Solution Description:

→ Ordinary is the analytical method that can be used to find the optimal weight vector ω for linear regression.

This method is based on the assumption that the response of a linear function:

$$y(x) = \omega^T x + \epsilon$$

where the residual error between model predictions and the true response

$$\varepsilon_i = y_i - \omega^T x_i$$

is normally distributed with mean μ & variance σ^2

The density function $P(y|x, \mu, \sigma^2)$ is also normally distributed with mean μ and variance σ^2. μ is a linear function of x : $\mu = \omega^T x_i$.

It is reasonable to assume that the noise component, σ^2, is the same for the entire dataset.

Assuming each training example is independent & identically distributed (iid) the likely hood function for the dataset is.

$$L(\mu, \sigma^2) = \prod_{i=1}^{m} \left[\left(\frac{1}{2\pi\sigma^2} \right)^{1/2} \cdot \exp\left(- \frac{(y_i - \mu)^2}{2\sigma^2} \right) \right]$$

The optimal w can be determined by applying MLE to RSS, that is by setting $\frac{\partial RSS}{\partial w} = 0$ & solving for w.

In matrix notation :

$$w_{OLS} = (X^T X)^{-1} X^T y$$

This is also known as the normal equation.

⇒ The complete startup code with following tasks is in file :
" Pujitha_kallu_Project-1_linear Regression.ipynb"

⇒ Apply the normal equation to find:

$$w_{OLS} = [w_0, w_1]^T$$

⇒ Print the predicted price of a 5000 square foot house. Remember to "normalize" the square footage first.

⇒ Plot the regression line $\hat{y}_{OLS}$ over the training examples. The line is defined by w_{OLS}