Goal:

We want to minimize a simple cost function using gradient descent. The function we'll minimize is the **Mean Squared Error (MSE)** for a linear regression model.

Simple Example:

We have a small dataset of 5 points that roughly follow a linear relationship:

- **Dataset**: $X = [1, 2, 3, 4, 5]$ $y = [3, 5, 7, 9, 11]$

We assume the true relationship is:

$$y = 2x + 1$$

Our goal is to find the parameters θ_0 (intercept) and θ_1 (slope) that minimize the error between our model's predictions and the true values of y.

Steps of Gradient Descent:

1. **Initialize Parameters**: Start with random initial values for θ_0 and θ_1. For simplicity, let's initialize them to 0.
2. **Hypothesis Function**: For a linear regression model, the hypothesis function is:

   $$h(x) = \theta_0 + \theta_1 \cdot x$$

   This is our model, and we want to adjust θ_0 and θ_1 to fit the data.

3. **Cost Function**: The cost function we want to minimize is the **Mean Squared Error (MSE)**:

   $$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^{m} (h(x_i) - y_i)^2$$

   where:

   - m is the number of data points (5 in this case).
   - $h(x_i)$ is the predicted value for the i-th data point.

- y_i is the actual value for the i-th data point.

4. **Gradient Calculation**: To update the parameters, we need to compute the **gradients** of the cost function with respect to θ_0 and θ_1:

$$\frac{\partial J}{\partial \theta_0} = \frac{1}{m} \sum_{i=1}^{m} (h(x_i) - y_i)$$

$$\frac{\partial J}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} (h(x_i) - y_i) \cdot x_i$$

5. **Update Parameters**: Using the learning rate α, we update the parameters in the direction that minimizes the cost function:

$$\theta_0 := \theta_0 - \alpha \cdot \frac{\partial J}{\partial \theta_0}$$

$$\theta_1 := \theta_1 - \alpha \cdot \frac{\partial J}{\partial \theta_1}$$

where α is the learning rate, a small value (say 0.1) that controls the size of each step.

6. **Repeat**: Repeat the process for a specified number of iterations (or until the cost stops changing significantly).