STAT 221 Homework 3

John Rapp Farnes / 405461225

3/3 2020

4.5

a)

We have for w_t: $E(w_t) = 0$ and for $h \geq 0$

$$\gamma_w(h) = cov(w_t, w_{t+h}) = \begin{cases} 1, & h = 0 \\ 0, & \text{otherwise} \end{cases}$$

as $w_t \in N(0, 1)$ i.i.d. Hence, as $E(w_t)$ constant and $\gamma_w(h)$ only depends on h, w_t is stationary.

For $x_t = w_t - \theta w_{t-1}$ we have $E(x_t) = E(w_t) - \theta E(w_{t-1}) = 0$, and for $h \geq 0$

$$\begin{aligned} \gamma_x(h) &= cov(w_t - \theta w_{t-1}, w_{t+h} - \theta w_{t+h-1}) \\ &= cov(w_t, w_{t+h}) - \theta cov(w_{t-1}, w_{t+h}) - cov(w_t, -\theta w_{t+h-1}) + cov(\theta w_{t-1}, \theta w_{t+h-1}) \\ &= \begin{cases} 1 + \theta^2, & h = 0 \\ -\theta, & h = 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Hence, as $E(x_t)$ constant and $\gamma_x(h)$ only depends on h, x_t is stationary.

b)

We have, using (4.16) in the book:

$$\begin{aligned} f(\omega) &= \sum_{h=-\infty}^{\infty} \gamma_x(h) e^{-2\pi i \omega h} \\ &= \gamma_x(-1) e^{2\pi i \omega} + \gamma_x(0) e^0 + \gamma_x(1) e^{-2\pi i \omega} \\ &= -\theta e^{2\pi i \omega} + 1 + \theta^2 - \theta e^{-2\pi i \omega} \\ &= 1 + \theta^2 - 2\theta (\frac{e^{2\pi i \omega} + e^{-2\pi i \omega}}{2}) \\ &= 1 + \theta^2 - 2\theta cos(2\pi\omega) \end{aligned}$$

4.9

```
library(astsa)
pg <- mvspec(sunspotz, log="no")

alpha <- 0.05
U = qchisq(alpha/2, 2)
```

```r
L = qchisq(1-alpha/2, 2)

max_index <- which(pg$spec == max(pg$spec))
freq <- pg$freq[max_index]

1/freq
```

```
## [1] 10.90909
```

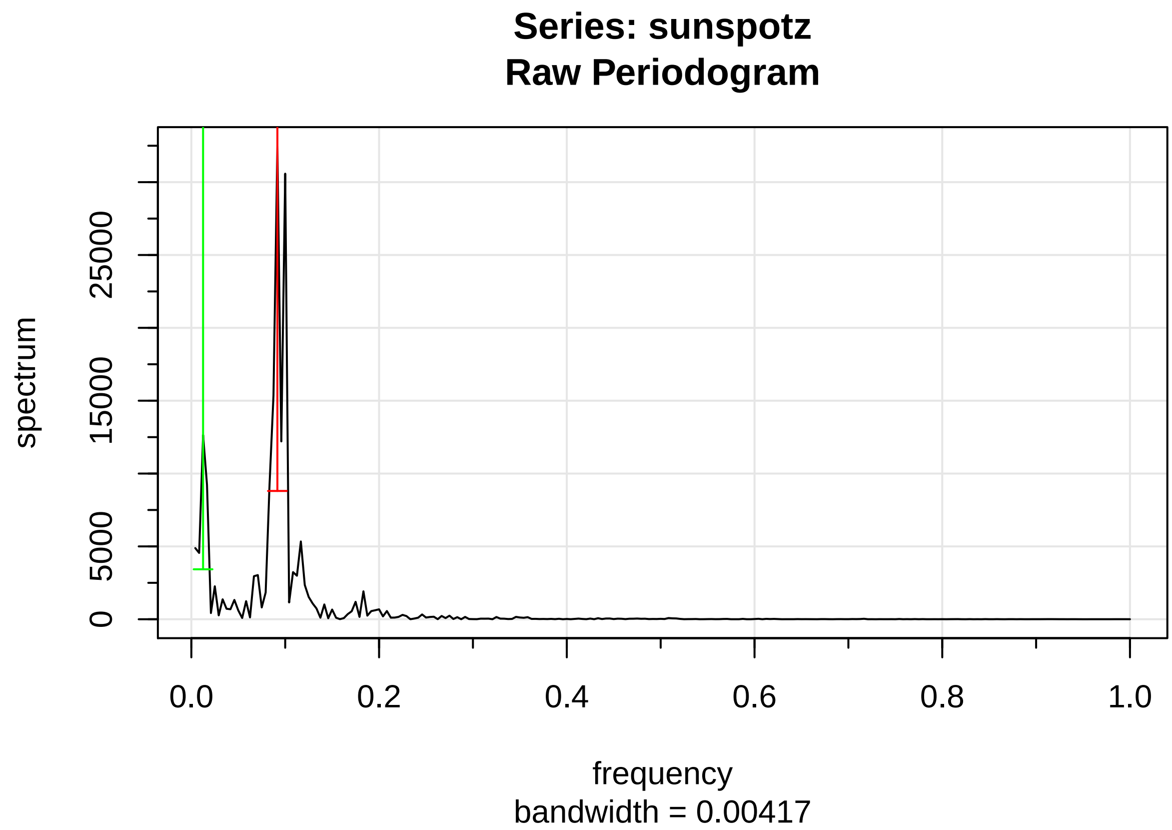
```r
conf <- c(2*pg$spec[max_index]/L,2*pg$spec[max_index]/U)
segments(x0=freq,y0=conf[1],x1=freq,y1=conf[2],col="red")
segments(x0=freq-0.01,y0=conf[1],x1=freq+0.01,y1=conf[1],col="red")


max_index <- which(pg$spec == max(pg$spec[which(pg$freq < 0.05)]))
freq <- pg$freq[max_index]

1/freq
```

```
## [1] 80
```

```r
conf <- c(2*pg$spec[max_index]/L,2*pg$spec[max_index]/U)
segments(x0=freq,y0=conf[1],x1=freq,y1=conf[2],col="green")
segments(x0=freq-0.01,y0=conf[1],x1=freq+0.01,y1=conf[1],col="green")
```



