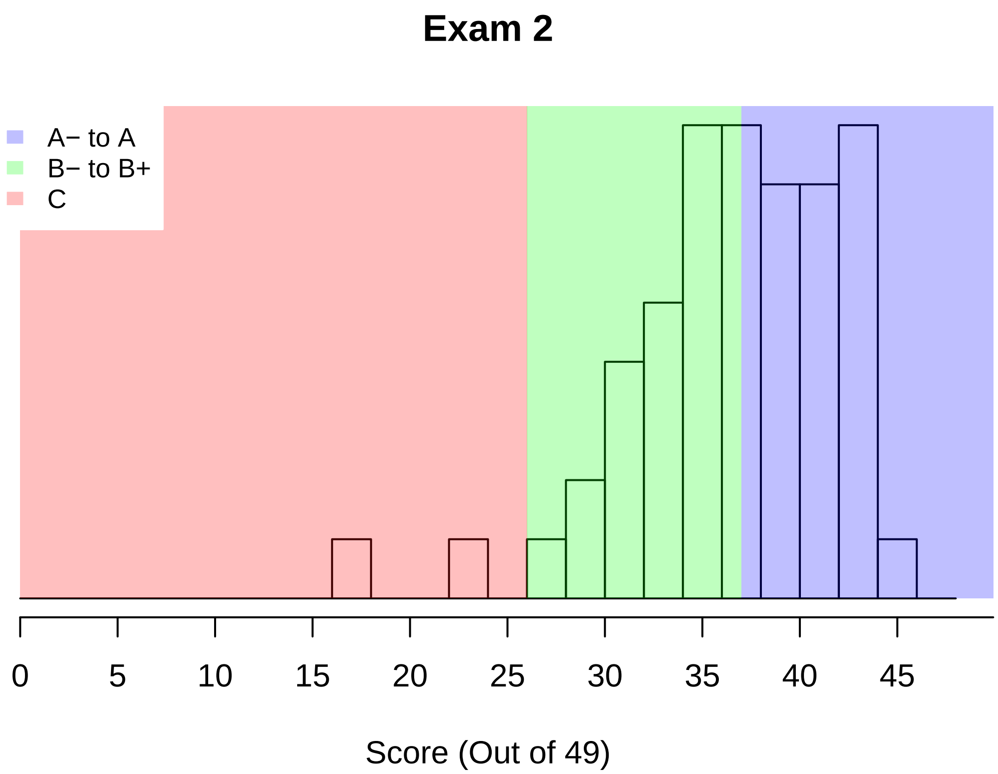
Exam 2

3/26/19

There are five questions, each of which has several parts. Neither the questions nor the parts are necessarily in order from easiest to most difficult. Make sure you have taken a look at and attempted all of the questions in the allotted time. Stop working and immediately turn in your exam when time has been called.

Name:	______________________
Question	Maximum Possible Points
1	14
2	9
3	6
4	8
5	10
Total	47

