ME698A

HW-4 Report

Name: Omkar Chavan

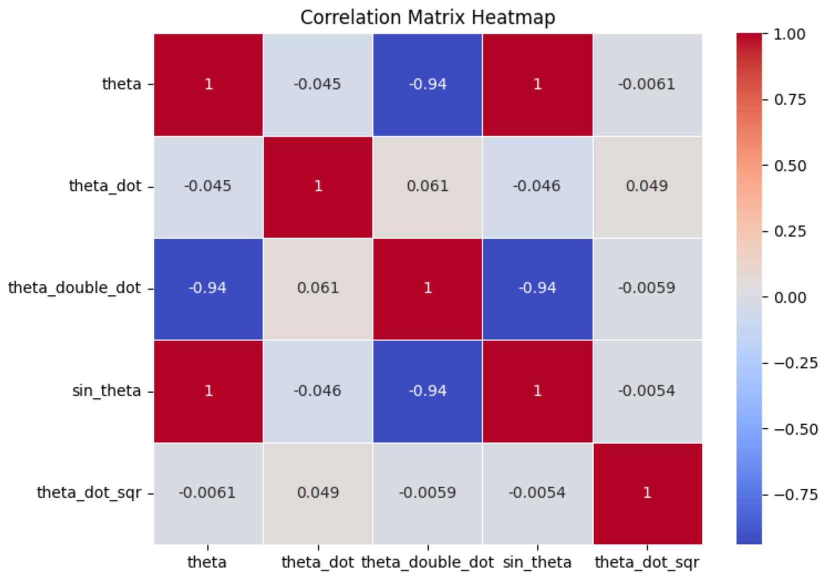
Roll No: 210280

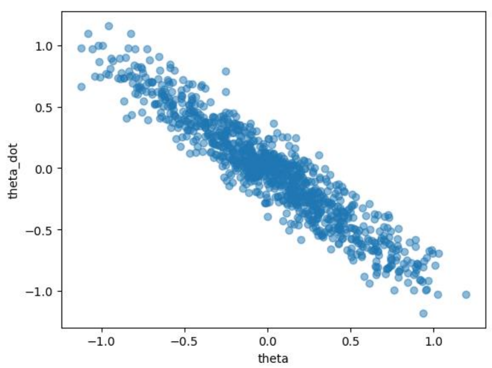
(a) Problem Statement:

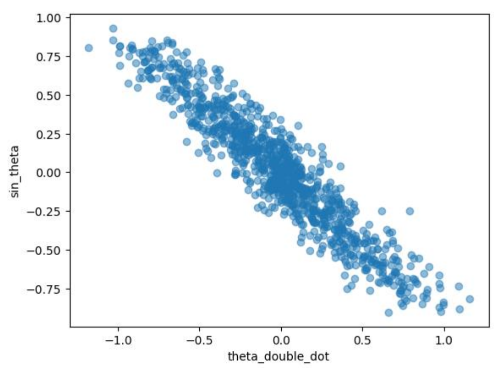
From given data file containing three columns: θ **(angular displacement)**, $\dot{\theta}$ **(angular velocity)** and $\ddot{\theta}$ **(angular acceleration)** at different time instants, finding the governing equation of the simple pendulum from the experimental data.

(b)Plots:

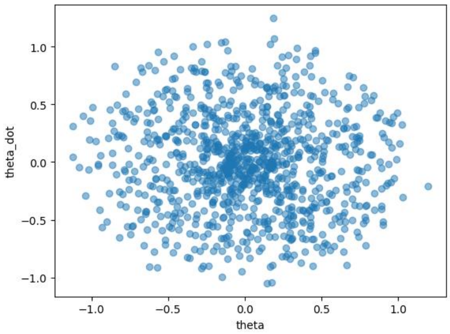
Correlation Heatmap:

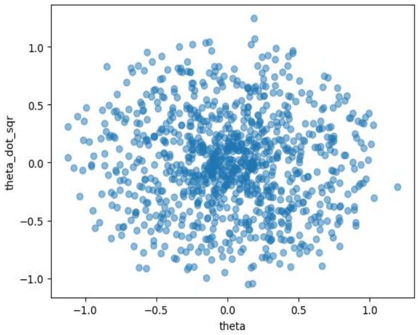


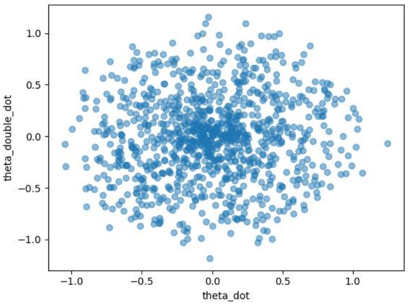




From these scatter plots, we can see that theta_double_dot has linear dependancy on theta and sin_theta. Also theta and sin_thet have similar variation (i.e. they have almost similar values- consistent with small angle approximation)







From these scatter plots we can see that there's no linear relation between theta_double_dot and theta , theta_dot. But there can be some polynomial relation between them.

<u>(c) Solution Procedure:</u>

1. As $\ddot{\theta}$ can be a function of θ, $\dot{\theta}$, $\dot{\theta}^2$ and $\sin\theta$, we will first calculate the values of $\dot{\theta}^2$ and $\sin\theta$.

2. From the correlation heatmap, we can see that there is almost -1 negative correlation between $\ddot{\theta}$ and $\sin\theta$, $\ddot{\theta}$ and θ. Also $\sin\theta$ and θ are also closely related to each other (obvious from small angle approximation). So we can generate model with either one of $\sin\theta$ and θ.

