

```
AR <- c(0.5, -0.5)
MA <- c(0.1)
ARMAtoMA(ar = AR, ma = MA, lag.max = 4)
```

```
## [1]  0.6 -0.2 -0.4 -0.1
```

■

3. Consider a stationary process $\{Y_t\}$. Show that if $\rho_1 < 0.5$ then ∇Y_t has a larger variance than Y_t.

We will show that $Var(\nabla Y_t) > Var(Y_t)$.

Since $\{Y_t\}$ is stationary, $Var(Y_t) = \gamma_0 = \sigma^2$ is a constant.

We have $Var(\nabla Y_t) = Var(Y_t - Y_{t-1})$.

By the properties of variance and letting $k = 1$,

$$\begin{aligned}
Var(\nabla Y_t) &= Var(Y_t) + Var(Y_{t-k}) - 2Cov(Y_t, Y_{t-k}) \\
&= \gamma_0 + \gamma_0 - 2\gamma_k \\
&= 2\gamma_0 - 2\gamma_k \\
&= 2(\gamma_0 - \gamma_k) \\
&= 2(\gamma_0 - \gamma_k)
\end{aligned}$$

Since $\rho_1 = \frac{\gamma_k}{\gamma_0}$, we have

$$\begin{aligned}
Var(\nabla Y_t) &= 2(\gamma_0 - \frac{\gamma_0}{\gamma_0}\gamma_k) \\
&= 2(\gamma_0 - \rho_k \gamma_0) \\
&= 2\gamma_0(1 - \rho_k)
\end{aligned}$$

Thus, for $\rho_k < 0.5$ for $k = 1$, then $2(1-\rho_k) > 1$, and hence when multiplied by γ_0, is larger than $Var(Y_t) = \gamma_0$.

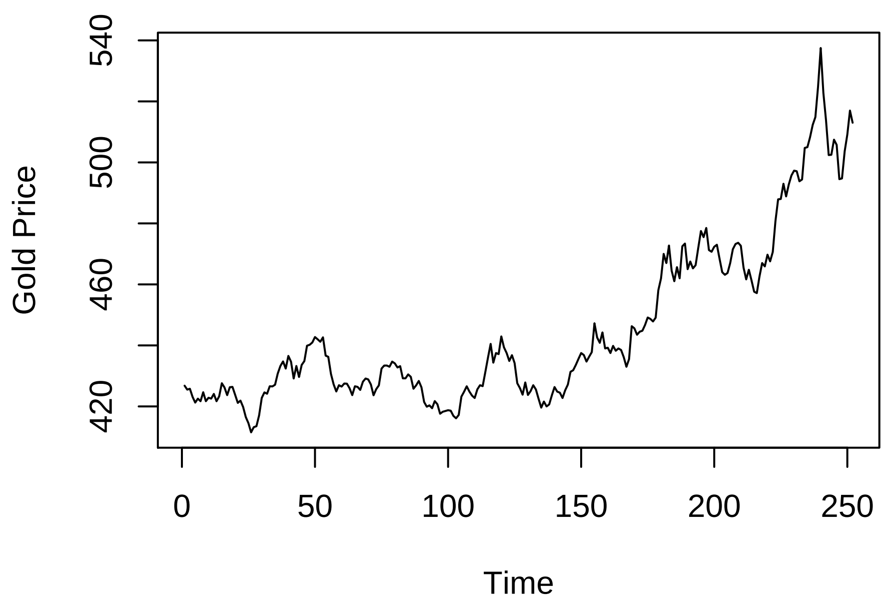
$\blacksquare$

4. The data set `gold` from the `TSA` library contains the daily price of gold for 252 trading days in 2005.

```r
data(gold)
```

(a) Construct a time series plot of the price of gold Y_t. What are the interesting features of this process?