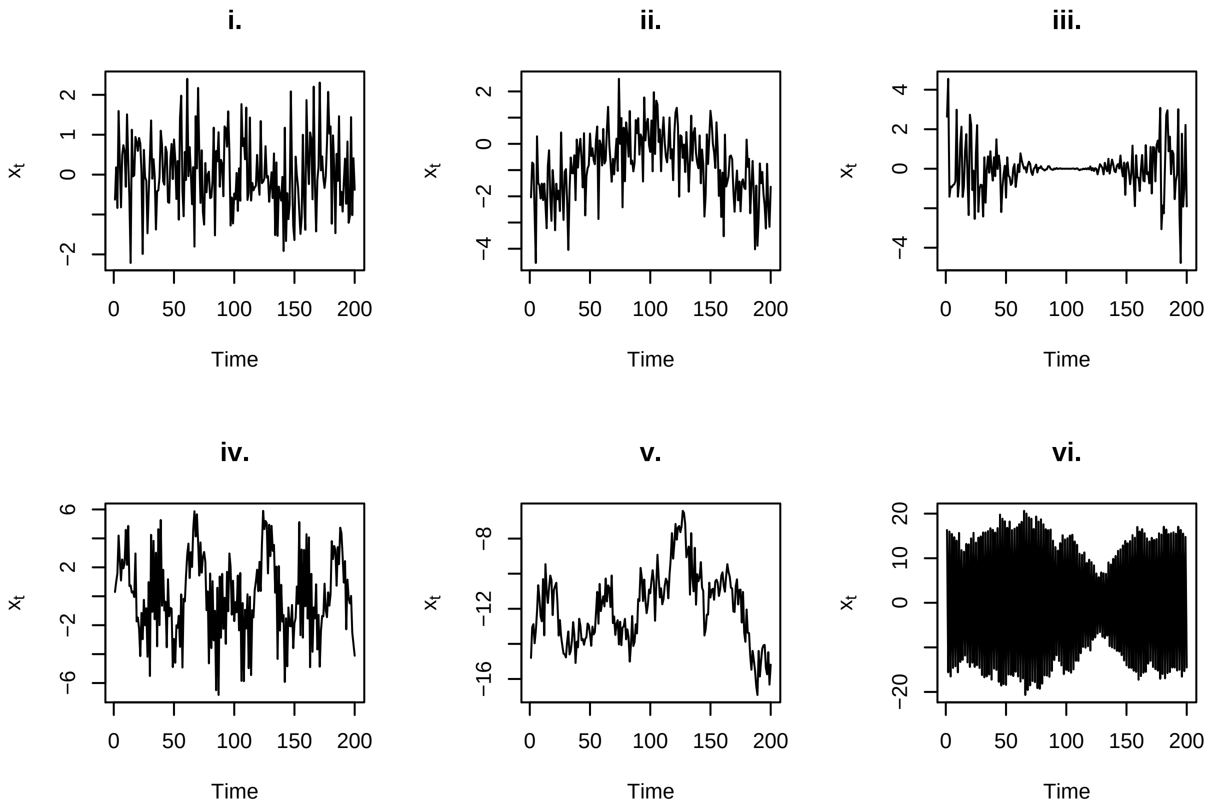
The figure below shows a histogram of scores. The mean was 36.76, the median was 37.5, and the standard deviation was 5.54. The rank correlation between the first and second exam scores was 0.58.



1. Assessing Stationarity

(a) Based on the following time series plots, indicate whether or not each time series appears to be stationary by filling in "yes" or "no."

Model	Appears Stationary?
i.	Y
ii.	N
iii.	N
iv.	Y
v.	N
vi.	N



(b) For each of the following models, indicate whether or not the model is approximately consistent with the null hypothesis of an augmented Dickey-Fuller (ADF) test, either with or without a linear time trend, by filling in "yes" or "no." Recall that we can write we can write any **ARMA**(p, q) as approximately an **AR**(k) process for some value of k. Then the approximate null hypothesis of an augmented ADF test without trend is that $x_t - x_{t-1}$ is a stationary **ARMA**(p, q) process, whereas the null hypothesis of an augmented ADF test with trend is $x_t - x_{t-1} - bt$ is a stationary **ARMA**(p, q) process for constant b.

- i. $x_t = w_t$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$;
- ii. $x_t = -10\left(t/n - 1/2\right)^2 + w_t$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$;
- iii. $x_t = 10\left(t/n - 1/2\right)^2 \times w_t$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$;
- iv. $x_t = \sum_{i=1}^5 u_i \cos\left(2\pi\left(i/6\right)t\right) + v_i \sin\left(2\pi\left(i/6\right)t\right)$, $u_i, v_i \overset{i.i.d.}{\sim} \mathcal{N}(0, i/12)$;
- v. $x_t = 0.5x_{t-1} + 0.25x_{t-2} + 0.25x_{t-3} + w_t$, , $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$;

- vi. $x_t = -x_{t-1} + w_t$, $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$.

