Fitting a model to measured power spectrum

A common way to model the noise power spectrum is as the sum of white and red ($1/\nu$) components. This can be parametrized as

$$N(\nu) = A\left[1 + \left(\frac{\nu}{\nu_*}\right)^\alpha\right], \tag{1}$$

where A is the white noise power, α is the slope of the red component, and ν_* is the "knee frequency", *i.e.* the frequency at which the white and red components have equal amplitude.

Figure 1 shows the timestream[1] and power spectrum for simulated noise with $A = 1$, $\nu_* = 1$ Hz, and $\alpha = -1$. From the right-hand panel, we can see that the power spectrum of the data (blue) clearly follows the model (orange), but there is a lot of variance from point to point.

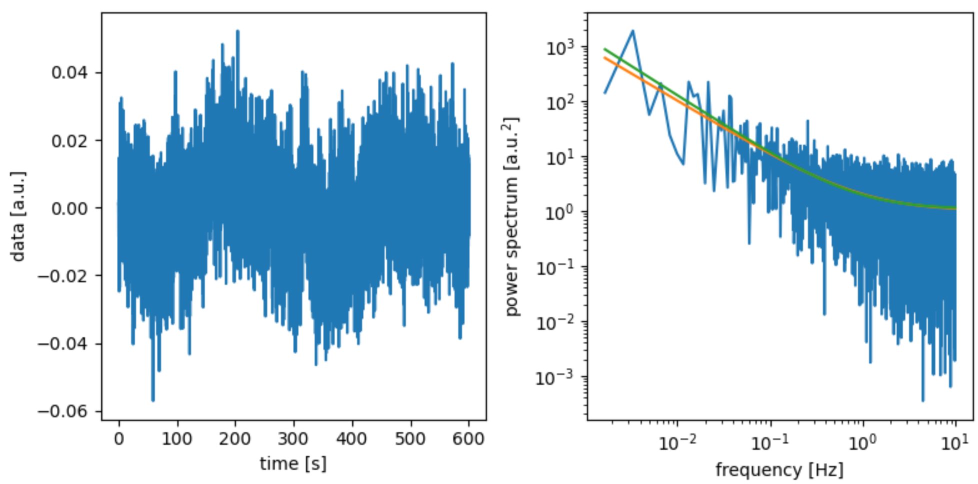

Figure 1: Simulated timestream data (left) and the power spectrum of these data (right). The power spectrum model used to generate the data is shown as the orange line in the right-hand panel. The green line is an attempt to fit the power spectrum by maximizing the likelihood.

If we know the model, *i.e.* equation 1 with the parameter values listed above, then we can write down a probability density function (pdf) for the power spectrum. The Fourier amplitude at frequency ν is a complex number with real and imaginary parts drawn independently from a Gaussian distribution with zero mean and variance equal to $N(\nu)/2$. Then we calculate the power spectrum by taking the absolute square of the Fourier amplitude, which should have a χ^2 pdf with two degrees-of-freedom and expectation value equal to $N(\nu)$. This pdf is given by

$$P_\nu(x) = \frac{1}{N_\nu}\exp\left(-\frac{x}{N_\nu}\right). \tag{2}$$

[1]For this document, I will talk about time domain data with power spectrum in the frequency domain, but it is exactly the same for spatial or angular power spectra.

I will leave it as an exercise for the reader to show that $P_\nu(x)$ is properly normalized and that $\langle x \rangle = N_\nu$.

This pdf is equivalent to the likelihood to obtain a particular value of the power spectrum, x, given model parameters A, ν_*, and α. So we can rewrite it as

$$\mathcal{L}_\nu(A, \nu_*, \alpha | x) = \frac{1}{N_\nu(A, \nu_*, \alpha)} \exp\left(-\frac{x}{N_\nu(A, \nu_*, \alpha)}\right) . \tag{3}$$

The only difference between equations 2 and 3 is that for equation 2 we are holding the model fixed and thinking about the probability for different values of x while for equation 3 we have fixed data realization x and allow the model to vary. Also, $P(x)$ is normalized so that it integrates to 1 over $x \in [0, \infty]$ while the normalization of the likelihood for integration over model parameters is unspecified.

If the timestream data are stationary[2], then each point in the power spectrum (each frequency) is independent of all the others. So that means that we can write a likelihood for the entire power spectrum as the product of the likelihoods for each point, $\mathcal{L} = \prod_\nu \mathcal{L}_\nu$, or better yet, the log-likelihood for the full power spectrum is the sum of the log-likelihoods at each frequency:

$$\log \mathcal{L} = \sum_\nu \log \mathcal{L}_\nu \tag{4}$$

$$= -\sum_\nu \frac{x}{N_\nu} - \sum_\nu \log N_\nu \, . \tag{5}$$

So now I can go back to my simulated data and power spectrum in Figure 1, write down the log-likelihood function, then try to find the parameters that minimize $-\log \mathcal{L}$. When I do this, I get the green curve on Figure 1 with parameters $A = 1.058$, $\nu_* = 0.853$ Hz, and $\alpha = -1.074$.

[2]Apodizing the data before Fourier transform breaks stationarity, but that probably isn't a big problem so long as the apodization isn't crazy.