Linear Models

Abhishek Swain

2023-09-16

Table of Contents

Linear Regression

The equations

Standard equation

$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$

- $\hat{y}$ is the predicted value.
- n is the number of features.
- x_i is the i^{th} feature value.
- θ_j is the j^{th} model parameter.

Vectorized equation

$\hat{y} = h_\theta(x) = \theta \cdot x$

- θ here is the model's parameter containing $\theta_0 \ldots \theta_n$.
- x is the features containing $x_0 \ldots x_n$ where x_0 is always 1.
- $\theta \cdot x$ is the dot product between θ & x. Both are colum vectors.

$$\hat{y} = \theta \cdot x = \theta^T x = [\theta_0 \ldots \theta_n] \begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n$$

Training the model

We need a loss function to train this model. Out loss funciton is MSE loss.

$MSE(X, h_\theta) = \frac{1}{m} \sum_{i=1}^{m} (\theta^T x^{(i)} - y^{(i)})^2$

Here (i) is the i^{th} training example.

The Normal Equation

Without using any optimization algorithm we also have a direct formula to get the parameters, this formula is the *Normal Equation*

$\hat{\theta} = (X^T X)^{-1} X^T y$

- $\hat{\theta}$ is the estimated parameter, $\hat{\theta} \approx \theta$
- y is is the target vector containing $y^{(1)} \dots y^{(m)}$