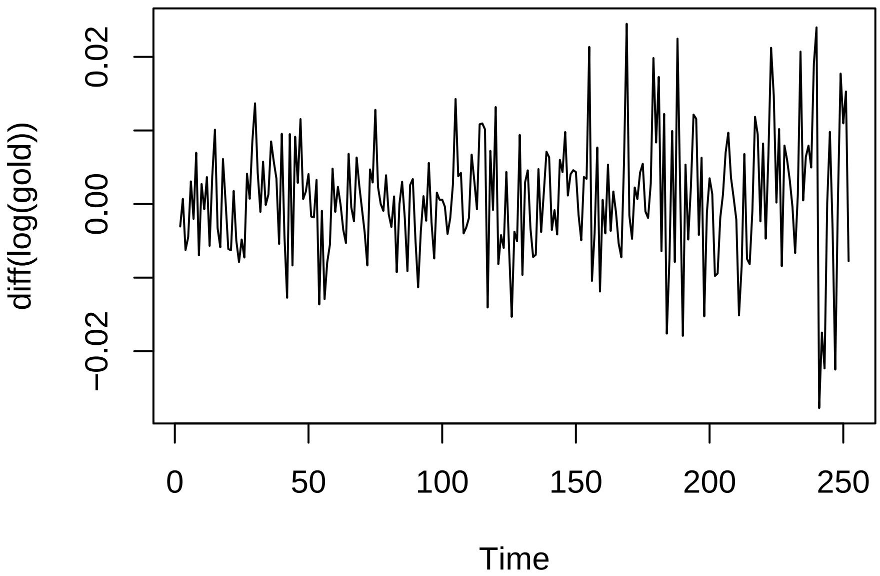
```r
par(cex=1)
plot(gold, ylab="Gold Price", pch=".")
```



The price of gold does not seem to be based on a deterministic trend, as we can see it begins to increase in variance after 150 days.

(b) Let $W_t = \nabla(\ln Y_t)$, the differences of the logarithms. Construct a time series plot of W_t. Does it look stationary?

```r
plot(diff(log(gold)))
```



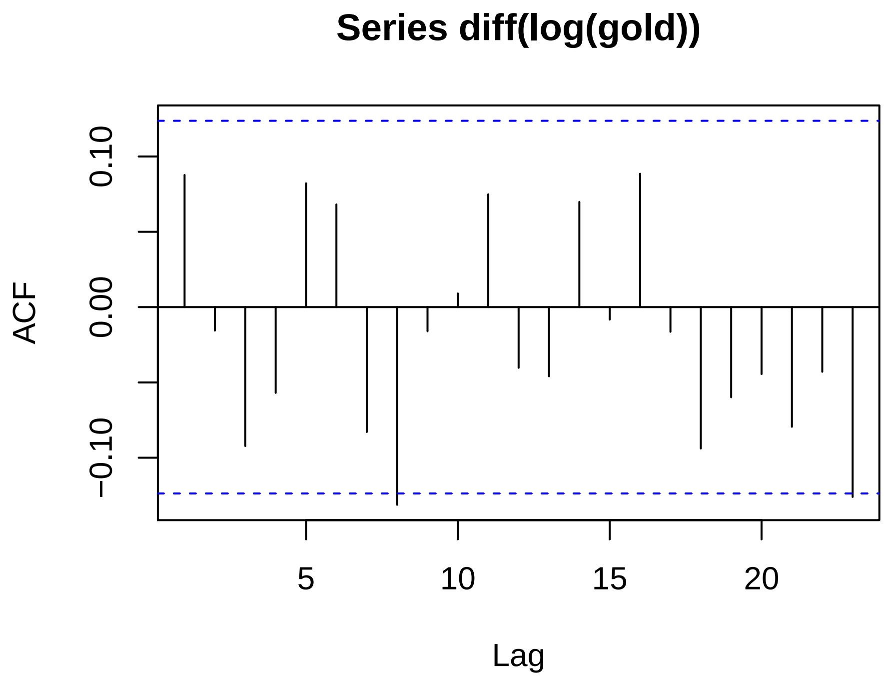

```r
stationary <- adf.test(diff(log(gold)), alternative = "stationary")
```

This looks heteroskedastic, and therefore not stationary, however, the variance is actually quite small, between -0.0277298 and 0.0244966.

Performing an Augmented Dickey-Fuller test for stationarity, we have a p-value of 0.01, so we reject the null hypothesis. The difference appears to be stationary.

(c) Use the sample ACF to investigate whether W_t is a white noise process.

