**Cost Function**

**Definition:**

* A **cost function** measures the accuracy of the hypothesis function by comparing the predicted values with the actual values.
* Specifically, the **mean squared error** (MSE) is a common cost function used in regression problems.

**Mathematical Expression:**

* The cost function J(θ0,θ1)J(θ0​,θ1​) is defined as:

J(θ0,θ1)=12m∑i=1m(y^i−yi)2=12m∑i=1m(hθ(xi)−yi)2J(θ0​,θ1​)=2m1​i=1∑m​(y^​i−yi)2=2m1​i=1∑m​(hθ​(xi)−yi)2

* Here, mm is the number of training examples, y^iy^​i is the predicted value, and yiyi is the actual value.
* hθ(xi)hθ​(xi) represents the hypothesis function with input xixi and parameters θ0,θ1θ0​,θ1​.

**Components:**

* **Squared Error**: The difference between the predicted value and the actual value is squared: (hθ(xi)−yi)2(hθ​(xi)−yi)2.
* **Mean**: The sum of the squared errors is divided by 2m2m to calculate the mean.
* **Halving Factor**: The mean is halved (multiplied by 1221​) as a computational convenience, especially during gradient descent.

**Interpretation:**

* The cost function quantifies how well the hypothesis function is performing by calculating the average squared difference between predicted and actual values.
* A lower cost indicates better performance, meaning the predictions are close to the actual values.
* It is called "squared error function" or "mean squared error" in different contexts.

**Conclusion:**

The cost function plays a vital role in supervised learning, guiding the optimization of the model's parameters. Specifically, the mean squared error provides a quantitative measure of how well the model's predictions align with the actual outcomes. It serves as the objective to be minimized during the training process, such as in gradient descent, to find the best-fitting model. Understanding the cost function is fundamental to both the implementation and interpretation of machine learning algorithms.