Assignment1_Fall2022

Alexander Tran

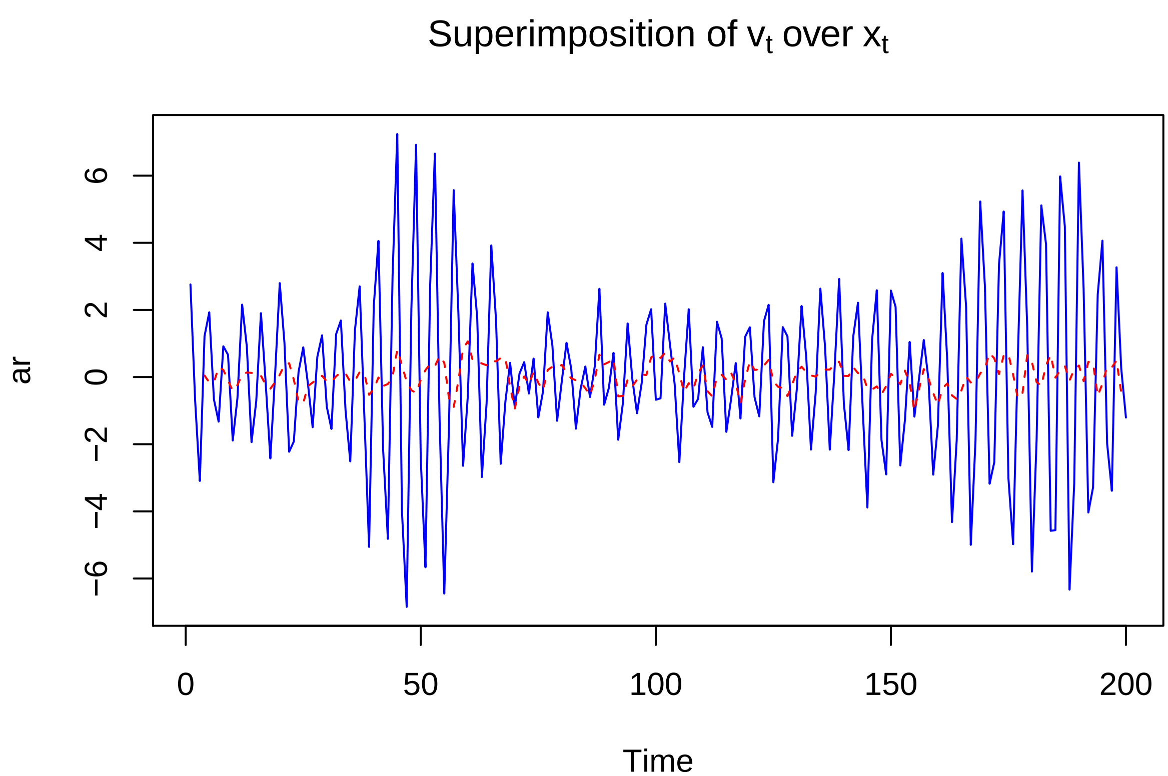
9/28/2022

Q1

```
set.seed(416)
# 1. Generate n = 200 observations from this process.
wn = rnorm(202,0,1) # 2 extra to avoid startup problems
ar = stats::filter(wn, filter=c(0,-.9), method="recursive")[-(1:2)] # remove first 2

# 2. Apply the moving average filter vt= (xt+ xt-1 + xt-2 + xt-3)/4 to the data you generated.
ma = stats::filter(ar, filter=rep(1/4,4), method="convolution", sides=1) # Moving average

# 3. Plot xt as a line and superimpose vt as a dashed line. Comment!
plot.ts(ar, main=bquote("Superimposition of"~ v[t] ~"over"~ x[t]), col="blue")
lines(ma, lty="dashed", col="red")
```



Looking at this plot, we can see that applying the moving average filter v_t has smoothed the series x_t fairly well, removing a good portion of the noise that was present.

