CEE7430 Homework 10: Frequency Domain Analysis

Due: 4/23/09

Reading:

Chapters 10-12 in Wei, W. S., Time Series Analysis, Univariate and Multivariate Methods, Addison Wesley, 478 p. 1990.

Focus on the following:
10.1. Overview
10.2. This is theoretical backup - skim
10.3. Finite sequences
10.4. Periodic sequences and power spectrum
10.5. Generalization to non periodic sequences
11.1.1. The spectrum of the autocovariance function
12.1.1 and 12.1.2. The periodogram
12.2. The sample spectrum
12.3. Skim this to get a sense for smoothing methods.

For the Colorado River at Lees Ferry, used in earlier homework, do the following using monthly flow data.

1. Compute the average across all years flow for each month as a representation of the seasonal cycle (i.e. 12 numbers). Find the fourier representation of this sequence using Wei's equation 10.3.2 (equation (S3) and (S4) in the presentation). Plot a graph of a_k, b_k, and sqrt(a_k^2+b_k^2) versus k for each valid value of k. Plot graphs of $a_k\cos(2\pi k t/n)$, $b_k\sin(2\pi k t/n)$ and the sum $a_k\cos(2\pi k t/n) + b_k\sin(2\pi k t/n)$ versus t for each valid value of k. Sum a few of these to see which components reproduce the majority of the variation in your data.

2. Find the complex fourier representation of this sequence using Wei's equation 10.3.4. Plot a graph of Re(c_k), Im(c_k) and $|c_k|$ versus k for each valid value of k. Plot graphs of $c_k e^{i w_k t}$ versus t for each valid value of k.

3. Use the R built in function `fft()` to obtain the Fourier transform of the same data. The R function returns its results using complex numbers. Compare the absolute value of the coefficients it produces with sqrt(a_k^2+b_k^2) calculated above. You should discover a similar pattern but different order of magnitude numbers due to a different Fourier transform formula being used by R. The formula that `fft()` in R uses is $c_k = \sum_{t=0}^{n-1} Z_t e^{-i\omega_k t}$. This does not have the 1/n in the definition from the Wei text. R also uses indexing of the data from 0 instead of 1 (common among C programmers). (Note, that in Fourier analysis there is not consistency

between references on the definitions so you need to be on the lookout for this question and make sure the formulas you work with are consistent. Relative comparisons are more important than actual values of Fourier coefficients so a scalar multiplier usually does not matter). In R the functions `Re()`, `Im()`, `Mod()` that work with complex numbers may also be useful. Plot a graph of Fourier coefficients from the built in function. Plot a graph of each Fourier component from the built in function.

Do the following problems using the full time series of monthly flows.

4. Find the Fourier representation of the full record using one of the approaches developed above. Plot a graph of the amplitude (sqrt(a_k^2+b_k^2) or $|c_k|$) versus w_k for each valid value of k. Label the frequency corresponding to the annual cycle. Plot a graph of the single Fourier component [either $c_k e^{iw_k t}$ or $a_k\cos(2 \pi k t/n) + b_k\sin(2 \pi k t/n)$] versus t, corresponding to the annual cycle. Plot a graph of the sum of the 5 Fourier components with largest amplitude and a period longer than the annual cycle.

5. The fourier representation of the full record should have n/2 terms (n~1080 months, 12 x 90) corresponding to frequencies between 0 and π. Evaluate the amplitude squared [(a_k^2+b_k^2) or $|c_k|^2$] for each component and plot amplitude squared versus frequency. This is the raw periodogram (Wei equation 12.1.4). Evaluate $\frac{1}{2}\sum_{k=1}^{n/2}(a_k^2 + b_k^2)$. This should be the same as the variance of Z_t. (Percivals theorem). If working with c_k evaluate $\sum_{k=1}^{n}|c_k/n|^2$. (The /n is because it is omitted in the R `fft()` implementation.).

6. Divide the interval between 0 to π into about 20 bins and average the amplitude squared [(a_k^2+b_k^2) or $|c_k|^2$] within each bin. Plot average amplitude squared versus bin frequency. This is a simple smoothing of the periodogram and an estimate of the power spectrum.

7. Use the R built in functions to estimate the power spectrum and compare and reconcile your results with the plots from questions 5 and 6 above. In R, the ts package has functions `spectrum()` and `spec.pgram()`. Use the spans argument to adjust the smoothing from the spectrum function, e.g. `spectrum(z,spans=c(5,7,12,20)).`

8. Calculate the autocovariance function and estimate the power spectrum by a direct Fourier transform of this. Compare and reconcile your results with the plots from questions 5 to 7 above.

9. Generate a long record using the best model you have available for generation of streamflow for this river from prior homework. Compute and plot the power spectrum for this synthetic long record and compare it to the power spectra from the data above.