LAB 5 – Polynomial Regression

In this lab session, we are going to implement polynomial regression with varying degrees using two independent variables: “Age”, and “Skill”. The target (or response) variable is once again “Salary”. We are finally going to visualize the results on a 3-dimensional plane.

- (50 pts) This is a non-linear model, so building the input matrix X is different than before. Implement a function which takes 3 parameters: $x_1,\ x_2$ and k (we could’ve tailored the function to different number of input variables, but we’re limiting ourselves to work on only 2 for ease of implementation). Construct an input matrix X as follows:

  - Compute every possible column of $x_1^i \odot x_2^j$ ($\odot$ stands for element-wise multiplication), where:
    - $0 \leq\ i\ \leq k,$
    - $0 \leq\ j\ \leq k,$
    - $i + j \leq k.$

  - Example: for $k = 3$, the columns would be:
    - $\{x_1^0 x_2^0,\ x_1,\ x_2,\ x_1^2,\ x_1 x_2,\ x_2^2,\ x_1^3,\ x_1^2 x_2,\ x_1 x_2^2,\ x_2^3\}$
      - $x_1^0 x_2^0$ corresponds to a vector of 1’s.

  - Combine every such column to form X, and return it.

- (10 pts) Once X is formed, we can use the standard linear algebra operations from earlier labs to compute predictions. Calculate and display the MSE of these predictions for $k = 2$ and $k = 3$. For both cases, x_1 is the “Age” column and x_2 is the “Skill” column. Here are the numbers:

```
MSE with degree 2: 11428217.365687706
MSE with degree 3: 7255552.463023132
```

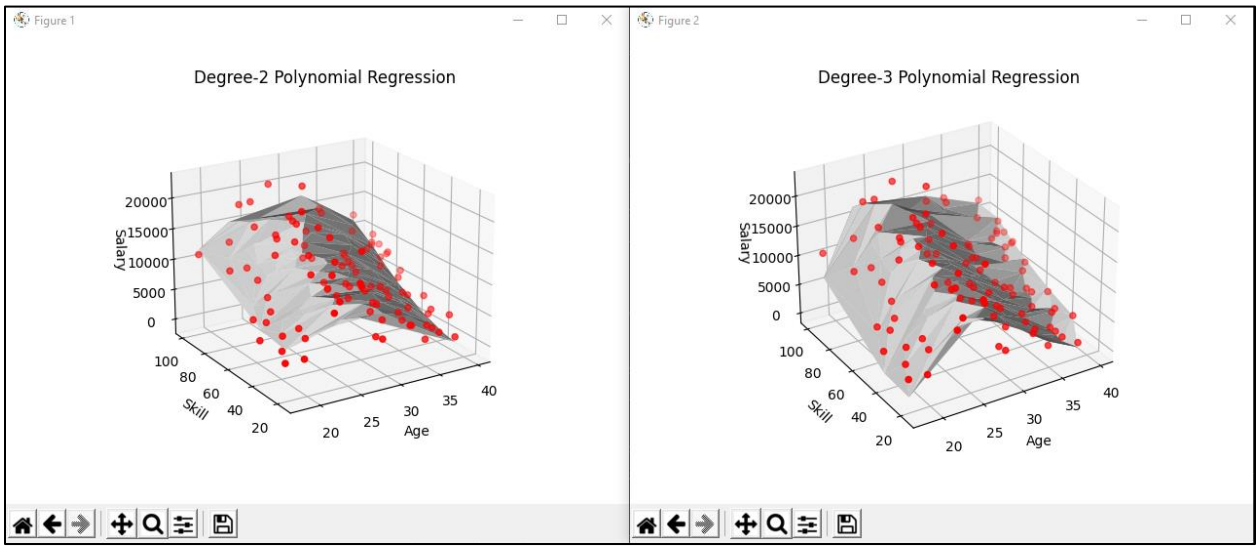
- (40 pts) Plotting:

  Since we have 2 independent variable and 1 response variable, the results can only be visualized in 3-dimensional space. Luckily, plotting elements in this space is very easy using `Matplotlib`. You might have to look at the documentation on 3D plotting to figure out how to plot certain elements in this space.

  Since we have two sets of predictions, do the following for each in separate figures:

  - Create a scatter plot using x_1 (x-axis), x_2 (y-axis) and y (z-axis).
  - Create a triangular surface plot using x_1 (x-axis), x_2 (y-axis) and $\hat{y}$ (z-axis).
    - $\hat{y}$ stands for the predictions you have computed in the previous step.
  - Set axis labels and plot title accordingly.
  - Finally, show the plot. It is important to show both plots simultaneously so we can compare the behavior more clearly.

  Here are how the plots look like (you can rotate these plots in any direction so feel free to look at the results from different angles):



Important note: These instructions require you to implement the computations (except plotting) manually. Any submissions which bypass such calculations will receive 0 points from corresponding parts.