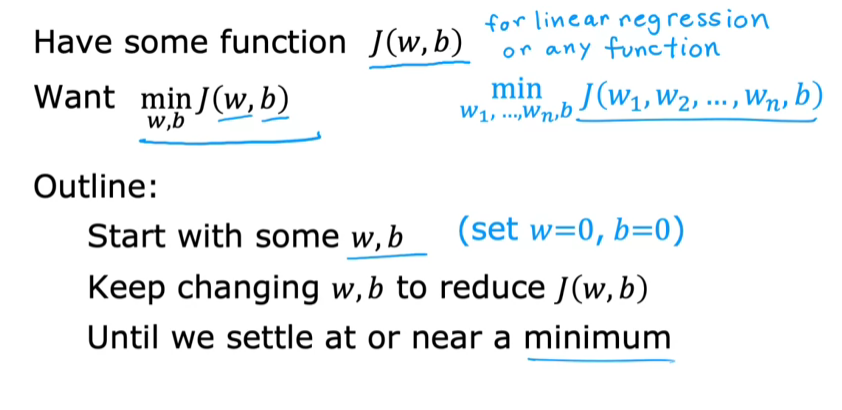
***GRADIENT DESCENT***

UNDERSTANDING GRADIENT DESCENT

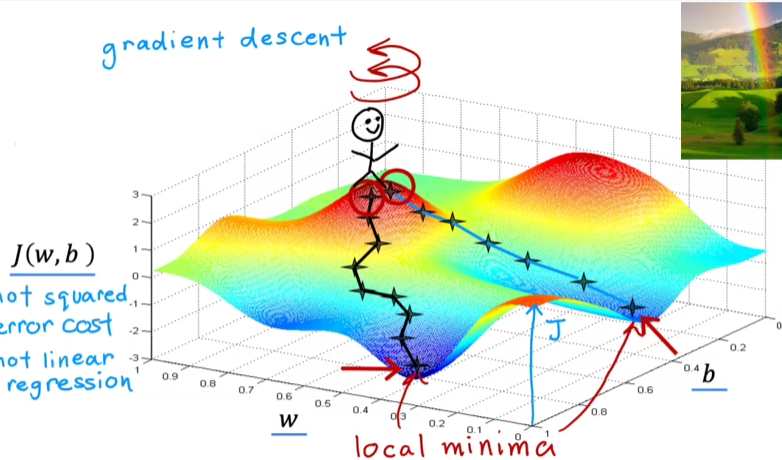
* **Gradient descent is an algorithm used to find the values of parameters (w and b) that minimize the cost function (j of w, b).**
* **It can be applied to various functions, not just linear regression, making it a versatile tool in machine learning.**

THE PROCESS OF GRADIENT DESCENT



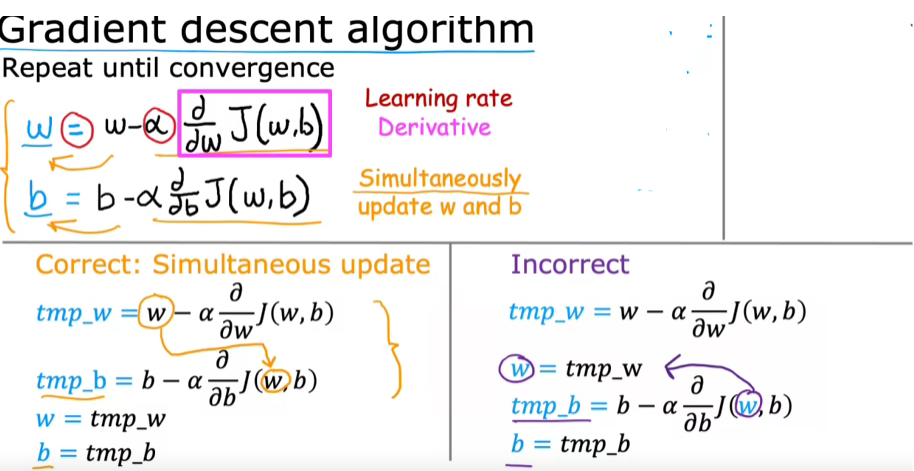
* **Start with initial guesses for parameters, commonly set to zero, and iteratively adjust them to reduce the cost function.**
* **The algorithm involves taking small steps in the direction of the steepest descent, akin to walking downhill on a hilly surface.**