Q2

1.

$$
\begin{aligned}
x_t &= \delta + x_{t-1} + w_t \\
&= \delta + (\delta + x_{t-2} + w_{t-1}) + w_t \\
&= 2\delta + x_{t-2} + \sum_{j=t-1}^{t} w_j \\
&= 3\delta + x_{t-3} + \sum_{j=t-2}^{t} w_j \\
&= \dots \\
&= \delta(t-1) + x_1 + \sum_{j=2}^{t} w_j \\
&= \delta t + x_0 + \sum_{j=1}^{t} w_j \\
&= \delta t + \sum_{j=1}^{t} w_j
\end{aligned}
$$

2. Given that $E[w_t] = 0$,

$$
\begin{aligned}
\mu_t = E[x_t] &= E[\delta t + \sum_{j=1}^{t} w_j] \\
&= E[\delta t] + E[\sum_{j=1}^{t} w_j] \\
&= \delta E[t] + \sum_{j=1}^{t} E[w_j] \\
&= \delta t
\end{aligned}
$$

$$
\begin{aligned}
\gamma(s,t) &= cov(x_s, x_t) \\
&= cov(\delta s + \sum_{i=1}^{s} w_i, \delta t + \sum_{j=1}^{t} w_j) \\
&= cov(\sum_{i=1}^{s} w_i, \sum_{j=1}^{t} w_j)
\end{aligned}
$$

When $s = t$,

$$
\begin{aligned}
cov(x_t, x_t) &= cov(\sum_{i=1}^{t} w_i, \sum_{j=1}^{t} w_j) \\
&= cov(w_1, w_1) + cov(w_2, w_2) + ... + cov(w_{t-1}, w_{t-1}) + cov(w_t, w_t) \\
&= t\sigma_w^2
\end{aligned}
$$

When $s = t + 1$,

$$\begin{aligned} cov(x_{t+1}, x_t) &= cov(\sum_{i=1}^{t+1} w_i, \sum_{j=1}^{t} w_j) \\ &= cov(w_1, w_1) + cov(w_2, w_2) + ... + cov(w_{t-1}, w_{t-1}) + cov(w_t, w_t) \\ &= t\sigma_w^2 \end{aligned}$$

When $s = t - 1$,

$$\begin{aligned} cov(x_{t-1}, x_t) &= cov(\sum_{i=1}^{t-1} w_i, \sum_{j=1}^{t} w_j) \\ &= cov(w_1, w_1) + cov(w_2, w_2) + ... + cov(w_{t-2}, w_{t-2}) + cov(w_{t-1}, w_{t-1}) \\ &= (t-1)\sigma_w^2 \end{aligned}$$

When $s = t + 2$,

$$\begin{aligned} cov(x_{t+2}, x_t) &= cov(\sum_{i=1}^{t+2} w_i, \sum_{j=1}^{t} w_j) \\ &= cov(w_1, w_1) + cov(w_2, w_2) + ... + cov(w_{t-1}, w_{t-1}) + cov(w_t, w_t) \\ &= \sigma_w^2 \end{aligned}$$

When $s = t - 2$,

$$\begin{aligned} cov(x_{t-2}, x_t) &= cov(\sum_{i=1}^{t-2} w_i, \sum_{j=1}^{t} w_j) \\ &= cov(w_1, w_1) + cov(w_2, w_2) + ... + cov(w_{t-3}, w_{t-3}) + cov(w_{t-2}, w_{t-2}) \\ &= (t-2)\sigma_w^2 \end{aligned}$$

Following the same pattern, we can compute other values of s to arrive at the following autocovariance function:

$$\gamma(s,t) = \begin{cases} t\sigma_w^2 & s - t \geq 0 \\ (s)\sigma_w^2 & 0 > s - t > -t \\ 0 & s - t \leq -t \end{cases}$$

3. The mean is not constant and depends on time t. Thus, we can conclude that x_t is not stationary.

4.

