CH5440 HW6

Prem Sagar S - AE14B021

May 1, 2017

Problem 1

Part a

To apply KPCA the Kernel chosen was Gaussian kernel.

$$K(x_i, x_j) = exp\left(-\frac{(x_i - x_j)^2}{width}\right)$$

Now PCA is applied in the feature space to get the regression matrix B.

$$p_{sat} = TB$$

where T is the scores matrix.

Using this regression matrix, the estimate for the saturated pressures $\hat{p}_{sat}$ is obtained. The norm of the difference of the true and the estimated saturated pressures is obtained. Also to obtain the regression matrix, we need to know the number of PCs to retain. Using these two as input variables, the KPCR function was written to return the errors.

The least error was observerd for:

- 5 PCs and width=4.6970

The plot below shows the error plot for 5 PCs.

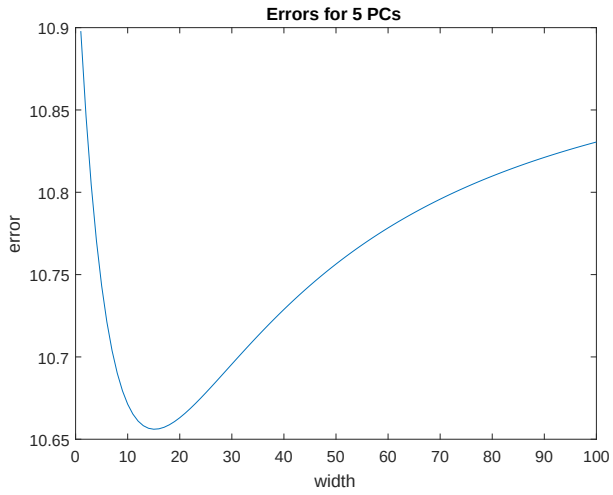


Figure 1: Error plot to obtain optimal width for the optimal PC

Part b

Using the above model, the saturation pressures were obtained for 55^0C and 100^0C .

	T=55	**T=100**
$\hat{p}_{sat}$	63.5442	104.69
p_{sat}	63.96445	246.2437

The predicted saturation pressures are quite close to the true saturation pressure within the provided temperature ranges, i.e., interpolation works fine. But data for T=100 was no given

The same can be observed from the plot below.

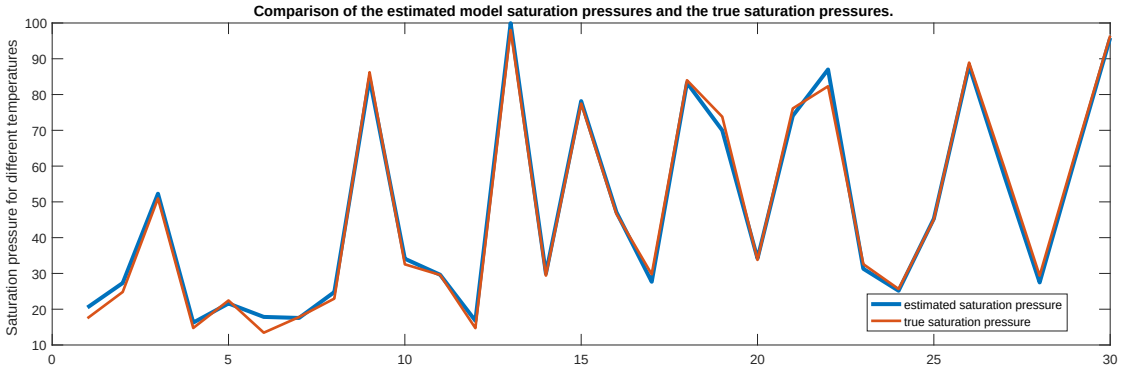


Figure 2: Model comparison with the test data

Problem 2

Part a

The first order ARX model is given by,

$$y_{k+1} = a_0 y_k + b_0 u_k$$

This was calculated using an OLS regression on $[y_{1000:2}]_{1\times 999} = \begin{bmatrix} a_0 & b_0 \end{bmatrix} \begin{bmatrix} y_{999:1} \\ u_{999:1} \end{bmatrix}_{2\times 999}$

$$\begin{bmatrix} a_0 & b_0 \end{bmatrix} = \begin{bmatrix} 0.4776 & 0.7228 \end{bmatrix}$$

Part b

The model in implicit form can be written as:

$$\begin{bmatrix} 1 & -a_1 & -b_1 \end{bmatrix} \begin{bmatrix} y_k \\ y_{k-1} \\ u_{k-1} \end{bmatrix} = 0$$

Assuming the variables have equal variances, the data matrix is constructed as below:

$$A == \begin{bmatrix} y_{1000:2} & y_{999:1} & u_{k:1} \end{bmatrix}$$

A has dimensions (999 samples x 3 variables).

