Squared Errors

\[ E = \frac{1}{2} \sum \sum [ y^m - \hat{y}^m ]^2 \]  
(SSE)  
sum of squared errors

Gradient is another word for slope or rate of change.  
Gradient descent can get into "local minima".  
One way to solve is "momentum".

error term: \[ S = (y - \hat{y})f'(h) \]  
activation function

\[ w_i = w_i + \eta \delta x_i \]

\[ h = \sum w_i x_i \]

Mean Squared Error (MSE)

\[ E = \frac{1}{2m} \sum \sum (y^m - \hat{y}^m)^2 \]