$$\begin{aligned} \rho_x(t-1, t) &= \frac{\gamma(t-1,t)}{\sqrt{\gamma(t-1,t-1)\gamma(t,t)}} \\ &= \frac{(t-1)\sigma_w^2}{\sqrt{(t-1)\sigma_w^2 t\sigma_w^2}} \\ &= \frac{(t-1)}{\sqrt{(t-1)(t)}} \\ &= \frac{\sqrt{(t-1)}}{\sqrt{t}} \\ &= \sqrt{\frac{(t-1)}{t}} \end{aligned}$$

$$\lim_{t \to +\infty} \sqrt{\frac{(t-1)}{t}} = 1$$

$\lim_{t\to+\infty}\sqrt{\frac{(t-1)}{t}} = 1$

5. To make this series stationary, we can take the first differences of x_t. The transformed series is given by:

$z_t = \triangle x_t = x_t - x_{t-1} = \delta + x_{t-1} + w_t - x_{t-1} = \delta + w_t$

Given that $E[w_t] = 0$,

$$\begin{aligned} \mu_t = E[z_t] &= E[\delta + w_t] \\ &= E[\delta] + E[w_t] \\ &= \delta \end{aligned}$$

$$\begin{aligned} \gamma(s,t) &= cov(x_s, x_t) \\ &= cov(\delta + w_i, \delta + w_j) \\ &= cov(w_s, w_t) \end{aligned}$$

It is easy to see that the autocovariance function is as follows:

$$\gamma(s,t) = \begin{cases} \sigma_w^2 & |s-t| = 0 \\ 0 & |s-t| > 0 \end{cases}$$

Thus, because the mean is independent of time t, and the autocovariance function depends on s and t only through their difference $|s-t|$, z_t is stationary.

Q3

1.

$$\begin{aligned}\bar{x} &= \frac{1}{n}\sum_{t=1}^{n} x_t \\ &= \frac{1}{10}(24+20+25+31+30+32+37+33+40+38) \\ &= 31\end{aligned}$$

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_{t+h}-\bar{x})(x_t-\bar{x})$$

$$\begin{aligned}\hat{\gamma}(0) &= \frac{1}{10}\sum_{t=1}^{10}(x_t-\bar{x})(x_t-\bar{x}) \\ &= \frac{1}{10}[(24-31)(24-31)+(20-31)(20-31)+(25-31)(25-31)+(31-31)(31-31)+(30-31)(30-31) \\ &+(32-31)(32-31)+(37-31)(37-31)+(33-31)(33-31)+(40-31)(40-31)+(38-31)(38-31)] \\ &= 37.8\end{aligned}$$

$$\begin{aligned}\hat{\gamma}(1) &= \frac{1}{10}\sum_{t=1}^{9}(x_{t+1}-\bar{x})(x_t-\bar{x}) \\ &= \frac{1}{10}[(20-31)(24-31)+(25-31)(20-31)+(31-31)(25-31)+(30-31)(31-31)+(32-31)(30-31) \\ &+(37-31)(32-31)+(33-31)(37-31)+(40-31)(33-31)+(38-31)(40-31)] \\ &= 24.1\end{aligned}$$

$$\begin{aligned}\hat{\gamma}(2) &= \frac{1}{10}\sum_{t=1}^{8}(x_{t+2}-\bar{x})(x_t-\bar{x}) \\ &= \frac{1}{10}[(25-31)(24-31)+(31-31)(20-31)+(30-31)(25-31)+(32-31)(31-31)+(37-31)(30-31) \\ &+(33-31)(32-31)+(40-31)(37-31)+(38-31)(33-31)] \\ &= 11.2\end{aligned}$$

$$\begin{aligned}\hat{\gamma}(3) &= \frac{1}{10}\sum_{t=1}^{7}(x_{t+3}-\bar{x})(x_t-\bar{x}) \\ &= \frac{1}{10}[(31-31)(24-31)+(30-31)(20-31)+(32-31)(25-31)+(37-31)(31-31)+(33-31)(30-31) \\ &+(40-31)(32-31)+(38-31)(37-31)] \\ &= 5.4\end{aligned}$$