**Example:**



**THE GRADIENT DESCENT ALGORITHM**

* **The gradient descent algorithm updates the parameter (w) by taking its current value and adjusting it using the learning rate (alpha) and the derivative of the cost function (J).**
* **The learning rate (alpha) controls the size of the steps taken during the optimization process, with smaller values leading to more gradual changes.**



* **Both parameters (w) and (b) are updated simultaneously to ensure that the changes are based on the same values from the previous iteration.**
* **The correct implementation involves calculating temporary values for both parameters before updating them, preventing any discrepancies in the calculations.**
* **The algorithm continues to update the parameters until it converges, meaning that the values of (w) and (b) stabilize and do not change significantly with further iterations.**

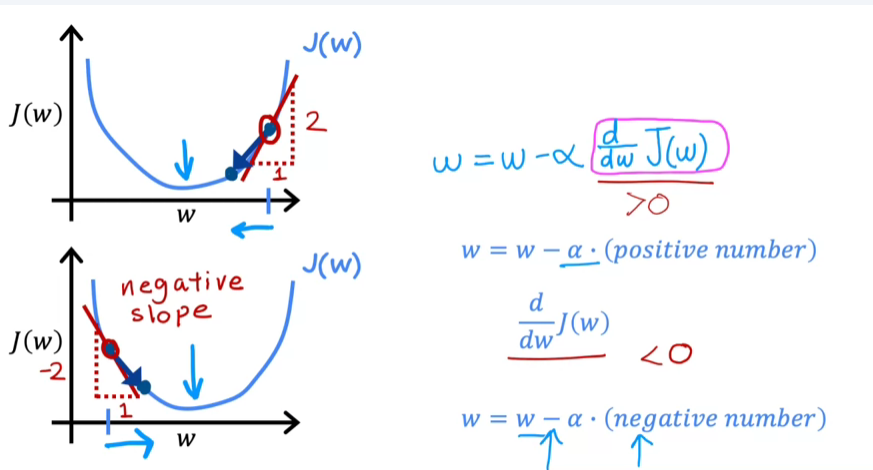
**THE INTUITION**

INTUITION BEHIND PARAMETER UPDATES

* **When the derivative is positive, the parameter w is decreased, moving left on the graph, which reduces the cost J, indicating progress toward the minimum.**
* **Conversely, if the derivative is negative, the parameter w is increased, moving right on the graph, which also leads to a decrease in cost J, demonstrating that gradient descent is effectively navigating toward the minimum.**

CHOOSING THE LEARNING RATE

* **The learning rate (Alpha) is crucial; if it's too small, updates will be slow, while if it's too large, it may overshoot the minimum. Understanding how to select an appropriate value for Alpha is essential for effective gradient descent implementation.**



**MORE ABOUT THE LEARNING RATE**

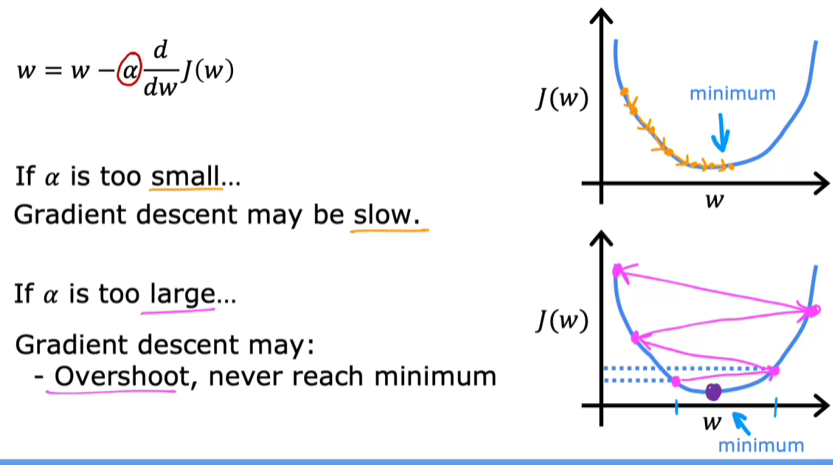
* **If the learning rate is too small, gradient descent will converge to the minimum, but very slowly, requiring many tiny steps.**