Model	ADF Null without Trend	ADF Null with Trend
i.	Y	Y
ii.	N	Y
iii.	N	N
iv.	N	N
v.	Y	Y
vi.	N	N

- i. $x_t - x_{t-1} = w_t - w_{t-1}$ is a stationary **ARMA**(p, q) process;
- ii. $x_t - x_{t-1} = 1/n^2 - 20\left(t/n - 1/2\right)/n + w_t - w_t$ is not a stationary **ARMA**(p, q) process because of the linear term in t but $x_t - x_{t-1} + bt$ is, where $b = 20/n^2$;
- iii. $x_t - x_{t-1} = 10\left(t/n - 1/2\right)^2 \times w_t - 10\left(\left(t-1\right)/n - 1/2\right)^2 \times w_{t-1}$ has nonconstant variance, so it will not be a stationary **ARMA**(p, q) process.
- iv. $x_t - x_{t-1}$ may be stationary but is not an **ARMA**(p, q) process.
- v. $x_t - x_{t-1} = -0.5\left(x_{t-1} - x_{t-2}\right) + 0.25\left(x_{t-2} - x_{t-3}\right) + 0.25\left(x_{t-3} - x_{t-4}\right) + w_t - w_{t-1}$ is a stationary **ARMA**(p, q) process.
- vi. $x_t - x_{t-1} = -\left(x_{t-1} - x_{t-2}\right) + w_t - w_{t-1}$ is a non-stationary **ARMA**(p, q) process.

(c) Based on (b) and in no more than one sentence, is it possible to detect all kinds of non-stationarity based on an augmented Dickey-Fuller test, with or without trend?

No, some kinds of non-stationarity like non-constant variance will not be detected by a Dickey-Fuller test, regardless of whether or not a trend is included.

2. Differencing and Correlation

(a) Express the lag-one autocorrelation function of ∇x_t denoted by $\frac{\gamma_{\nabla x}(1)}{\gamma_{\nabla x}(0)} = \frac{\mathrm{E}[\nabla x_t \nabla x_{t-1}]}{\mathrm{E}\left[\nabla x_t^2\right]}$ in terms of the autocovariance function of a mean-zero time series x_t.

$$\rho_{\nabla x}(1) = \frac{\gamma_{\nabla x}(1)}{\gamma_{\nabla x}(0)} = \frac{\mathrm{E}\left[\nabla x_t \nabla x_{t-1}\right]}{\mathrm{E}\left[\nabla x_t^2\right]} = \frac{2\gamma_x(1) - \gamma_x(0) - \gamma_x(2)}{2\left(\gamma_x(0) - \gamma_x(1)\right)}$$

(b) If $x_t = w_t$, where $w_t \overset{i.i.d.}{\sim} \mathcal{N}(0, 1)$, then $\gamma_x(0) = 1$ and $\gamma_x(h) = 0$ for $h > 0$. How does the lag-one autocorrelation of x_t compare to the lag-one autocorrelation of ∇x_t?

$$\rho_{\nabla x}(1) = -\frac{1}{2}$$

(c) If $x_t = 0.9x_{t-1} + w_t$, where $w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1 - 0.9^2\right)$. Then then $\gamma_x(h) = 0.9^h$. How does the lag-one autocorrelation of x_t compare to the lag-one autocorrelation of ∇x_t?

$$\rho_{\nabla x}(1) = \frac{2\left(\frac{9}{10}\right) - 1 - \frac{81}{100}}{2\left(1 - \frac{9}{10}\right)} = \frac{\frac{180-100-81}{100}}{\frac{20}{100}} = -\frac{1}{20}.$$

(d) Fill in the blank with "stronger," "weaker," or "possibly stronger or weaker." If a time series x_t is differenced, the differenced time series ∇x_t will have possibly stronger or weaker lag-one autocorrelations than the undifferenced raw time series. This is about how the **magnitudes** of the autocorrelations compare, not the values!