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

$$\begin{aligned}\hat{\rho}(0) &= \frac{\hat{\gamma}(0)}{\hat{\gamma}(0)} \\ &= \frac{37.8}{37.8} \\ &= 1\end{aligned}$$

$$\begin{aligned}\hat{\rho}(1) &= \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} \\ &= \frac{24.1}{37.8} \\ &= 0.638\end{aligned}$$

$$\begin{aligned}\hat{\rho}(2) &= \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)} \\ &= \frac{11.2}{37.8} \\ &= 0.296\end{aligned}$$

$$\begin{aligned}\hat{\rho}(3) &= \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)} \\ &= \frac{5.4}{37.8} \\ &= 0.143\end{aligned}$$

2.

$$t_{\rho(h)} = \frac{\hat{\rho}(h)}{\sigma_{\hat{\rho}(h)}}$$

$$\sigma_{\hat{\rho}(h)} = \frac{1}{\sqrt{n}}$$

$$\begin{aligned}t_{\rho(1)} &= \frac{\hat{\rho}(1)}{\sigma_{\hat{\rho}(1)}} \\ &= \frac{0.638}{(1/\sqrt{10})} \\ &= 2.018\end{aligned}$$

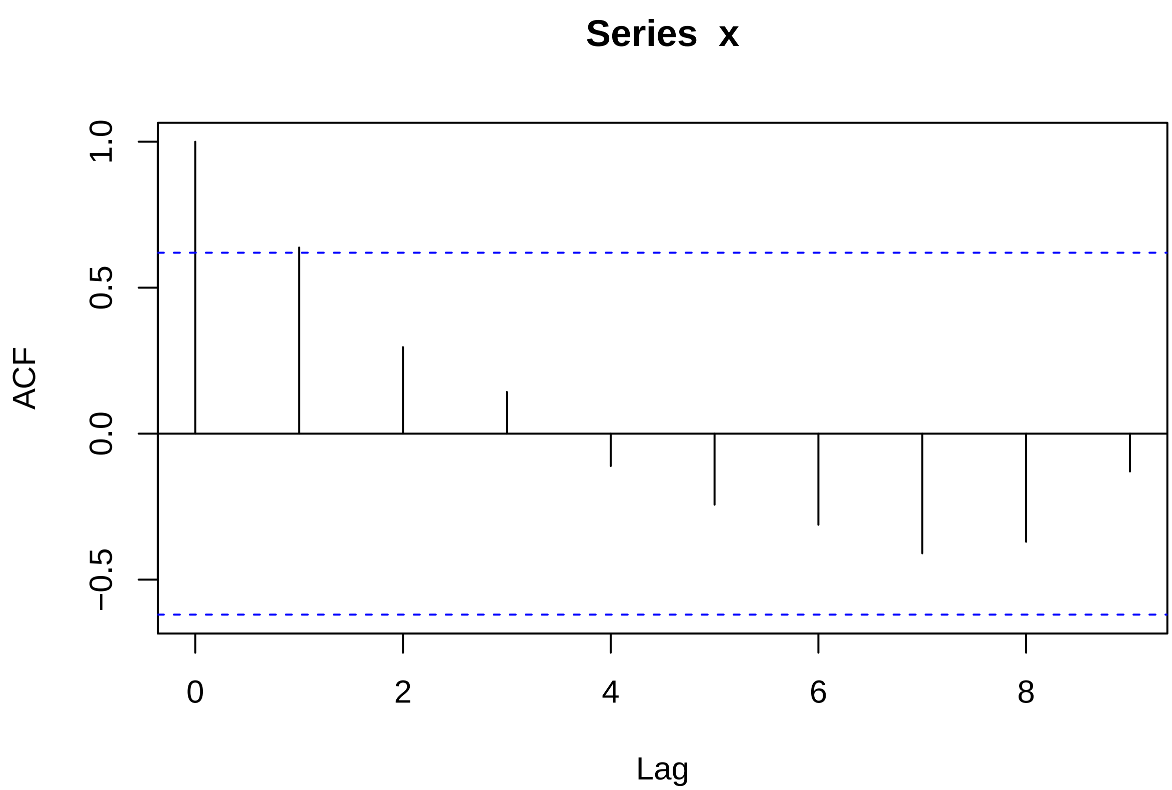
$t_{\rho(1)}$ is greater than $z_{0.025} = 1.96$, so we reject the null hypothesis that the theoretical autocorrelation at lag $h = 1$ equals zero at the 5% significance level.

