





Back to Week 2

imes Lessons

This Course: Machine Learning

Prev

Next

Gradient Descent For Multiple Variables

Gradient Descent for Multiple Variables

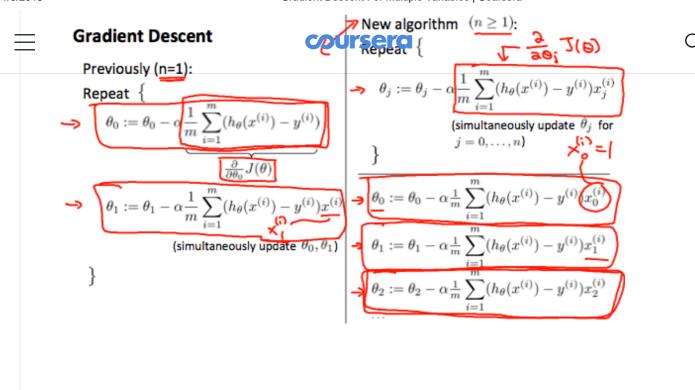
The gradient descent equation itself is generally the same form; we just have to repeat it for our 'n' features:

$$egin{aligned} ext{repeat until convergence: } \{ \ heta_0 &:= heta_0 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_0^{(i)} \ heta_1 &:= heta_1 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_1^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_1 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_2^{(i)} \ heta_2 &:= heta_2 - a \, heta_2 - a \, heta_2 &:= heta_2 - a \, heta_2 - a \, heta_2 &:= heta_2 - a \, heta_2 &:= heta_2 - a \, heta_2 &:= hea_2 - a \, heta_2 &:= hea_2 - a \, heta_2 &:= hea_2 - a \, heta_2$$

In other words:

$$egin{aligned} ext{repeat until convergence: } \{ \ heta_j := heta_j - lpha \, rac{1}{m} \sum_{i=1}^m (h_ heta(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)} \qquad ext{for j} := 0... n \end{aligned}$$

The following image compares gradient descent with one variable to gradient descent with multiple variables:



✓ Complete