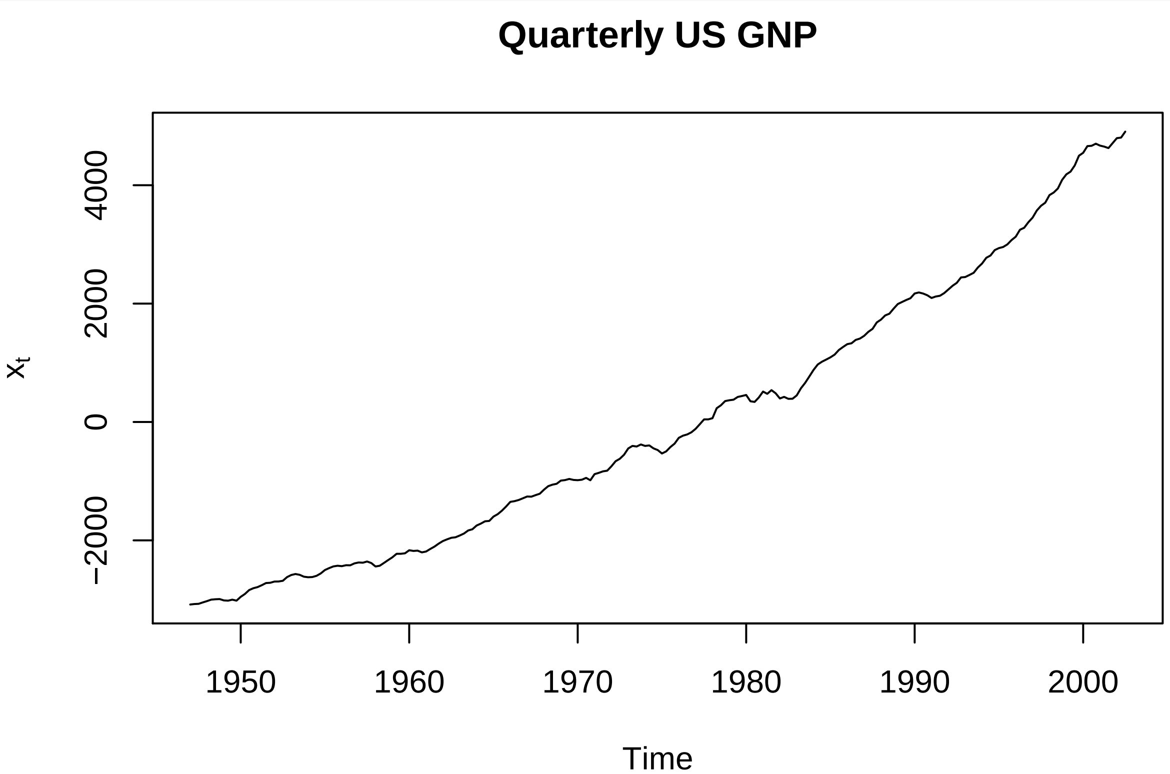
3. ARIMA

This question will ask you to analyze the `gnp` data from the `astsa` package, which gives the quarterly United States GNP from the first quarter of 1947 to the third quarter of 2002.

```
library(astsa)
data(gnp)
time <- time(gnp)
gnp <- c(gnp) - mean(gnp)
n <- length(gnp)
```

The data is plotted below.

```
plot(time, gnp, main = "Quarterly US GNP", ylab = expression(x[t]), xlab = "Time", type = "l")
```



(a) Does x_t appear stationary? If not, do you think ∇x_t would appear stationary, based on the plot of the data?

The time series x_t does not appear stationary because there is a linear trend present, but it looks like ∇x_t might be (because differencing will eliminate a linear trend).