3.

```r
#1
x <- c(24, 20, 25, 31, 30, 32, 37, 33, 40, 38)
acf(x, lag.max=3, plot=FALSE) # produce the acf values
```

```
##
## Autocorrelations of series 'x', by lag
##
##     0     1     2     3
## 1.000 0.638 0.296 0.143
```

```r
#2
acf(x, lag.max = NULL, plot = TRUE)
```



We see in the plot that the peak at lag 1 is outside of the two-standard errors interval, thus making it significant. Seeing this we reject the null hypothesis at the 5% significance level.

Q4

1.

```
set.seed(416)

wn = rnorm(502,0,1) # 2 extra to avoid startup problems
ma = stats::filter(wn, filter=rep(1/3,3), method="convolution", sides=2)
ma = ma[-(1:1)][(1:500)] # remove first and last values
acf(ma, lag.max=20, plot=FALSE)
```

```
##
## Autocorrelations of series 'ma', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000  0.645  0.315 -0.025  0.002  0.004  0.016 -0.002  0.031  0.028  0.038
##     11     12     13     14     15     16     17     18     19     20
##  0.025  0.052  0.068  0.068  0.047  0.019 -0.007 -0.037 -0.028  0.002
```

The theoretical autocorrelation function is given by:

$$\rho(h) = \begin{cases} 1 & h = 0 \\ \frac{2}{3} & |h| = 1 \\ \frac{1}{3} & |h| = 2 \\ 0 & |h| \geq 2 \end{cases}$$

The sample ACF is very similar to the actual ACF. The autocorrelation values at lag 0 is 1.0, as expected, the values at lags 1 and 2 are also very close to their expected values of 0.667 and 0.333, and the values past lag 2 are all very close to 0, as we expect.

2.

```
set.seed(416)

wn = rnorm(52,0,1) # 2 extra to avoid startup problems
ma = stats::filter(wn, filter=rep(1/3,3), method="convolution", sides=2)
ma = ma[-(1:1)][(1:50)] # remove first and last values
acf(ma, lag.max=20, plot=FALSE)
```

```
##
## Autocorrelations of series 'ma', by lag
##
##      0      1      2      3      4      5      6      7      8      9     10
##  1.000  0.462 -0.043 -0.433 -0.294 -0.135 -0.025  0.171  0.134 -0.007 -0.187
##     11     12     13     14     15     16     17     18     19     20
## -0.122  0.026  0.147  0.174  0.159  0.037 -0.070 -0.157 -0.114 -0.097
```

Changing n significantly affects the results. While we observe that the autocorrelation value at lag 0 is still 1, the values at lags 1 and 2 are significantly lesser than their expected values. And while some values past lag 2 are near-zero, we still see many unexpected values such as those at lags 3, 4, 5, 7, 8, 10, 11, etc.

Q5

$$
\begin{aligned}
\mu_t = E[x_t] &= E[cos(2\pi(\frac{t}{12} + \phi))] \\
&= \int_{-2}^{2} xcos(2\pi(\frac{x}{12} + \phi))dx \\
&= 0
\end{aligned}
$$

$$
\begin{aligned}
\gamma(s,t) &= cov(x_s, x_t) \\
&= cov(cos(2\pi(\frac{s}{12} + \phi)), cos(2\pi(\frac{t}{12} + \phi)))
\end{aligned}
$$

This process is weak stationary as its mean is independent of time t, and the autocovariance function depends on s and t only through their difference $|s - t|$.