**Effects of a Small Learning Rate**

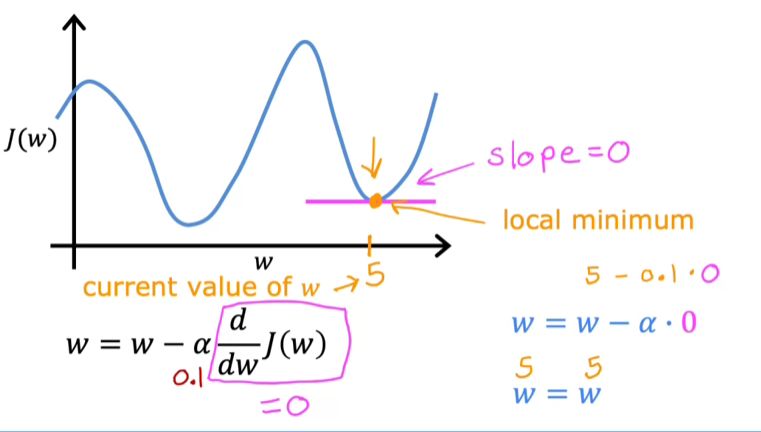
* **With a small learning rate, each update results in minimal changes, leading to a slow decrease in the cost function.**
* **This slow convergence means that it may take a long time to approach the minimum, which can be inefficient.**

**Consequences of a Large Learning Rate**

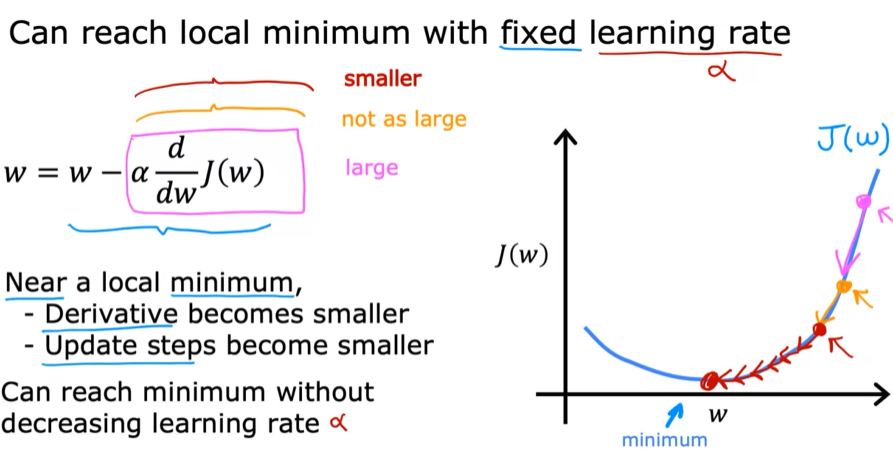
* **Conversely, if the learning rate is too large, gradient descent may overshoot the minimum, causing the cost to increase instead of decrease.**
* **A large learning rate can lead to divergence, where the algorithm fails to converge to a solution, moving further away from the minimum with each step.**