The periodogram shows two frequency bands with peaks. The first and highest one is around the 11 year period, and the second around an 80 year period. The 11 year period has several peaks close to it, indicating that the cycle is centered around 11 years but irregular. The same goes for the 80 year one but to a lesser extent. Both appear to be significantly higher than neighboring baseline values.

These frequency bands can be seen clearer by smoothing the periodogram, e.g. with a Daniell(2,2) filter:

```r
k <- kernel("daniell", c(2,2))
pg <- mvspec(sunspotz, kernel=k, log="no")

l <- 1/sum(k$coef^2)

alpha <- 0.05
U = qchisq(alpha/2, 2*l)
L = qchisq(1-alpha/2, 2*l)

max_index <- which(pg$spec == max(pg$spec))
freq <- pg$freq[max_index]

1/freq
```

```
## [1] 10.90909
```

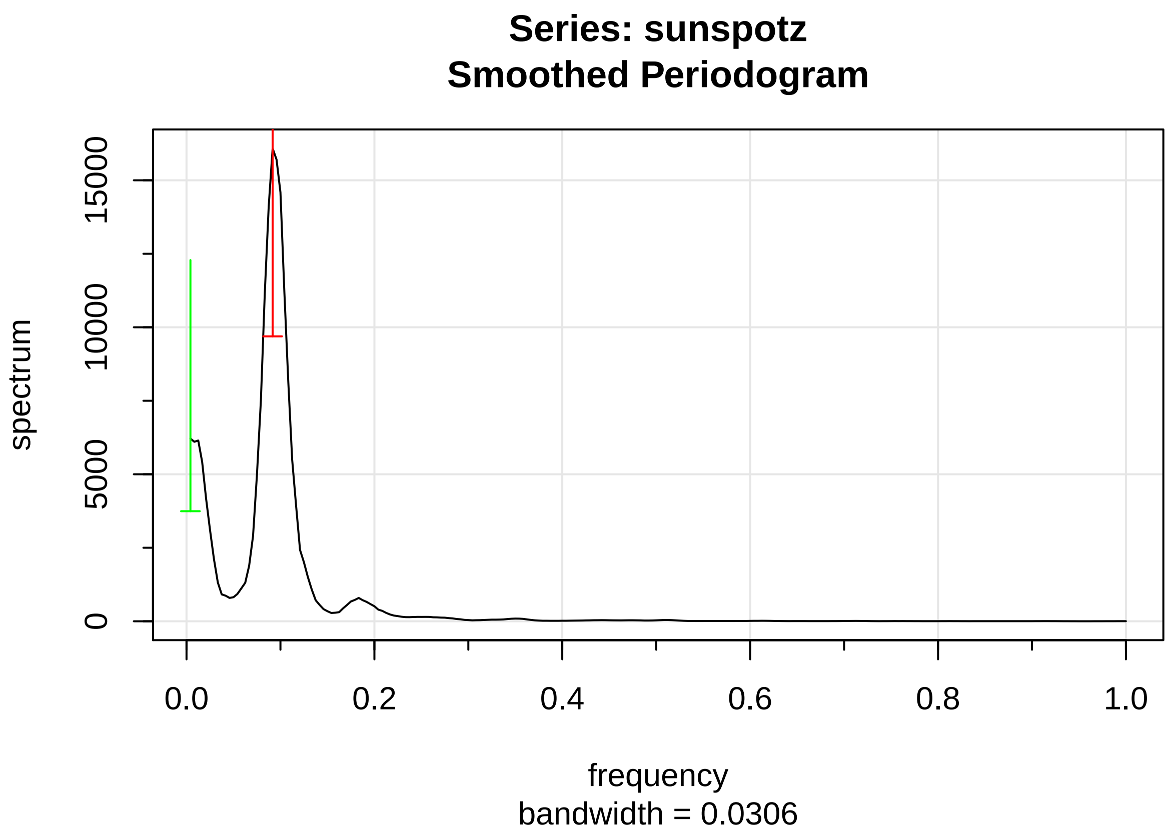
```r
conf <- l*c(2*pg$spec[max_index]/L,2*pg$spec[max_index]/U)
segments(x0=freq,y0=conf[1],x1=freq,y1=conf[2],col="red")
segments(x0=freq-0.01,y0=conf[1],x1=freq+0.01,y1=conf[1],col="red")


max_index <- which(pg$spec == max(pg$spec[which(pg$freq < 0.05)]))
freq <- pg$freq[max_index]

1/freq
```

```
## [1] 240
```

```r
conf <- l*c(2*pg$spec[max_index]/L,2*pg$spec[max_index]/U)
segments(x0=freq,y0=conf[1],x1=freq,y1=conf[2],col="green")
segments(x0=freq-0.01,y0=conf[1],x1=freq+0.01,y1=conf[1],col="green")
```