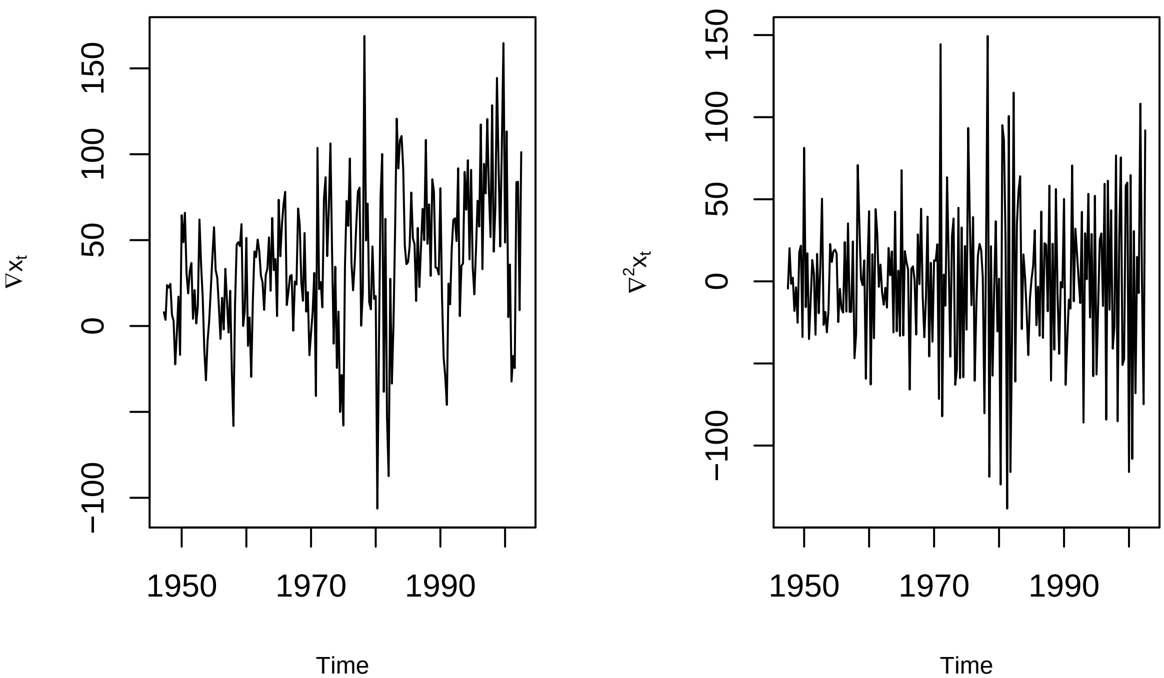
(b) How many times does a level-0.05 augmented Dickey-Fuller test indicate that we should difference the data? The value returned by `ndiffs` is printed after the code chunk.

```
library(forecast)
ndiffs(gnp, test = "adf", alpha = 0.05, type = "level")
```

```
## [1] 1
```

Once.

(c) The first and second differences are plotted below. Based just on the plots alone, how many times do you think we should difference the data?



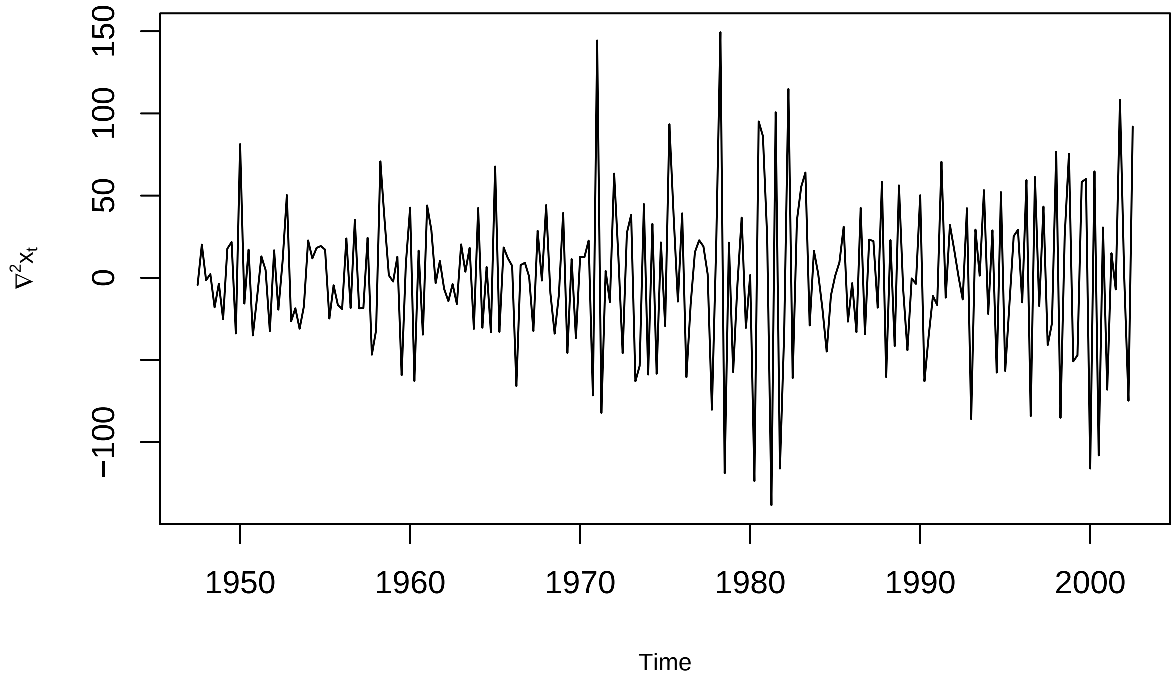
Twice.