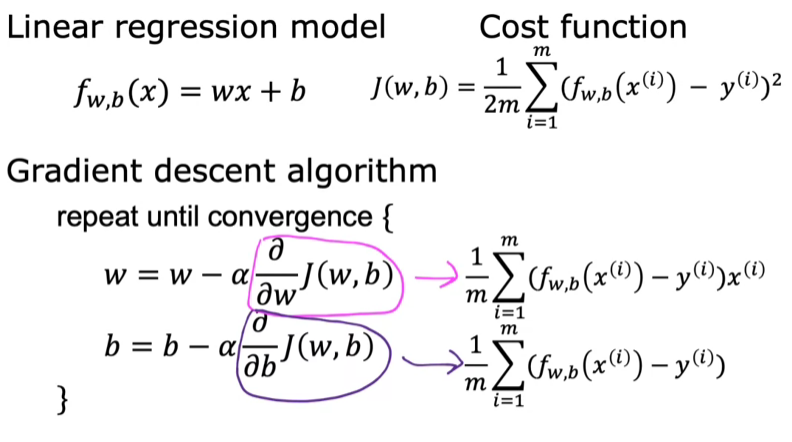
**Check out this J(w) vs. w plot:**



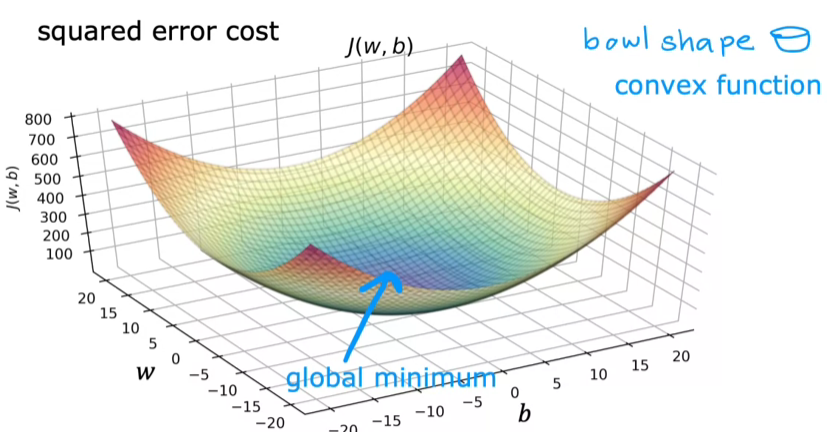
**Even though if we keep the alpha fixed, the decrease in w might become slow because of low gradient value.**



**GRADIENT DESCENT FOR LINEAR REGRESSION MODEL**

****

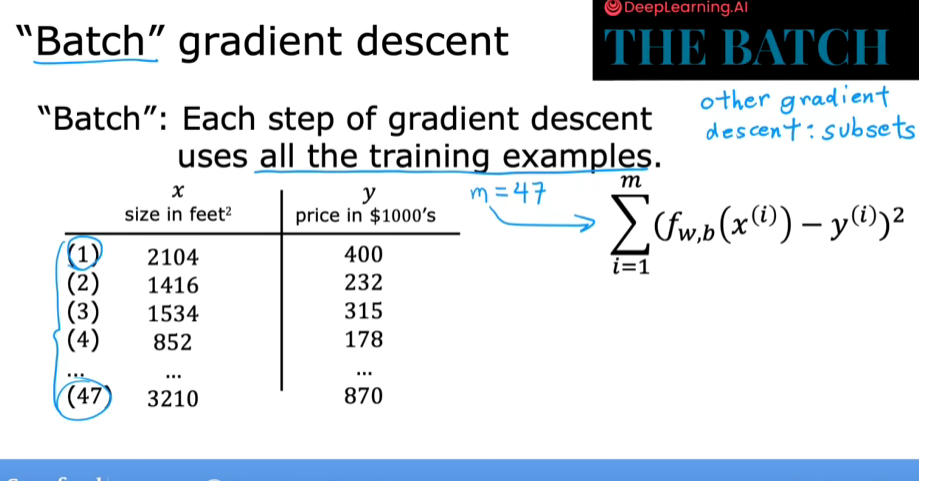
* The squared error cost function is a convex function, meaning it has a single global minimum, which ensures that gradient descent will converge to the optimal solution as long as the learning rate is appropriately chosen.



**RUNNING GRADIENT DESCENT MODEL**

BATCH GRADIENT DESCENT

* Batch gradient descent refers to using the entire training dataset to compute the updates for the parameters at each step, ensuring a comprehensive approach to optimization.
* This method contrasts with other versions of gradient descent that utilize smaller subsets of data for updates.



* Once the model is trained, it can be used to make predictions, such as estimating house prices based on size.