CS 5751 – Spring 2018 – Homework 1
Assigned: 01/16/2018
Due: 01/23/2018
Total points: 100 pts. + BONUS
Submit a soft copy to canvas. Your homework will be graded during the lab session on which it is due.

Objectives: The objectives of this homework are the following:

- Review Taylor series expansions.
- Review basic concepts about multivariate calculus: the gradient, its properties, and the Hessian.
- Setting up Jupyter with R and python.
- Learn how to use R and Python (numpy, pandas, jupyter) for simple exploratory data analysis.

Notes:

- This homework is to be done individually. You may discuss with your classmates, but the work that you write must be your own.
- If you choose Python, then you need to make sure that you have the following packages: numpy, pandas, scikit-learn, matplotlib, jupyter. See the slides posted on canvas to learn how to install python and R. If you are using Linux or Mac, you can install these packages by typing the following at the command prompt: pip install numpy, pandas, sklearn, matplotlib, jupyter.
- I have marked some exercises with (*), (**) or (***) to indicate that they could be slightly more difficult than the others, and that can take more time to solve.

Activity 1: (20pts.) (Taylor Series) Using Python, write a Jupyter notebook named yourLastName_hw1_q1.ipynb containing the answers to the following tasks:

a) (5 pts.) Write down the Taylor series expansion for a *fifth order* Taylor series expansion of $f(x)$ around x_0, where $f(x): \mathbb{R} \longrightarrow \mathbb{R}$. You can write math in a Jupyter notebook if you mark a cell as "markdown" (press escape and then type m, or go to the top menu) and then type your math formula between pairs of dollar signs $$ e^{-x^2}$$.

b) (10 pts.) (*) Write code that replicates Figure 1 below. Your plot needs to have labels for the axes, a title and a legend.

c) (5 pts.) What do you observe in Figure 1? Why is that?

Activity 2: (20pts.) (Derivatives) Consider the function $z = f(x, y) = x^2 + y^2$. Now, write a pdf file named yourlastname_hw1_q2.pdf containing the answers to the following tasks:

i. (5 pts.) Compute the gradient $\nabla f(x, y)$. For this task you need to explain every step of your computation of the gradient. You cannot simply write the vector.

ii. (5 pts.) Compute the Hessian of $f(x, y)$. For this task you need to explain every step of your computation of the Hessian. You cannot simply write it the matrix.

iii. (5 pts.) Write the second order Taylor series expansion for $f(x, y)$ around $x_0 = 0$. Explain your answer.

iv. (5 pts.) Consider Figure 2, which is a plot of $f(x, y)$. What do you observe in this plot? In which direction does the gradient point? Why does this happen?

Note that the pdf that you'll submit for this activity could be either typed in Latex/Word, or you can write it by hand and then scan it. If you do the latter, make sure that the contrast of your PDF file is adequate so as to make it legible.

Activity 3: (20pts.) (Simple Exploratory Data Analysis with R) Using R and the Dow Jones Index Dataset that you can download from the UCI dataset repository web site https://archive.ics.uci.edu/ml/datasets/Dow+Jones+Index, write a Jupyter notebook named yourLastName_hw1_q3.ipynb that performs the following tasks:

a) (10 pts.) Create a matrix of scatter plots showing the correlations between all possible pairs of attributes. Using the scatter plots, comment on which pairs of attributes exhibit the highest correlation. The axes of your plots must be properly labeled.
b) (10 pts.) Plot the correlation matrix. How do you interpret this plot?

Note that in order to use Jupyter with R you need to do a little bit of setup R.

Activity 4: (20pts.) (Simple Exploratory Data Analysis with Python) Write a Jupyter notebook named yourLastName_hw1_q4.ipynb to repeat Activity 3, but using Python and pandas.

Activity 5: (20pts.) (Probability) Suppose we have an unfair die such that Pr(1) = 0.2, Pr(2) = 0.3, Pr(3) = 0.1, Pr(4) = 0.1, Pr(5) = 0.1, Pr(6) = 0.2. Now do the following tasks. Note that you cannot simply write the value; you need to write the complete formula and then solve it step by step:

a) (6 pts.) What is the expected value of the face on which the die will land?
b) (7 pts.) What is the variance of the value of the face on which the die will land?
c) (7 pts.) Suppose that you throw that same die, but you can't see where it landed. Someone else tells you that the top face of the die is even number. What is the probability that the top face is "2"? And the probability that it's "4"? And the probability that it's "2" or "4"?

