LINEAR LEAST SQUARES AND SIGMA CLIPPING

Due date: 3/27/2024, 9am

In class we have shown that fitting a (linear) model to the data using (linear, weighted) least squares is akin to finding the inverse of the quadratic form $\mathbf{A}^\top\mathbf{W}\mathbf{A}$. This is essentially what off-the-shelf least squares methods do under the hood. That said, χ^2 cost function is only appropriate when data meet certain conditions, and "blemishes" can cause problems. We talked about heteroskedasticity, serial correlation, multi-collinearity and model mis-specification.

1. Write a generative linear function for an N-dimensional model. Incorporate options to add noise and data "blemishes" from above.
2. Define the cost function and *thoroughly* explore its behavior as a function of data "blemishes".
3. Sample over the solution using MCMC and determine parameter posteriors. Again, *thoroughly* explore the posteriors as a function of deviations from classical assumptions.
4. Legendre polynomials appear in the solution of the Laplace equation in spherical coordinates:

   $$\nabla^2 f(r) = 0; \qquad f(r) = \sum_n \left(A_n r^n + B_n r^{-(n+1)}\right) P_n(\cos\theta).$$

   For the given boundary conditions, $f(r)$ can be solved for coefficients A_n and B_n. A convenient way to generate polynomials P_n is by the following recursion formula:

   $$(n+1)P_{n+1} = (2n+1)xP_n - nP_{n-1}, \quad P_0 = 1.$$

   Legendre polynomials are a very good choice for fitting non-oscillatory data because of their orthogonality.

   §

   Using sigma-clipping, find the continuum of the absorption spectrum by fitting it with a Legendre series. You can either use your own spectrum or download one from the course homepage.
5. Download a *Kepler* or a *TESS* light curve of your favorite object and use sigma-clipping to detrend it.