3. Hypothesis Selection:

Selected 3 hypothesis:

(i) Assumed that $\ddot{\theta}$ is a function of θ only. Direct correlation between θ and $\ddot{\theta}$.(i.e. dropping $\dot{\theta}$, $\sin\theta$, $\dot{\theta}^2$ and then fitting our equations)

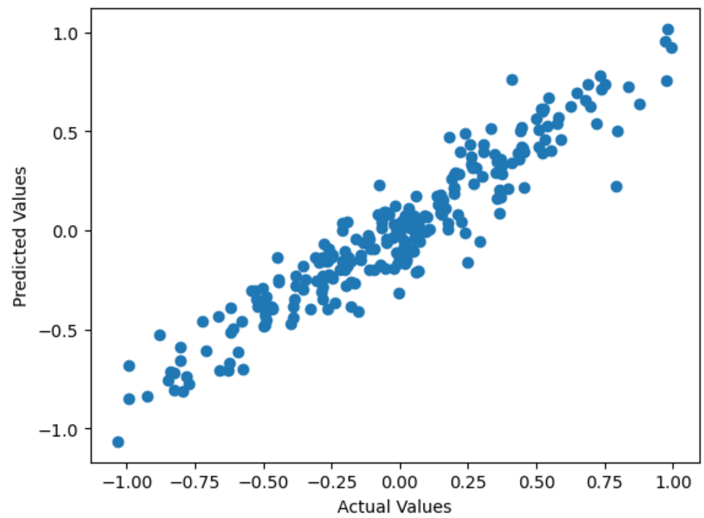
(ii) Assumed that $\ddot{\theta}$ is a function of $\dot{\theta}$, $\sin\theta$, $\dot{\theta}^2$(dropping θ)

(iii) Assumed that $\ddot{\theta}$ is a function of $\sin\theta$ only. Direct correlation between θ and $\ddot{\theta}$.(i.e. dropping $\dot{\theta}$, θ, $\dot{\theta}^2$ and then fitting our equations). As for such a small angles $\sin\theta$ can be approximated as theta and there's strong negative correlation(-0.94) between ***sinθ*** and $\ddot{\boldsymbol{\theta}}$, assuming direct correlation between both(i.e. deleting all columns except $sin\theta$ and $\ddot{\theta}$)

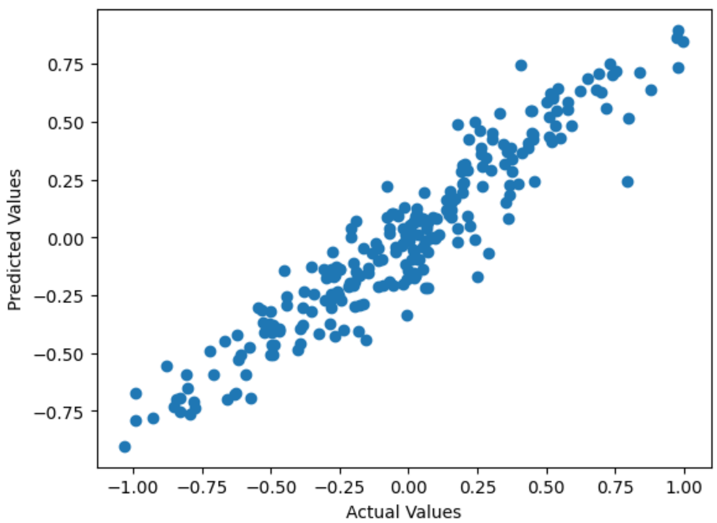
4. For each hypothesis, we have trained a **Linear Regression** model with **Ridge Regularization**. The hyperparameters in our model are **degree of polynomial** and **values of lambda in Ridge Regularization**. To find the optimal value of hyper-parameters, we have use **Grid Search k-fold Cross Validation** with **k=5** on the training data and selected the parameters for which the model has the highest R^2 score.

(d) Results:

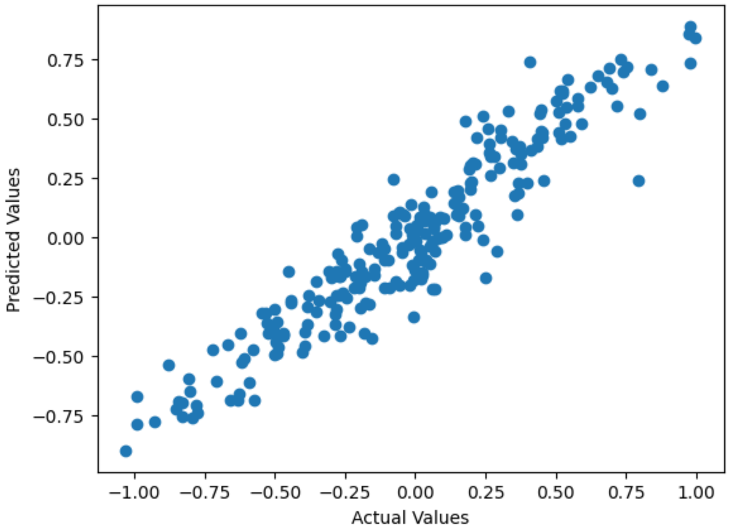
1. Hypothesis 1



2. Hypothesis 2



3. Hypothesis 3:



Hypothesis	Cross Validation Mean Score	MSE
1	0.8844216346284585	0.018396359254199992
2	0.8833787960128191	0.018922838419323882
3	0.8834803527772529	0.018770371034009976

(d) Conclusion:

Clearly hypothesis 1 is the one with highest score and least mean square error.

So in conclusion the model will be:

$\ddot{\theta} = w_0 + w_1 \theta$

This is consistent with actual equation(for small θ) which is:

$$\frac{d^2\theta}{dt^2} + \frac{g}{L}\theta = 0$$

We will get similar results if we had put **sin*θ*,** instead of ***θ*,** as they have close correlation and for small ***θ*, *sinθ*** is almost equal to ***θ*.**