ME 793 - Assignment 2

Department of Mechanical Engineering, IIT Bombay

Spring 2022

Due Date: 8:30 AM, Jan 31, 2022, Marks 20

Assignment Date: 2:00 PM, Sunday, Jan 23, 2022

Objective and Instructions

1. The objective is to perform principal component analysis (PCA) via first principle and then using a library and compare the results. In this exercise a tabular data is provided. You may find Tutorial 2 and 3 to be good starting points.
2. Submit Jupyter Notebook / Google Colab notebook and the corresponding pdf file to Moodle.
3. For marking Answer No. and to make a commentary, use Markdown cells in your notebook.

Q 1. A dataset with 10 measurements of length, width and thickness of a Si wafer is given.

length	width	thickness
7	4	3
4	1	8
6	3	5
8	6	1
8	5	7
7	2	9
5	3	3
9	5	8
7	4	5
8	2	2

For the following, use the library functions of PCA from *Scikit learn* or *numpy* only for the questions in which you have been asked to do so.

(a) Write a function for determining PCs of the above dataset X. Standardize your data i.e. use zero mean and normalized data using the "Standardize" function shown in Tutorial.

(b) Show the principal vectors i.e. matrix P.

(c) Show the transformed data Y.

(d) Determine the variances along the principal directions.

(e) Determine the principal axes using *Scikit learn* and compare with your solution. Does your solution compare well with that from the python library functions ? Why or why not ?

(f) Determine the principal axes using *numpy* and compare with your solution. Does your solution compare well with that from the python library functions ? Why or why not ?

(g) How many PCs are sufficient to represent the data in reduced dimensions with 95 % accuracy. Show how did you come up with you answer.

–end–