(d) Indicate whether or not your conclusions in (b) are the same. If they are not the same, explain how many times **you** think we should difference the data (give a reason for the number of differences you think you should take).

Either once or twice could be accepted as answers, depending on your reasoning!

4. ARCH/GARCH

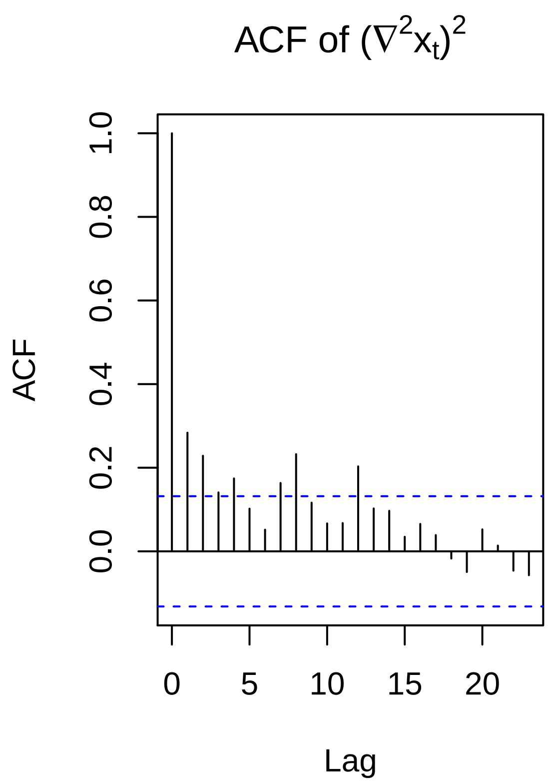
We're going to keep working with the `gnp` data from the `astsa` package, and continue to focus on the second differences $\nabla^2 x_t$, which are plotted again below.

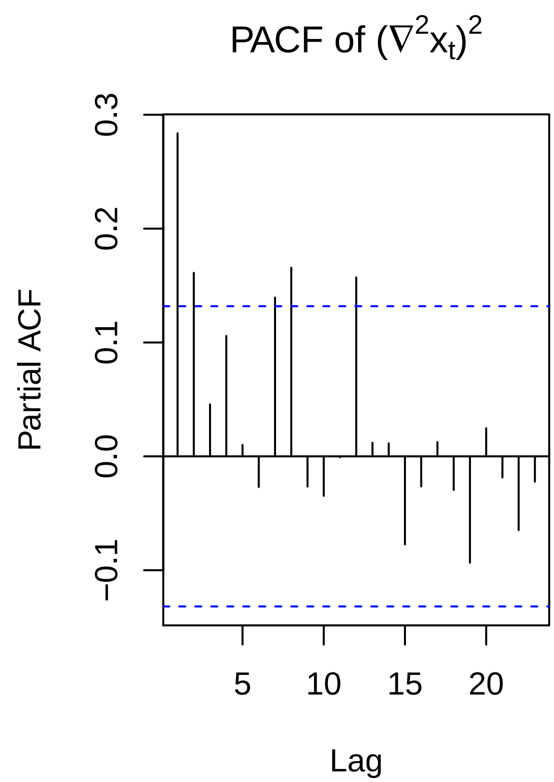


(a) Based on the above plot, is there evidence of variance nonstationarity?

