## 🌟 OLS vs Gradient Descent (Easy Comparison with Formulas)

| Thing | Ordinary Least Squares (OLS) | Gradient Descent |
|---|---|---|
| Idea | Finds best answer **directly** | Finds best answer **step by step** |
| How it works | Uses **math formula once** | Updates values **again and again** |
| Weight update | ❌ No updates | ✅ Updates every step |
| Learning rate | ❌ Not needed | ✅ Needed |
| Best for | Small datasets | Large datasets |
| Used in Deep Learning | ❌ No | ✅ Yes |

---

## 🧮 Formulas

### 📌 OLS (Closed-Form Solution)

**Simple Regression**

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

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

**Multiple Regression (Matrix Form)**

$$\mathbf{w} = (X^T X)^{-1} X^T y$$

📌 Gives **final weights in one shot**.

---

### 📌 Gradient Descent (Iterative Updates)

**Cost Function**

$$J(w,b) = \frac{1}{n}\sum(y - \hat{y})^2$$

**Weight Update**

$$w := w - \alpha\frac{\partial J}{\partial w}$$

**Bias Update**

$$b := b - \alpha\frac{\partial J}{\partial b}$$

📌 Repeats until error becomes small.

## MULTIPLE LINEAR REGRESSION
## ONE ITERATION OF GRADIENT DESCENT (WITH 2 FEATURES)

---

**Problem Setup**

**Features:**
$x_1 = 1$
$x_2 = 2$

**Actual output:**
$y = 5$

**Initial weights and bias:**
$w_1 = 1$
$w_2 = 1$
$b = 0$

**Learning rate:**
$\alpha = 0.1$

---

## Step 1: Prediction ( $\hat{y} = wx + b$ )

Formula:
$\hat{y} = w_1x_1 + w_2x_2 + b$

Calculation:
$\hat{y} = (1 \times 1) + (1 \times 2) + 0$
$\hat{y} = 3$

---

## Step 2: Loss (Mean Squared Error for one sample)

Formula:
$\text{Loss} = (y - \hat{y})^2$

Calculation:
$\text{Loss} = (5 - 3)^2$
$\text{Loss} = 4$

---

## Step 3: Gradients (Partial Derivatives)

**Gradient with respect to $w_1$:**

Formula:
$\partial L/\partial w_1 = -2(y - \hat{y})x_1$

Calculation:
$\partial L/\partial w_1 = -2(5 - 3)(1)$
$\partial L/\partial w_1 = -4$

---

**Gradient with respect to $w_2$:**

Formula:
$\partial L/\partial w_2 = -2(y - \hat{y})x_2$

Calculation:
$\partial L/\partial w_2 = -2(5 - 3)(2)$
$\partial L/\partial w_2 = -8$

---

**Gradient with respect to bias b:**

Formula:
$\partial L/\partial b = -2(y - \hat{y})$

Calculation:
$\partial L/\partial b = -2(5 - 3)$
$\partial L/\partial b = -4$

---

## Step 4: Gradient Descent Update (One Iteration)

**Update $w_1$:**

Formula:
$w_1 = w_1 - \alpha(\partial L/\partial w_1)$

Calculation:
$w_1 = 1 - 0.1(-4)$
$w_1 = 1.4$

---

**Update $w_2$:**

Formula:
$w_2 = w_2 - \alpha(\partial L/\partial w_2)$

Calculation:
$w_2 = 1 - 0.1(-8)$
$w_2 = 1.8$

---

**Update bias b:**

Formula:
$b = b - \alpha(\partial L/\partial b)$

Calculation:
$b = 0 - 0.1(-4)$
$b = 0.4$

---

**Final Parameters After One Iteration**

$w_1 = 1.4$
$w_2 = 1.8$
$b = 0.4$