The file *ENSO.txt* contains weather data. The independent variable is the monthly averaged atmospheric pressure differences between Easter Island and Darwin, Australia and the dependent variable is time in months. This difference drives the trade winds in the southern hemisphere. The data contains three cyclical signals and should be fit to the model

$$\begin{aligned} y_\theta(t) = \theta_1 + \theta_2 \cos(2\pi t/12) + \theta_3 \sin(2\pi t/12) + \theta_5 \cos(2\pi t/\theta_4) \\ +\theta_6 \sin(2\pi t/\theta_4) + \theta_8 \cos(2\pi t/\theta_7) + \theta_9 \sin(2\pi t/\theta_7) \end{aligned}$$

The annual cycle is the strongest and corresponds to parameters θ_2 and θ_3 in the model. (Observe that these parameters scale Fourier terms with periods of 12 months.) Two other cycles with unknown periods (given by θ_4 and θ_7) are described by the model. These cycles correspond to the El Niño and the Southern Oscillation.

One of the challenges in nonlinear fits is finding a reasonable starting guess for an optimization algorithm. In this problem you will use the data to reason about reasonable initial guesses for the parameters.

a. Argue that θ_1 corresponds to the average value of the data and use that to suggest a reasonable starting guess for θ_1.

b. Argue that θ_2, θ_3, θ_5, θ_6, θ_8 and θ_9 must correspond to the size of the variation in the signal from the mean and suggest a reasonable range of starting guesses for these parameters.

c. Argue that θ_4 and θ_7 correspond to the period of the two other signals in the data. What is the shortest period you could reasonably infer from data? What is the longest period you could reasonably infer from the data? Based on these arguments, suggest a range of starting guesses for these parameters.

d. Denote the model map for this learning problem by $\boldsymbol{F}(\theta)$. Evaluate the $\|\partial \boldsymbol{F}/\partial\theta_4\|$ and $\|\partial \boldsymbol{F}/\partial\theta_7\|$ for values of θ_4 and θ_7 that are within the range you suggested in the previous part and compare them to values evaluated well outside that range. What do you observe about the norm of these derivatives when the parameters θ_4 and θ_7 are pushed to extreme values outside your suggested range??

e. Fit this model to the data. Begin your fit from a few different starting points, including points near your intial guesses for θ_4 and θ_7, as well as those far away. Can you choose initial parameter values for which your fitting algorithms fail.

f. Plot the data and your model at the best fit parameters. Calculate the Fisher Information Matrix at the best fit and report the variance in each of the parameter estimates. Don't forget to scale your results by σ. Assume that σ^2 is the sum of the squared error divided by the degrees of freedom