Bonus: (10 pts.) (**) Using Python or R, plot the function $z = f(x, y) = x^2 + y^2$. Then, in the same figure, plot the gradient. Remember that the gradient is a vector defined at each point (x, y). Your plot should look like Figure 2. For this task the function quiver can help you plot the gradient. Remember that your plot needs to have labels for the axes and a title.

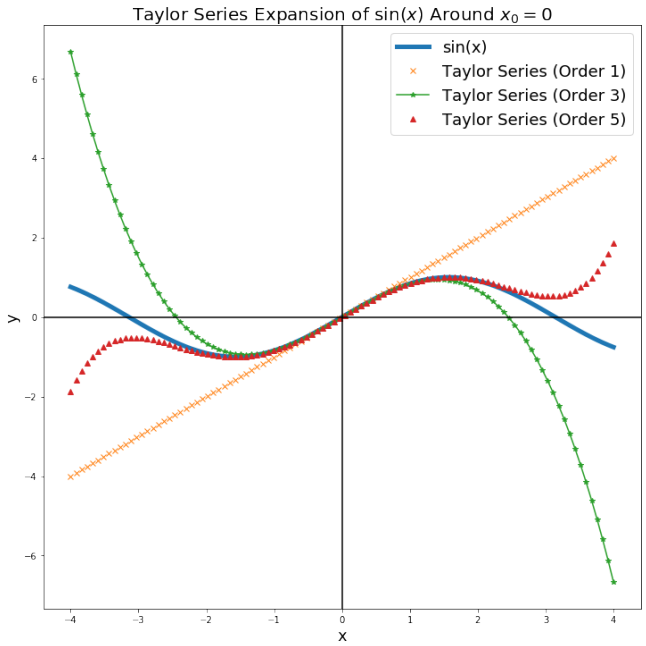

Figure 1. Taylor series expansion of sin(x) around 0

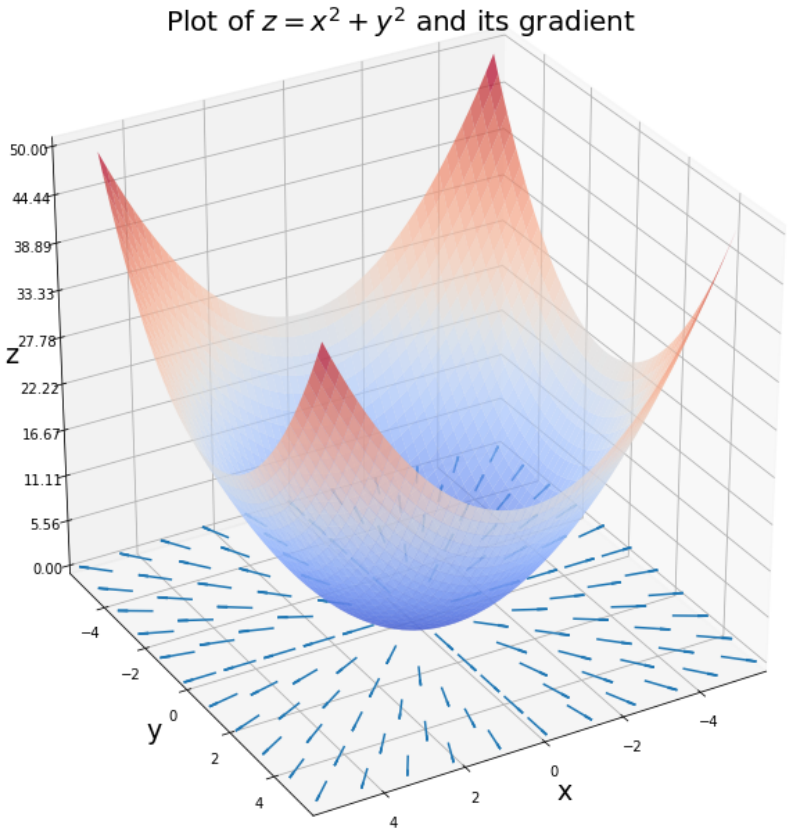

Figure 2. Plot of z=x^2+y^2 and its gradient at the bottom