Homework 5

Due: Thursday 3/13/19 by 12:00pm (noon)

Spectral Analysis

1. Show that $\gamma_x(h) = \sum_{k=1}^{q} \sigma_k^2 \cos(2\pi\omega_k h)$ if

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

You will find the following trigonometric identities helpful:

- $\cos(a) = \cos(-a)$;
- $\sin(-a) = -\sin(a)$;
- $\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$.

2. In class we computed the scaled periodogram of an observed time series $\boldsymbol{x}$ by hand. I have made a little R function called `scaled periodogram` that does what we did in class.

```
scaled.periodogram <- function(x) {
  n <- length(x)
  # Get number of columns in our design matrix
  Z <- matrix(nrow = n, ncol = n)

  # First column is always the intercept!
  Z[, 1] <- 1
  for (i in 2:n) {
    if (i%%2 == 0) {
      Z[, i] <- cos(2*pi*floor(i/2)*1:n/n)
    } else {
      Z[, i] <- sin(2*pi*floor(i/2)*1:n/n)
    }
  }
  linmod <- lm(x~Z-1)

  # Let's record the coef magnitudes
  m <- ifelse(n%%2 == 0, n/2, (n - 1)/2 + 1)
  coef.mags <- numeric(m)
  for (i in 1:length(coef.mags)) {
    if (i == 1) {
      coef.mags[i] <- coef(linmod)[1]^2
    } else if (i == length(coef.mags) & n%%2 == 0) {
      coef.mags[i] <- coef(linmod)[length(coef(linmod))]^2
    } else {
      coef.mags[i] <- sum(coef(linmod)[1 + 2*(i - 2) + 1:2]^2)
    }
  }
  return(list("coef.mags" = coef.mags, "freqs" = 0:(m - 1)/n, "Z" = Z))
}
```

(i) Consider an $\mathbf{AR}(1)$ model with $\phi_1 = 0.5$ and $\sigma_w^2 = 1$, and consider time series of lengths $n = 25$ and $n = 100$. Simulate $1,000$ time series for each value of n according to this model, and record the scaled periodogram for each.

(a) Make a pair of plots in a single plot window using `par(mfrow = c(1, 2))` - one for each value of n. For both plots, use the range $(-20, 5)$ on the y-axis and $(0, 0.5)$ on the x-axis. In each plot for a single value of n, plot the average log scaled periodogram on against the frequency across all of the simulations. Also plot the log scaled periodograms against frequency for ten randomly selected simulations.

(b) Based on the average log scaled periodogram, does any specific frequency dominate the periodogram of this $\mathbf{AR}(1)$ process? Answer with at most one sentence.

(c) In one sentence, when n gets bigger, do the log scaled periodograms become smoother?

(d) Redo part (a) for an $\mathbf{AR}(1)$ model with $\phi_1 = 0$ and $\sigma_w^2 = 1$.

(e) Describe the shape of the average log periodogram from (d) in one sentence.

(f) Consider the following $\mathbf{AR}(p)$ models, all with $\sigma_w^2 = 1$.

i. $p = 1$, $\phi_1 = 0.99$
ii. $p = 2$, $\phi_1 = 0.04$, $\phi_2 = 0.92$
iii. $p = 2$, $\phi_1 = 0.04$, $\phi_2 = -0.92$
iv. $p = 3$, $\phi_1 = 0.42$, $\phi_2 = -0.29$, $\phi_3 = 0.15$

Set $n = 100$ and simulate $1,000$ time series for each value of n according to this model, and record the scaled periodogram for each. Using the range $(-6, 3)$ on the y-axis and $(0, 0.5)$ on the x-axis, plot the average log scaled periodogram on against the frequency across all of the simulations for each model.

(g) In at most one sentence, comment on how the $\mathbf{AR}(p)$ parameters affect the shape of the periodogram based on (h).