Observations - actual real-world data

$x^i, y^i = [0, 0]\,[1, 2]\,[2, 1]\,[3, 7]\,[4, 4]\,[5, 12]$

Prediction - for a given x, what is y?

$prediction = mx + b$

Error (aka "Cost Function") - how wrong am I?

$error^i = (prediction^i - observation^i)^2$

$$\sum_{i=1}^{m} error^i = \sum_{i=1}^{6} (prediction^i - observation^i)^2$$

Chain Rule... written the easy way.

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$$

(thank you, Leibniz)

Partial Derivatives - which way should I go?

$error^i = (prediction^i - observation^i)^2$

$error^i = (mx^i + b - observation^i)^2$

(using the chain rule...)

$derror/dm = 2 \cdot (mx^i + b - observation^i) \cdot x^i$

$derror/db = 2 \cdot (mx^i + b - observation^i) \cdot 1$