$Q_8)$

a) $\hat{z}_{t+1} = \theta_1 z_t + \theta_2 z_{t-1} \cdots + \theta_M z_{t-M+1}$

Last M terms of the time series are used to predict the next term of the time series.

$$\hat{z}_{M+1} = \theta_1 z_M + \theta_2 z_{M-1} \cdots + \theta_M z_1$$

$$\hat{z}_{M+2} = \theta_1 z_{M+1} + \theta_2 z_M \cdots + \theta_M z_2$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$\hat{z}_{100} = \theta_1 z_{99} + \theta_2 z_{98} + \cdots + \theta_M z_{100-M}$$

- We use last M terms so M is called the lag of the model.
- We have 100 − M equations to find the appropriate values of $\theta_1, \theta_2 \cdots \theta_M$.
- We can model this as a least square problem where we have to find the least square solution to $A\theta = b$.

b) Where $A = \begin{bmatrix} z_M & z_{M-1} & \cdots & z_1 \\ z_{M+1} & z_M & \cdots & z_2 \\ \vdots & \vdots & & \vdots \\ z_{99} & z_{98} & \cdots & z_{100-M} \end{bmatrix}$

$$A \in \mathbb{R}^{(100-M) \times M}$$

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{bmatrix} \qquad \theta \in \mathbb{R}^{m \times 1}$$

$$b = \begin{bmatrix} z_{m+1} \\ z_{m+2} \\ \vdots \\ z_{100} \end{bmatrix} \qquad b \in \mathbb{R}^{(100-m) \times 1}$$

(c) As we move one step forward in the time series, the previous row is shifted by one position and oldest data is removed and one more ~~dep~~ datapoint is added.

$$R_i = [\, z_{i+m-1} \; z_{i+m-2} \; \cdots \; z_i \,]$$

$$R_{i+1} = [\, z_{i+m} \; z_{i+m-1} \; \cdots \; z_{i+1} \,]$$

As we can see from R_i, oldest datapoint (z_i) is removed and z_{i+m} is added.

We can also observe that all diagonals of the matrix A contain same values.

$$\begin{bmatrix} z_m & z_{m-1} & z_{m-2} & \cdots & z_1 \\ z_{m+1} & z_m & z_{m-1} & \cdots & z_2 \\ z_{m+2} & z_{m+1} & z_m & \cdots & z_3 \\ \vdots & & \ddots & & \vdots \\ z_{99} & z_{98} & z_{97} & \cdots & z_{100-m} \end{bmatrix}$$

(d) As no two rows or coloums of the matrix are the same and we're assuming datapoints do not follow a particular order. we can safely say that

$$\text{rank}(A) = \min(M, 100-M)$$

$M \rightarrow$ number of coloumns

$100-M \rightarrow$ number of rows.