

explain only what i asked , we will move later towards the multiple linear regression using OLS



Below is the strict, complete mathematical derivation of Simple Linear Regression using Ordinary Least Squares (OLS), following exactly the structure you specified.

## 1. Problem Definition

Let

$$\{(x_i, y_i)\}_{i=1}^n, \quad x_i \in \mathbb{R}, y_i \in \mathbb{R}$$

**Inputs**

$$X = \{x_1, x_2, \dots, x_n\}, \quad Y = \{y_1, y_2, \dots, y_n\}$$

**Outputs**

$$\hat{\theta} = (\hat{w}, \hat{b}) \in \mathbb{R}^2$$

**Model Assumption**

$$y_i = wx_i + b + \varepsilon_i$$

**Objective Function**

$$\min_{w,b} J(w, b)$$

where

$$J(w, b) = \sum_{i=1}^n (y_i - (wx_i + b))^2$$

## 2. Mathematical Setup

**Parameters**

$$w \in \mathbb{R}, \quad b \in \mathbb{R}$$

**Prediction Function**

$$\hat{y}_i = f(x_i; w, b) = wx_i + b$$

## Residuals

$$e_i = y_i - \hat{y}_i = y_i - (wx_i + b)$$

## Loss Function

$$J(w, b) = \sum_{i=1}^n e_i^2$$


---

## 3. Derivation of the Core Equations

### 3.1 Expand the Objective Function

$$J(w, b) = \sum_{i=1}^n (y_i - wx_i - b)^2$$


---

### 3.2 Partial Derivative w.r.t. $w$

$$\frac{\partial J}{\partial w} = \sum_{i=1}^n 2(y_i - wx_i - b)(-x_i)$$

$$\frac{\partial J}{\partial w} = -2 \sum_{i=1}^n x_i(y_i - wx_i - b)$$


---

### 3.3 Partial Derivative w.r.t. $b$

$$\frac{\partial J}{\partial b} = \sum_{i=1}^n 2(y_i - wx_i - b)(-1)$$

$$\frac{\partial J}{\partial b} = -2 \sum_{i=1}^n (y_i - wx_i - b)$$


---

## 4. Solving for Optimal Parameters

Set gradients equal to zero:

---

## 4.1 Normal Equation for $w$

$$\sum_{i=1}^n x_i(y_i - wx_i - b) = 0$$

Expand:

$$\sum_{i=1}^n x_i y_i - w \sum_{i=1}^n x_i^2 - b \sum_{i=1}^n x_i = 0$$


---

## 4.2 Normal Equation for $b$

$$\sum_{i=1}^n (y_i - wx_i - b) = 0$$

Expand:

$$\sum_{i=1}^n y_i - w \sum_{i=1}^n x_i - nb = 0$$


---

## 4.3 Solve the System

From second equation:

$$b = \frac{1}{n} \sum_{i=1}^n y_i - w \frac{1}{n} \sum_{i=1}^n x_i$$

Define:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$b = \bar{y} - w\bar{x}$$

Substitute into first equation:

$$\begin{aligned} \sum x_i y_i - w \sum x_i^2 - (\bar{y} - w\bar{x}) \sum x_i &= 0 \\ \sum x_i y_i - w \sum x_i^2 - \bar{y} \sum x_i + w\bar{x} \sum x_i &= 0 \\ w \left( \sum x_i^2 - n\bar{x}^2 \right) &= \sum x_i y_i - n\bar{x}\bar{y} \end{aligned}$$

## 4.4 Closed-Form Solution

$$\hat{w} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{b} = \bar{y} - \hat{w}\bar{x}$$

## 5. Final Formula Summary

### Scalar Form

$$\hat{y}_i = \hat{w}x_i + \hat{b}$$

### Optimal Parameters

$$\hat{w} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \hat{b} = \bar{y} - \hat{w}\bar{x}$$

## Vector / Matrix Form (Simple Linear Case)

Let

$$\mathbf{X} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} w \\ b \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

## 6. Computational Form

### Closed-Form Computation

$$\hat{w} = \frac{\sum (x_i y_i) - n \bar{x} \bar{y}}{\sum (x_i^2) - n \bar{x}^2}$$

$$\hat{b} = \bar{y} - \hat{w}\bar{x}$$

## Gradient-Based Update Rules (Optional)

$$w^{(t+1)} = w^{(t)} - \eta \left( -2 \sum x_i (y_i - w^{(t)} x_i - b^{(t)}) \right)$$

$$b^{(t+1)} = b^{(t)} - \eta \left( -2 \sum (y_i - w^{(t)} x_i - b^{(t)}) \right)$$

---

End of derivation.

When you are ready, we can **extend this derivation to Multiple Linear Regression (OLS)** by lifting scalars to vectors and matrices without changing the underlying optimization structure.

◻ ⚡ ❌ ⚡ ...

nothing

Understood.

◻ ⚡ ❌ ⚡ ...