Batch Gradient Descent
for Predicting CreditScore using Age

Tran Quoc Thai
ID student: 2370759

March 23, 2025

1 Introduction

In this report, we apply batch gradient descent to optimize the linear regression model for predicting `CreditScore` using `Age` as input. The hypothesis function is given by:

$$h_\theta(x) = \theta_0 + \theta_1 x. \tag{1}$$

We start with initial parameters $\theta_0 = 500$ and $\theta_1 = 5$, and perform one iteration of gradient descent with a learning rate $\alpha = 0.01$.

2 Cost Function and Gradients

The cost function for linear regression is given by:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i)^2, \tag{2}$$

where m is the number of training samples, x_i represents the input feature (Age), and y_i represents the target variable (CreditScore).

The gradients for θ_0 and θ_1 are computed as follows:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i), \tag{3}$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x_i) - y_i) x_i. \tag{4}$$

3 Gradient Descent Updates

We update θ_0 and θ_1 using the gradient descent update rule:

$$\theta_j := \theta_j - \alpha \frac{\partial J}{\partial \theta_j}. \tag{5}$$

3.1 Computation of θ_0 Gradient

First, we compute $h_\theta(x)$ for each training sample:

$$\begin{aligned}
h_\theta(35) &= 500 + 5(35) = 675, \\
h_\theta(28) &= 500 + 5(28) = 640, \\
h_\theta(45) &= 500 + 5(45) = 725, \\
h_\theta(31) &= 500 + 5(31) = 655, \\
h_\theta(52) &= 500 + 5(52) = 760, \\
h_\theta(29) &= 500 + 5(29) = 645, \\
h_\theta(42) &= 500 + 5(42) = 710, \\
h_\theta(33) &= 500 + 5(33) = 665.
\end{aligned}$$

The sum of errors is:

$$\begin{aligned}
\sum(h_\theta(x_i) - y_i) &= (675 - 720) + (640 - 650) + (725 - 750) + (655 - 600) \\
&\quad + (760 - 780) + (645 - 630) + (710 - 710) + (665 - 640) \\
&= -45 - 10 - 25 + 55 - 20 + 15 + 0 + 25 \\
&= -5.
\end{aligned}$$

Dividing by $m = 8$:

$$\frac{1}{8}\sum(h_\theta(x_i) - y_i) = \frac{-5}{8} = -0.625. \tag{6}$$

Updating θ_0:

$$\theta_0 := 500 - (0.01 \times -0.625) = 500 + 0.00625 = 500.00625. \tag{7}$$

3.2 Computation of θ_1 Gradient

The weighted sum of errors is:

$$\begin{aligned}
\sum(h_\theta(x_i) - y_i)x_i &= (675 - 720) \cdot 35 + (640 - 650) \cdot 28 + (725 - 750) \cdot 45 \\
&\quad + (655 - 600) \cdot 31 + (760 - 780) \cdot 52 + (645 - 630) \cdot 29 \\
&\quad + (710 - 710) \cdot 42 + (665 - 640) \cdot 33.
\end{aligned}$$

Calculating each term:

$$
\begin{aligned}
(-45) \cdot 35 &= -1575, \\
(-10) \cdot 28 &= -280, \\
(-25) \cdot 45 &= -1125, \\
(+55) \cdot 31 &= 1705, \\
(-20) \cdot 52 &= -1040, \\
(+15) \cdot 29 &= 435, \\
(0) \cdot 42 &= 0, \\
(+25) \cdot 33 &= 825.
\end{aligned}
$$

Summing the terms:

$$-1575 - 280 - 1125 + 1705 - 1040 + 435 + 0 + 825 = -3000 + 2965 = -35. \tag{8}$$

Dividing by $m = 8$:

$$\frac{1}{8} \sum (h_\theta(x_i) - y_i)x_i = \frac{-35}{8} = -3.75. \tag{9}$$

Updating θ_1:

$$\theta_1 := 5 - (0.01 \times -3.75) = 5 + 0.0375 = 5.0375. \tag{10}$$

4 Conclusion

After one iteration of batch gradient descent, the updated parameters are:

$$
\begin{aligned}
\theta_0 &= 500.00625, \\
\theta_1 &= 5.0375.
\end{aligned}
$$

The positive update in both parameters indicates that the initial values underestimated the true relationship between `Age` and `CreditScore`. The updated parameters indicate that θ_0 slightly increases and θ_1 moves in the positive direction. This suggests that the initial model underestimated the baseline CreditScore and underestimated the contribution of Age to CreditScore prediction. Further iterations will refine these values to minimize the cost function.

Batch gradient descent was applied to optimize the parameters of a linear regression model for predicting `CreditScore` using `Age`. The parameter updates indicate an adjustment towards a more accurate prediction model. Repeated iterations will continue refining the model until convergence.