Problems Week 6

Due Wednesday, November 14, 2018

This problem set is to be completed independently, without collaborating with your classmates. At the top of your problem set, please write and sign the following statement: "I certify that this represents my own work and that I have not worked with classmates or other individuals to complete this assignment." If you have questions, you may e-mail me.

1. **Compute two spectra.** For this problem set, please generate a 10,000 element data set with Gaussian white noise. Use your white noise to generate a second data set using an autoregressive (AR1) process. In Matlab you can do this as follows:

```
a=randn(10000,1);
b(1)=a(1);
for i=2:length(a)
  b(i)=.5*b(i-1)+a(i);
end
```

   Now compute spectra for the white noise and autoregressive data sets by breaking the data up into segments. (You can do this by reshaping your data vector as a 500×20 matrix, which will give you 20 segments, each 500 points long.) What are the differences between the spectra?

2. **Add error bars to your spectra.** Use the χ^2 formulation to compute uncertainties for your spectra, and show this on your plot.

3. **Use Monte Carlo simulation to verify the χ^2 error bar.** To check the error bars, you'll want to generate multiple ensembles of data and compute spectra for each of them. I would suggest using 200 matrices that are 500×20 elements each. Compute spectra for each of them. This will provide you with 200 independent realizations of the spectrum, and you can use these 200 values to study the range of possible spectra that you could obtain from white noise. Examine the pdf of your 200 values. (Actually, in the white noise case, because the statistics of at Gaussian white noise are the same at all frequencies, you can merge all the frequencies to produce a larger ensemble.) Is the pdf consistent with your expectations for a variable with a χ^2 distribution?

   Now, for each frequency, sort the 200 realizations of the pdf by size (e.g. using the "sort" command in Matlab. Since you are looking for the 95% confidence range, you'll want to find the limits that exclude the lowest 2.5% and the highest 2.5% of your data—presumably the 6th and 195th points in each sorted set of pdfs.) What is the ratio of the upper limit to the lower limit? Are the error bars derived using the Monte Carlo process consistent with those from question 1?

4. **Evaluate whether windowing alters degrees of freedom.** Now use 50% overlapping segments and apply a triangle window to your data. Repeat the analysis in

question 3. Show that the triangle window alters the uncertainties consistent with the predicted change in degrees of freedom discussed in Lecture 11.

5. **Bonus: Derive the reduction in degrees of freedom.** Use the formulation from Lecture 11 to derive the effective degrees of freedom for spectra computed with a triangle window with 50% overlap. Is this result consistent with your results from question 4?