Guidelines for Solving Linear Regression Problems

Eng. Ahmed Métwalli

February 2025

1 Gradient Descent Method

1.1 Step 1: Define the Model

- Assume a simple linear relationship: $Y = aX + b$
  - Identify the dataset and corresponding values of X and Y.

1.2 Step 2: Initialize Parameters

- Choose initial values for a and b (typically set to zero).
  - Set a learning rate α to control step size.
  - Define the number of iterations to update the parameters.

1.3 Step 3: Compute the Loss Function

- Use the Mean Squared Error (MSE) loss function:

$$J(a, b) = \frac{1}{n} \sum_{i=1}^{n} (Y_i - (aX_i + b))^2 \tag{1}$$

1.4 Step 4: Compute Gradients

- Calculate the gradients for a and b:

$$\frac{\partial J}{\partial a} = -\frac{2}{n} \sum X_i (Y_i - (aX_i + b)) \tag{2}$$

$$\frac{\partial J}{\partial b} = -\frac{2}{n} \sum (Y_i - (aX_i + b)) \tag{3}$$

1.5 Step 5: Update Parameters

- Apply the gradient descent update rules:

$$a = a - \alpha \frac{\partial J}{\partial a} \tag{4}$$

$$b = b - \alpha \frac{\partial J}{\partial b} \tag{5}$$

1.6 Step 6: Iterate Until Convergence

- Repeat the update step for the specified number of iterations or until the parameter changes are minimal.

2 Normal Equation Method

2.1 Step 1: Define the Normal Equation

- Express the solution in matrix form:

$$\theta = (X^T X)^{-1} X^T Y \tag{6}$$

2.2 Step 2: Construct the Feature Matrix

- Define the design matrix X by including a column of ones for the intercept term.

2.3 Step 3: Compute Matrices

- Calculate $X^T X$ and $X^T Y$.

2.4 Step 4: Compute the Inverse

- Find $(X^T X)^{-1}$ if it exists.

2.5 Step 5: Compute Parameters

- Multiply the inverse matrix by $X^T Y$ to obtain the coefficients.