

Gradient Decent Pseudocode

Step	Description
1	Initialize parameters $\theta_0, \theta_1, \dots, \theta_n$ randomly or with zeros
2	Set the learning rate α and number of iterations T
3	For each iteration $t = 1$ to T :
4	Compute the cost function: $J(\theta) = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$
5	For each parameter θ_j (for $j = 0, 1, \dots, n$):
6	Compute the gradient:
7	$\frac{\partial J(\theta)}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)}$
8	Update the parameter: $\theta_j := \theta_j - \alpha \frac{\partial J(\theta)}{\partial \theta_j}$
9	Repeat steps 4-8 until convergence or for T iterations
10	Return the optimized parameters θ