Computer Intensive Statistical Methods STATISTICS NUS
Fall 2017 – 2018 Zhigang YAO

Due on 2 November 2017 **Assignment**

Choose Any Two Exercises and Do Your Best

Exercise 1.
This exercise asks you to perform the SVD example of the handwritten digits 3's that I have shown in lecture (pg 18 of chapter 5). The data can be downloaded in an R package named 'ElemStatLearn'. This is a package that contains most of the data sets that have been used in the book named "The Elements of Statistical Learning".

- You will need to extract the images of 3's from the data 'zip.test' and 'zip.train'. Although the actual number of 3's that you use is not crucial, you need to try your best to extract as many as possible.
- It is suggested to use a two-component model, while you may try some other models.
- Read `https://cran.r-project.org/web/packages/ElemStatLearn/ElemStatLearn.pdf` before doing your analysis.

Exercise 2.
Implement the principal curve algorithm to fit a curve for the included data (principalcurve-data.txt).

- This is a bivariate data. Each data is a point on the 2d plane.
- For your information, the principal curve should look like the curve in the attached file (principal curve sample).

Exercise 3.
Write a program for principal flow to fit a curve for the spherical data (principalcurve-data.csv).

- This is a spherical data. Each data point is a point on the sphere.
- Your principal flow should start from the mean of the data. This means that you will need to write your algorithm to find the mean.
- For your information, the principal flow should look like the black curve in the attached file (principal flow sample).

Your submission should be sent in PDF format to the folder named **Assignment submission** on IVLE before the deadline.

The name of your submission is restricted to the format as ID.pdf, where ID is your Matric No.

You are only allowed to submit one PDF for this assignment