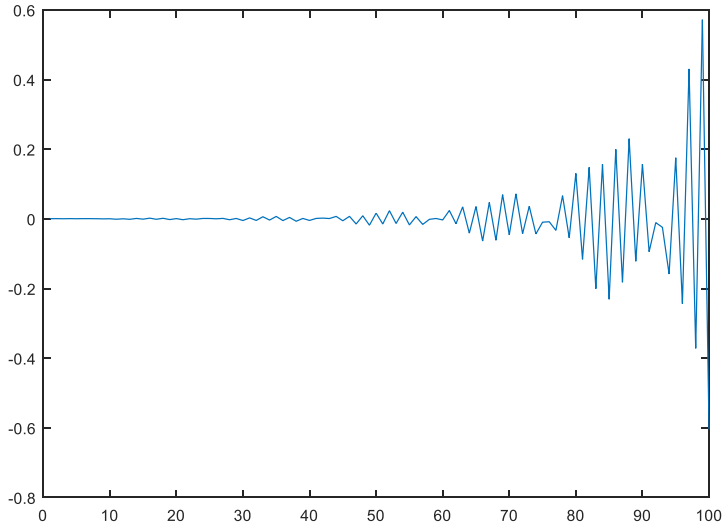
Programming Assignment 3: Earthquake is coming!

Problem Description

- In an earthquake sensing station, the researchers install a very sensitive sensor to do the early detection of earthquake events. An informal research reports that every earthquake actually starts from micro-level vibrations. If the sensor can capture the signal of the micro-level vibrations at the very beginning, and do a prediction on the growing scale of vibrations, then people may have more time to evacuate before an earthquake causes large damages.
- Suppose that you are sent a series of vibration signals from the sensitive sensor. The signal looks like the following plot. You are required to build a prediction model from the signal (download the data here) and predict when the vibration scale will exceed 10^{10} which is Leveled as 10 according to the Richter magnitude scale.



Hint

- You can use the least mean square error method to build a r-order predictor. A r-order predictor assumes the following prediction model:

$$y(t+1) = a_0 + a_1 y(t-r-1) + a_2 y(t-r-2) + \ldots + a_{r-1} y(t-1) + a_r y(t)$$

  where y(t) denotes the predicted scale at time t and a_0, a_1, …, a_r are the prediction coefficients. Using the model and the received data y(0)~y(100), you can estimate the prediction coefficients by the following linear system, which can be solved by the least mean square error (LMSE) method.

$$\begin{bmatrix} 1 & y(1) & \ldots & y(r) \\ 1 & y(2) & \ldots & y(r+1) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & y(k) & \ldots & y(r+k-1) \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_r \end{bmatrix} = \begin{bmatrix} y(r+1) \\ y(r+2) \\ \vdots \\ y(r+k) \end{bmatrix}$$

- Once you have got the prediction coefficients, you can use the prediction model to predict the incoming signals and the time the signal scale becomes larger than the specified threshold.
- Note: you have to decide the order **r** before estimating the coefficients because a correct value for the order r is very important to the prediction accuracy. One trial-and-error way to find a better order is to try guessing a value for r (ranging between 3-20) and use the leading r-1 samples to estimate the prediction coefficients by the LMSE. Then, use the coefficients to predict the remaining samples. By summing the differences between your predictions and the true data, you can use the differences as the goodness (the smaller, the better) of this predictor. Having tried many different values for the order **r**, choose the best one among them.