

Project2 Report: Polynomial Regression for GPA Estimation

Simeon Babatunde

Problem Description

We want to create a polynomial hypothesis function that will predict a student's grade point average based on how many minutes per week they study and how many ounces of beer they consume per week. Using the polynomial equation $y = w_0 + w_1x_1 + w_2x_2 + w_3x_1x_2 + w_4x_1^2 + w_5x_2^2$

Data Description

The first line of the dataset contains a single integer indicating how many lines of data are in the file. Each line after that contains three tab-separated real values that represent minutes studying/week, ounces of beer/week, and semester grade point average..

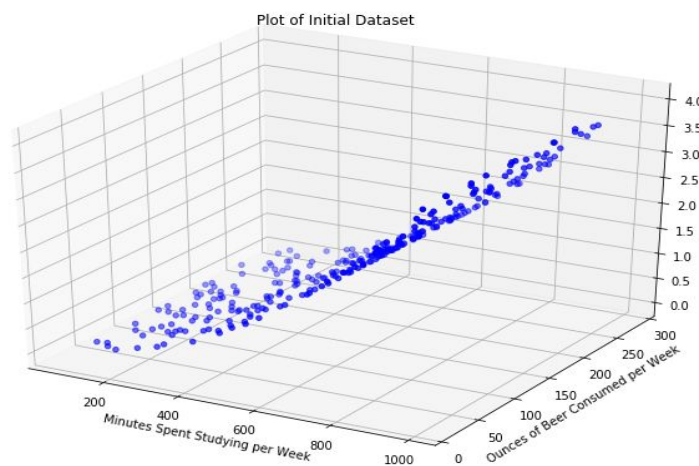


Figure 1. The Initial Dataset

Feature Scaling

In order to speed up gradient descent convergence, it is necessary to ensure the features are on a similar scale. We scaled the dataset using the standardization technique. $(x_i - \mu) / \sigma$. The feature scaling was carried out after randomizing the dataset.

Model Training

While training the regression model, we computed the error $J(w)$ to inform the quality of the weights that were being used in the hypothesis function. Figure 2a shows a plot of error J and number of iterations it takes to minimize the error while Figure 2b shows the tail end of the plot to show the error state before convergence.

The initial values chosen were; $w_0 = 50$ $w_1 = 72$ $w_2 = 48$ $w_3 = 5$ $w_4 = 12$ $w_5 = 10$, $\alpha = 0.1$
 $J=5853.292274732861$

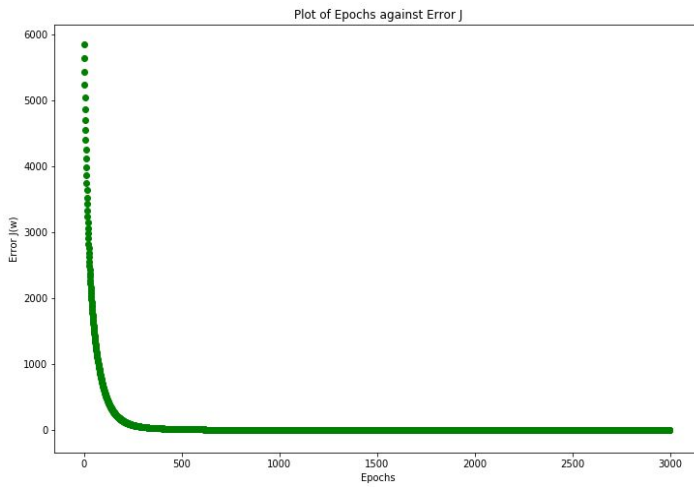


Figure 2a. Error J against Number of Iterations

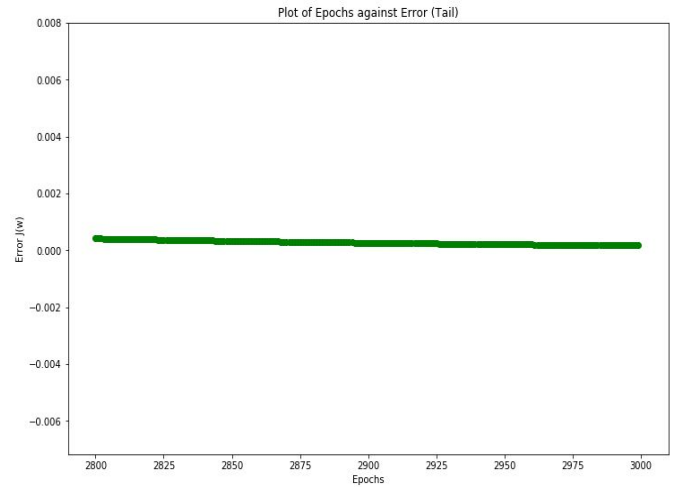


Figure 2b. Tail section of the plot

Results

The final values for $\alpha = 0.01$, $w_0 = 1.204926783686873$, $w_1 = 1.1689073231292413$, $w_2 = -0.0006834056528370415$, $w_3 = -0.0014895007193559378$, $w_4 = 0.27818631239172503$, $w_5 = -0.012474841502375435$. The algorithm went through 3000 iterations to yield a final error value on training data of $J = 0.00017296560478599694$

Finally, the value of J on test data was $J = 0.00015148778367610205$