The smoothened periodogram reveals the same 11 year period peak, however the 80 year period peak is not as clear as there is no local maxima in the lower frequencues, rather the periodogram appears to be highest around 0. Both peaks appear significant.

4.19

```
pg <- spec.ar(sunspotz, log="no")

max_index <- which(pg$spec == max(pg$spec))
freq <- pg$freq[max_index]

1/freq
```

```
## [1] 10.39583
```

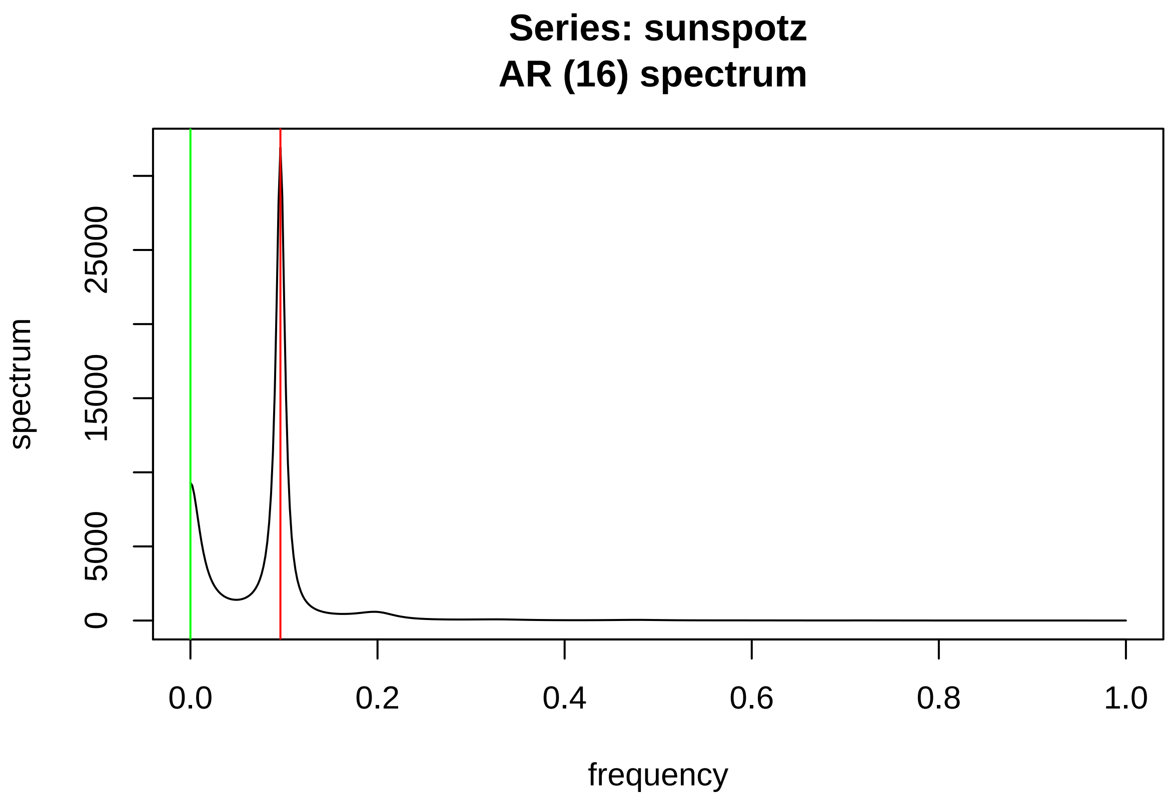
```
abline(v=freq, col="red")


max_index <- which(pg$spec == max(pg$spec[which(pg$freq < 0.05)]))
freq <- pg$freq[max_index]

1/freq
```