Yes - the variance appears to be increasing as time passes.

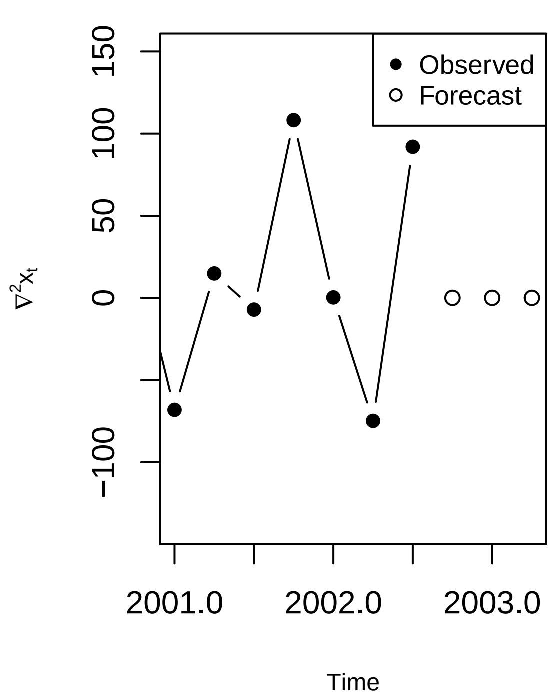
(b) Let's consider a $\mathbf{GARCH}(m, 0)$ model for $\nabla^2 x_t$. Based on the plotted ACF and PACF of the squared second differences $\left(\nabla^2 x_t\right)^2$, what would you choose for m?

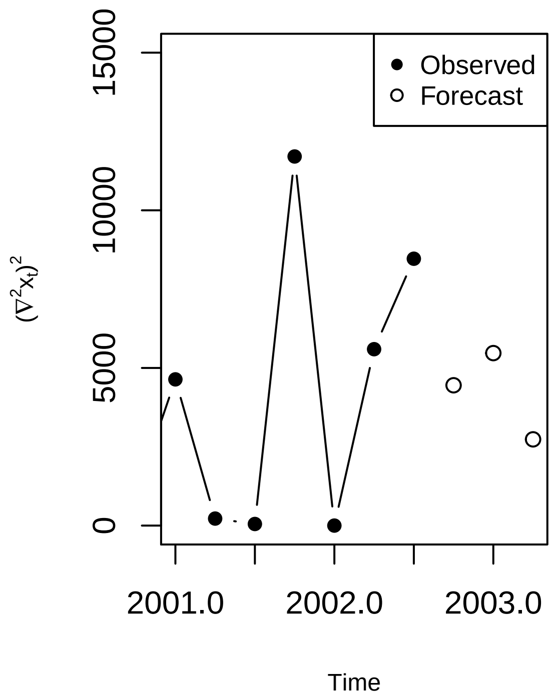




A $\mathbf{GARCH}(m, 0)$ model for $\nabla^2 x_t$ corresponds to an $\mathbf{AR}(m)$ model for the magnitudes $\left(\nabla^2 x_t\right)^2$. We can select the order of an $\mathbf{AR}(m)$ model from the PACF by selecting an order equal to the last lag for which the sample partial autocorrelation is outside of the 95% interval for $\gamma_x(h) = 0$. This means we would choose $m = 12$ here.

(c) Suppose we fit a $\textbf{GARCH}(m,0)$ model for the correct value of m based on part (b) to the second differences. Forecasts of $\nabla^2 x_t$ and $\left(\nabla^2 x_t\right)^2$ based on this $\textbf{GARCH}(m,0)$ model are given in the plots below. Based the plotted forecasts, does assuming a GARCH model for $\nabla^2 x_t$ help us forecast $\nabla^2 x_t$ or the magnitudes $\left(\nabla^2 x_t\right)^2$?