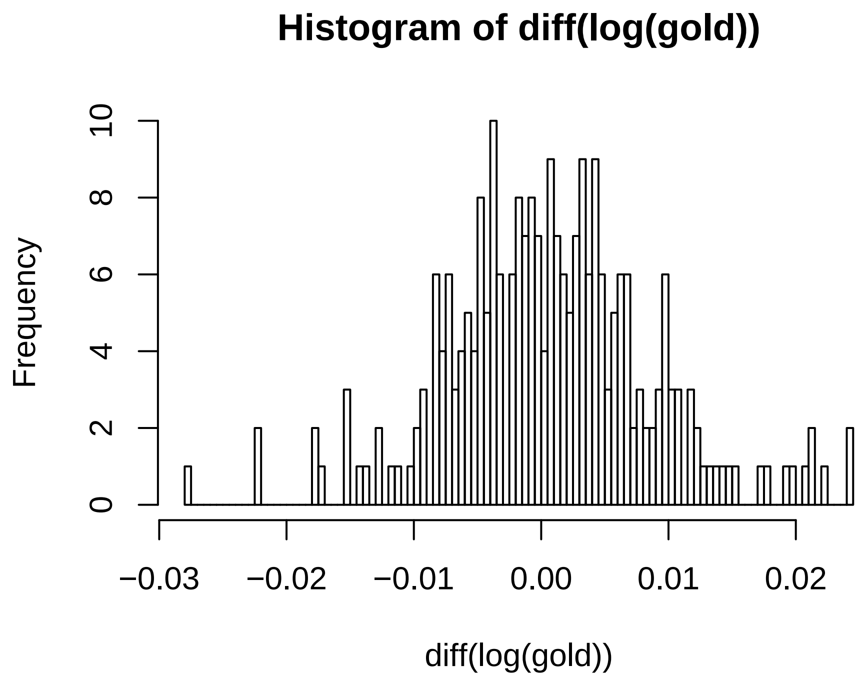
```r
acf(diff(log(gold)))
```



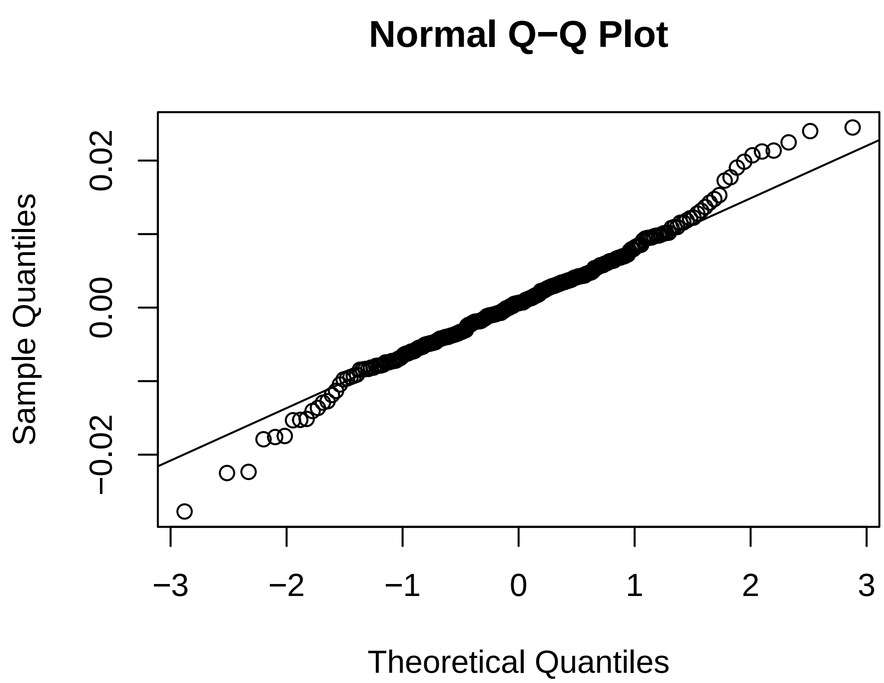
The ACF of the log difference appears to follow a white noise process.

(d) Investigate whether W_t is a normal white noise process.

```
hist(diff(log(gold)), breaks=100)
```




```r
qqnorm(diff(log(gold)))
qqline(diff(log(gold)))
```



```
gold.test <- shapiro.test(diff(log(gold)))
```

Although there are some outliers at the extremes, running a Shapiro-Wilk test for normality results in a p-value of $0.0151904 < 0.05$, so we can say this is a normal white noise process.

■