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

```
abline(v=freq, col="green")
```



The auto AR fitting algorithm in the astsa package found that an AR(16) process had the lowest BIC value and was used as the parametric estimator. Similar to the non-parametric periodogram, the spectrum has a peak around an 11 year period, with a slightly lower estimate for the peak (~10.4). The low frequency peak is not as easily decernable as there is no local maxima, rather the maximum value is at frequency 0.

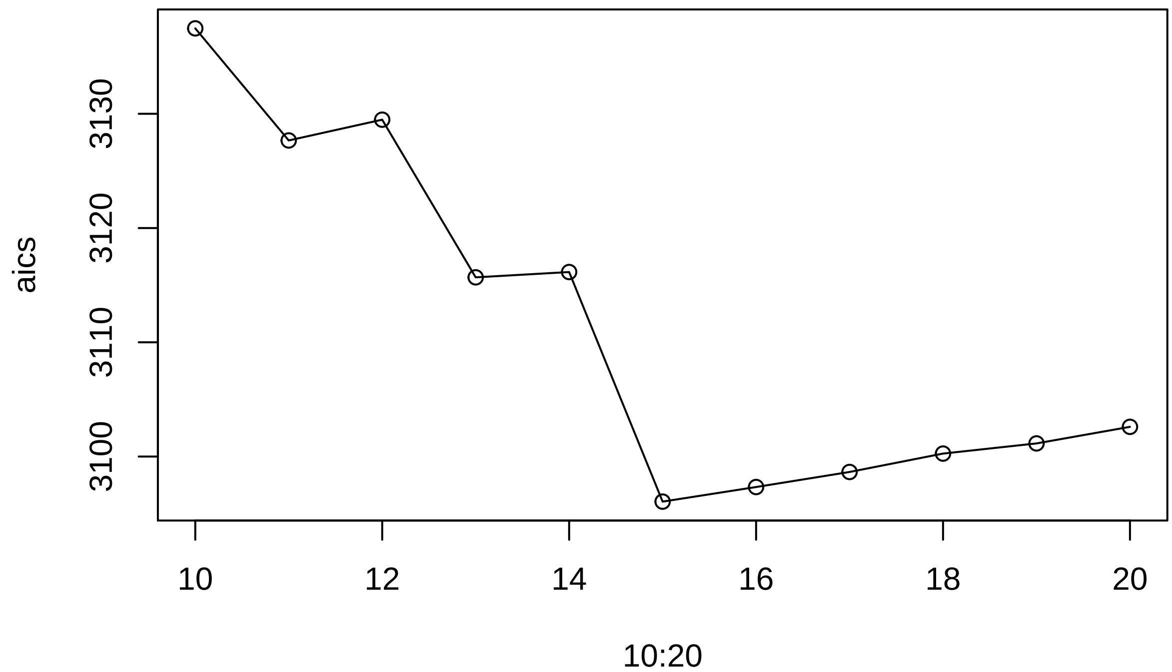
As a note: comparing AIC values of different AR(p) models also results in p close to 16 being the best fitting model:

```
# Showing a subset of possible p values in order to achieve reasonable computing time
p <- 10:20

models <- lapply(p, function(p) {
  return(arima(sunspotz, order=c(p,0,0)))
})

aics <- unlist(lapply(models, function(model) {
  return(model$aic)
}))

plot(10:20, aics, type="o")
```




```r
models[which(aics == min(aics))]
```

```
## [[1]]
##
## Call:
## arima(x = sunspotz, order = c(p, 0, 0))
##
## Coefficients:
##          ar1      ar2     ar3      ar4     ar5      ar6     ar7      ar8
##       1.8349  -1.4446  1.1172  -1.0693  0.8927  -0.8847  0.8767  -0.7256
## s.e.  0.0455   0.0958  0.1160   0.1254  0.1321   0.1358  0.1393   0.1414
##          ar9     ar10    ar11     ar12    ar13     ar14    ar15  intercept
##       0.6044  -0.6222  0.6228  -0.5209  0.3973  -0.3525  0.2215    50.6999
## s.e.  0.1397   0.1367  0.1335   0.1266  0.1177   0.0978  0.0465     5.8293
##
## sigma^2 estimated as 45.55:  log likelihood = -1531.03,  aic = 3096.06
```