CH5440 - MULTIVARIATE DATA ANALYSIS
ASSIGNMENT-6 SUBMISSION
CH18B035
A GUNAVARDHAN REDDY

Q1.

A) Constructing appropriate data matrix and applying LPCA, we estimate the model coefficients as

$$y^*[k] - \mathbf{0.5493}y^*[k-3] = \mathbf{2.0235}u^*[k-3] - \mathbf{1.9036}u^*[k-5]$$

B) Performing LPCA on 100 sets generated by bootstrap we get -
Mean of the coefficients as

coefficients	Mean	Standard deviation	CI	Significance
y[k]	1.0000	0	[1, 1]	**significant**
y[k-1]	-0.0315	0.0732	[-0.2004, 0.1758]	In-significant
y[k-2]	-0.0845	0.0599	[-0.2212, 0.0812]	In-significant
y[k-3]	-0.5624	0.0588	[-0.6985, -0.3990]	**significant**
y[k-4]	-0.0445	0.0198	[-0.0898, 0.0082]	In-significant
y[k-5]	0.0142	0.0222	[-0.0276, 0.0609]	In-significant
u[k]	-0.0178	0.0383	[-0.1066, 0.0684]	In-significant
u[k-1]	0.0525	0.0374	[-0.0371, 0.1233]	In-significant
u[k-2]	0.0600	0.0395	[-0.0330, 0.1408]	In-significant
u[k-3]	-2.0230	0.0381	[-2.1054, -1.9362]	**significant**
u[k-4]	0.0534	0.1539	[-0.3906, 0.4244]	In-significant
u[k-5]	1.9246	0.1334	[1.5758, 2.2228]	**significant**

Hence, Input order is **5**
Output order is **3**.
So Delay is 5 which are all estimated correctly.
Estimated error variance is obtained to be **0.1369**

C) i)Estimate obtained from lag 10 is **6** using Hypothesis testing
Hence Delay = 10 - 6 + 1 = **5**
Model coefficients estimated as

$$y^*[k] - \mathbf{0.5505}y^*[k-3] = \mathbf{2.0260}u^*[k-3] - \mathbf{1.9026}u^*[k-5]$$

Estimated error variance is obtained to be **0.1361**

ii)Estimate obtained from lag 15 is **11** using Hypothesis testing
Hence Delay = 15 - 11 + 1 = **5**
Model coefficients estimated as

$$y^*[k] - \mathbf{0.5472}y^*[k-3] = \mathbf{2.0238}u^*[k-3] - \mathbf{1.897}u^*[k-5]$$

Estimated error variance is obtained to be **0.1356**

Q2.

A) Given error variance, we can scale the data accordingly - that is divide by there respective error standard deviation.

In the code the matrix is multiplied by inverse of diagonal matrix of error standard deviation. DPCA is applied with lag 10, then from hypothesis testing we get d = 6 then LPCA gives-

Model coefficients estimated as

y*[k] - **0.5825**y*[k-3] = **2.0256**u*[k-3] - **1.9170**u*[k-5]

B) Assuming we know nothing about the model or error variances
We can proceed using iterative PCA combination,

$$\min_{\Sigma_e}(N-L)\log|\hat{\mathbf{A}}^{(i)}\boldsymbol{\Sigma}_e(\hat{\mathbf{A}}^{(i)})^T| + \sum_{k=1}^{N-L}(\mathbf{r}^{(i)}[k])^T(\hat{\mathbf{A}}^{(i)}\boldsymbol{\Sigma}_e(\hat{\mathbf{A}}^{(i)})^T)^{-1}\mathbf{r}^{(i)}[k]$$

With the constraint of only two values for error covariance matrix.
Steps:

- Create the lagged data matrix(with some lag = L), take initial guess of the two error std covariance matrix of data matrix.
- In an iterative procedure with nfact <= nvar-2 (as there are only two errors std unknown to us)varing from nvar-2 to nvar-L in each iteration use fmincon to the above objective function.
- Before that scale the data matrix then perform svd on it, use the last nvar-nfact eigen factors for the model scale it back and perform fmincon.
- Initial estimates are only **two** values. Modifying the objective function as only two values are sent(modification create the expanded form of error matrix inside the objec val function)
- This constraints the error covariance matrix to what we need exactly.
- Exit the loop if the taken nfact corresponds to the null hypothesis rejection.
- Then we know the variances and the value of d(process order). we can find the model.

For **lag = 10**
From this procedure we get estimated error variance = [**0.0848 0.9331**]

- As we obtained error variances we can proceed with scaling and hypothesis testing which gives value of **d = 6 = process order**

Using this estimated variance-
Model coefficients estimated as

y*[k] - **0.5841**y*[k-3] = **1.9961**u*[k-3] - **1.8519**u*[k-5]