Consider a univariate stochastic process y_t which follows a first-order moving average representation

$$y_t = \epsilon_t - \theta \epsilon_{t-1} \tag{1}$$

where $\{\epsilon_t\}$ is an i.i.d. process distributed $\mathcal{N}(0,1)$ and $\theta > 1$

1. Write a program to simulate this process. Simulate 100 periods when $\theta = 2$. (Draw ϵ_{-1} from $\mathcal{N}(0,1)$ and $\theta > 1$)

2. Argue that ϵ_t cannot be expressed as a linear combination of y_{t-j} for $j \geq 0$ where the sum of the squares of the weights is finite. **Hint:** Note that

$$\begin{aligned} \epsilon_t &= y_t + \theta \epsilon_{t-1} \\ &= y_t + \theta y_{t-1} + \theta^2 \epsilon_{t-2} \\ &\vdots \end{aligned}$$

   if you keep repeating what happens to the coeficients on y_{t-j}?

3. Represent (1) using as a state space system using our Kalman filter. What are the matrices A, C, and G? How do they depend on θ? **Hint:** x_t will be a vector.

4. Assuming initial beliefs about ϵ_{t-1} are $\mathcal{N}(0,1)$. Write a program to apply this kalman filter to an arbitray sequence $\{y_t\}_{t=0}^T$ for an arbitray θ.

5. Write a program to compute the log-likelihood of sequence $\{y_t\}_{t=0}^T$ for an arbitray θ. **Hint:** Using the Kalman filter the uncoditional distribution of y_0 is $\mathcal{N}(G\hat{x}_0), \Omega_0)$. The distribution of y_t conditional on the past histories of $y^{t-1} = (y_{t-1}, y_{t-2}, \ldots, y_0)$ is $\mathcal{N}(G\hat{x}_t, \Omega_t)$. The likelihood of a sequence of data $(y_T, \ldots, y_0)$ can then be constructed from the conditional likelihoods

$$f(y_T, \ldots, y_0) = f(y_T | y^{T-1}) f(y_{T-1} | y^{T-2}) \cdots f(y_1 | y_0) f(y_0) \tag{2}$$

   and similarly for the log likelihood.

6. For your simulation in part 1. Plot the log likelihood as a function of θ. Where is it maximized? What happens if the simulation length is 1000 periods?