Assuming a GARCH model helps us forecast the magnitudes $\left(\nabla^2 x_t\right)^2$.

(d) In one sentence, explain why your answer in (c) makes sense given what we know about $\textbf{GARCH}(m,0)$ models. Hint: compare the autocorrelation function for $\nabla^2 x_t$ compared to the autocorrelation function of $\left(\nabla^2 x_t\right)^2$ under a $\textbf{GARCH}(m,0)$ model.

Under a $\textbf{GARCH}(m,0)$ model, the values $\nabla^2 x_t$ are uncorrelated but the magnitudes $\left(\nabla^2 x_t\right)^2$ are correlated, so past values of $\left(\nabla^2 x_t\right)^2$ help us forecast future values of $\left(\nabla^2 x_t\right)^2$ but past values of $\nabla^2 x_t$ do not help us forecast future values of $\nabla^2 x_t$.

5. Spectral Analysis

(a) Suppose

$$x_t = \sum_{k=1}^{r} v_k \cos\left(2\pi\omega_k t\right) + u_k \sin\left(2\pi\omega_k t\right),\ v_k, u_k \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_k^2\right)$$

$$y_t = \sum_{k=1}^{r} c_k \cos\left(2\pi\omega_k t\right) + d_k \sin\left(2\pi\omega_k t\right),\ c_k, d_k \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \tau_k^2\right)$$

The spectral density function of x_t is $f\left(\omega_k\right) = \sigma_k^2$ and the spectral density function of y_t is $g\left(\omega_k\right) = \tau_k^2$.

(a) Write out $z_t = ax_t + by_t$ in terms of a, b, the v_k's, the u_k's, the c_k's,the d_k's, the $\cos(2\pi\omega_k t)$'s and the $\sin(2\pi\omega_k t)$'s.

$$\begin{aligned} z_t &= ax_t + by_t \\ &= \sum_{k=1}^{r} (av_k + bc_k)\cos(2\pi\omega_k t) + (au_k + bd_k)\sin(2\pi\omega_k t) \end{aligned}$$

(b) Describe how to define e_k, f_k, and ν_k in terms of a, b, the v_k's, the u_k's, the c_k's, and the d_k's such that the following holds for z_t:

$$z_t = \sum_{k=1}^{r} e_k \cos(2\pi\omega_k t) + f_k \sin(2\pi\omega_k t)\,,\; e_k, f_k \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \nu_k^2\right).$$

$$\begin{aligned} e_k &= av_k + bc_k \\ f_k &= au_k + bd_k \\ \nu_k &= \sqrt{a^2\sigma_k^2 + b^2\tau_k^2} \end{aligned}$$

(c) What is the spectral density function $h(\omega_k)$ of z_t?

$$h(\omega_k) = a^2\sigma_k^2 + b^2\tau_k^2.$$

(d) Suppose that

$$\begin{aligned} x_t &= \phi_1 x_{t-1} + w_t, w_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_w^2\right) \\ y_t &= \psi_1 y_{t-1} + v_t, v_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_v^2\right). \end{aligned}$$

Write out $z_t = ax_t + by_t$, substituting $\phi_1 x_{t-1} + w_t$ in for x_t and $\psi_1 y_{t-1} + v_t$ in for y_t and collecting terms that correspond to the same time point.

$$\begin{aligned} z_t &= ax_t + by_t \\ &= a\phi_1 x_{t-1} + b\psi_1 y_{t-1} + aw_t + bv_t. \end{aligned}$$

(e) If we don't assume anything about the values of a, b, ϕ_1, and ψ_1, can we know whether or not z_t will be an $\mathbf{AR}(p)$ process? Hint: If z_t is an $\mathbf{AR}(p)$ process, we can find values $\gamma_1, \dots, \gamma_p$ such that $z_t = \gamma_1 z_{t-1} + \dots + \gamma_p z_{t-p}$. Just answer yes or no.

No - unless ϕ_1 and ψ_1 are equal or a and b are equal, we're not going to be able to write z_t as a linear function of its past values z_{t-1}.