Problems Week 6

Due Tuesday, November 9, 2021

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
```

   or in Python:

```
# number of point
N = 10000
# generate random numbers with normal dist. (Gaussian white noise)
x = np.random.randn(N)
# build red noise (here the time series will have a "memory")
auto = np.empty_like(x)
auto[0] = x[0]
for i in range(1,N):
    auto[i] = 0.5*auto[i-1] + x[i]
```

   Now compute spectra for the white noise and autoregressive data sets by breaking the data up into segments with 50% overlaps (with no windowing). (You can do this by reshaping your data vector as a 500×20 matrix, which will give you 20 segments, each 500 points long, and a 500×19 matrix containing the 50% overlaps.) Following the procedure that you've used before, add error bars to your spectra. What are the differences between the spectra?

2. **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. (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?

If you'd like, you can check your results using the red spectrum as well. One difference is that the white spectrum exhibits the same behavior at all frequencies allowing you to average in frequency. That isn't possible with the red spectrum, though you can average the ratio between the upper and lower bounds of the uncertainty estimate.

3. **Monte Carlo process for a Hanning window.** Repeat the Monte Carlo process to estimate the χ^2 error bar for data using a Hanning window for each segment and 50% overlap. Are the results consistent with what you would expect based on the discussion in class?

4. **Monte Carlo process for a cosine window.** Finally, repeat the Monte Carlo process for a cosine window ($\cos(\pi t/T)$, where $-T/2 < t < T/2$). How many effective degrees of freedom do you obtain for 50% overlap.

5. **Bonus.** Derive the analytic solution for a cosine window with 50% overlap. Does it agree with your numerical results?