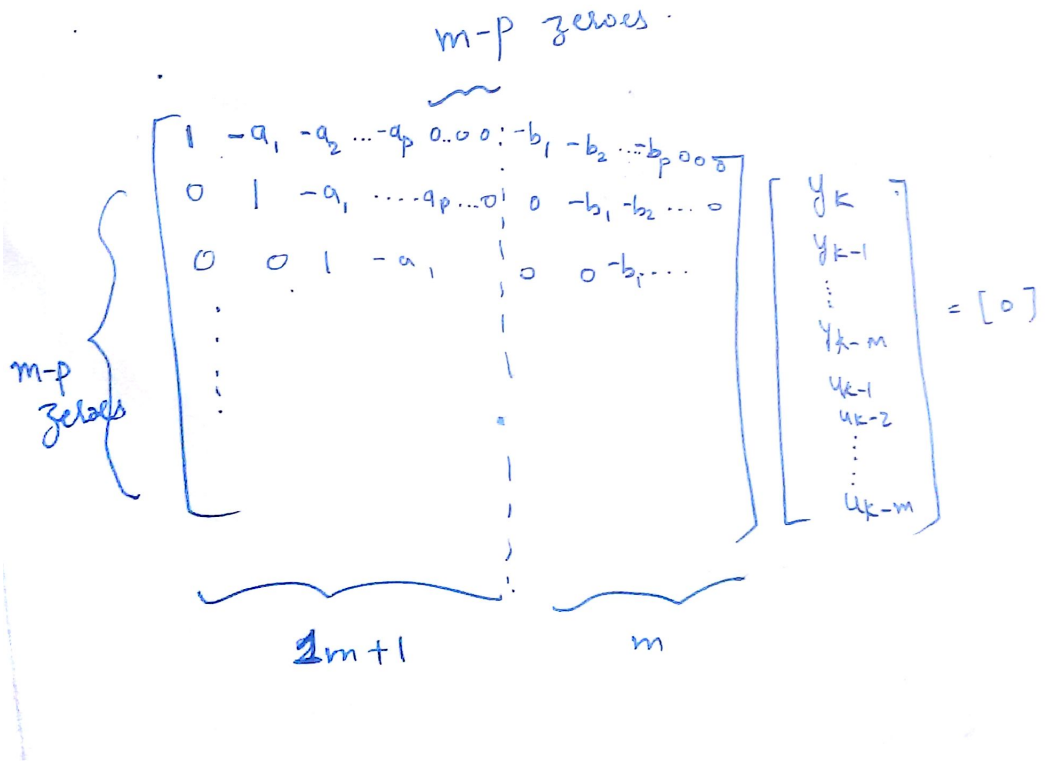
Performing the svd on this and since we know that only one constraint exists, the model parameters can be obtained by taking the last eigen vector.

$$\begin{bmatrix} a_1 & b_1 \end{bmatrix} = \begin{bmatrix} 0.4971 & 0.7430 \end{bmatrix}$$

Part c

Let the stacking order be **m.** The true model order be p.

The constraint matrix can be written as below.



For y_k, the first element of first row is 1 and the next p elements are the model coefficients a_i, followed by **m-p** zeroes. This is followed by p model coefficients for the input b_i, then followed by **m-p** zeroes. Now we can stack the constraint matrix upto **m-p** rows to so as to estimate upto y_{k-m+p}.

The size of the constraint matrix is therefore (m-p+1)×(2m+1).

So the number of constraints is **m-p+1**

If the true order of the model is 1, the number of constraints is m-1+1=**m.**

Part d

The structure of A is as below: ($C\begin{bmatrix} y \\ u \end{bmatrix}=0$)

$$\begin{bmatrix}
1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -b_0
\end{bmatrix}$$

If this is written in explicit form,

$$\hat{A}y = \hat{B}u$$

Structure of A:

$$A_{struc} = \begin{bmatrix} 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -a_0 \end{bmatrix}$$

The data matrix is made by stacking 10 previous outputs and 10 previous inputs for the kth output. This was done starting from 1000:11 outputs. This gives a data matrix of size 21×990. Performing PCA on this and knowing that there are 10-1+1=10 constraints, the implicit constraint matrix $\hat{C}$ (10*21). From this $\hat{A}$ can be obtained by taking the first 11 columns. This is upto a rotation and must be transformed to the structure of A.

$$\hat{A}^T * M = A^T_{trueStruc}$$

$$y^T \hat{A}^T = u^T \hat{B}^T$$

Post multiplying by the rotation matrix,

$$y^T \hat{A}^T M = u^T \hat{B}^T M$$

$$y^T A^T_{trueStruc} = u^T B^T_{trueStruc}$$

$$A_{trueStruc} y = B_{trueStruc} u$$

The rotation matrix was calculated to be:

1	-12.5729	3.1373e-...	7.1801e-...	-0.0206	-0.4630	6.7973e-...	20.5857	-3.9184e...	1.3255
-0.1422	1	-0.4119	-1.3722	-0.0068	-0.5492	0.0487	146.2018	94.4895	-7.9152
0.0206	-9.5972	1	1.8358	0.0089	-0.0805	0.0050	55.1808	30.5003	-2.0460
-0.0433	46.0716	-0.5402	1	-0.0379	-0.6604	-0.0861	22.1981	-301.9435	56.8933
0.0177	22.4700	-0.0720	-0.2623	1	16.4914	0.0154	1.7657	-10.8337	2.1107
-0.0054	3.6651	0.3769	-0.8653	0.0114	1	-1.8799	8.5762	15.2008	-2.8759
-0.0070	-42.5681	-0.2584	-0.3902	0.0022	-9.9571	1	-7.1559	-1.8387	-0.1771
-0.0023	7.7317	-0.2636	81.9052	-0.0822	0.0187	-0.0137	1	4.7448	-0.8138
-0.0047	2.6029e...	0.6416	0.3483	-0.0072	-0.3535	0.0454	0.0464	1	-0.1744
-0.0032	-2.2003e...	30.6079	0.7155	6.2505e-...	-0.0664	0.0295	0.4021	-5.2624	1

Now comparing the rotated A with the true A, we can get the model parameters. But all the elements are not similar to the structure. Hence an average of the corresponding elements were taken to get an estimate for the model parameters.

$$a_0 = 1.5296e - 17$$

$$b_0 = 43.2882$$

This is a very different result obtained as compared to OLS ans TLS in part b here the stacking order was 1. So overstacking results in an erroneous model identification. The main issue why this happens is that it's assumed that the error variances of all the y_i and u_i are equal while this is not true at all.