

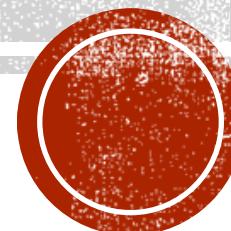
A landscape photograph of a rocky mountain peak. Three colored paths are overlaid on the image: a red path, a green path, and a purple path. The paths represent different gradient descent methods: Batch Gradient Descent, Mini-batch Gradient Descent, and Stochastic Gradient Descent respectively. The paths start from the bottom left and follow the contours of the mountain peak towards the top right.

- Batch Gradient
- Mini-batch Gradient
- Stochastic Gradient



بسم الله الرحمن الرحيم

GRADIENT DESCENT



$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$



Target = 10

Predicted = -10

Loss = 20

w₁ = 3

w₂ = 4

w₃ = 5

$$\frac{dz}{dw_1} = \frac{dz}{dy} * \frac{dy}{dg} * \frac{dg}{dw_1} = -10$$

$$\frac{dz}{dw_2} = \frac{dz}{dy} * \frac{dy}{dw_2} = w_3 * 1 = w_3 = 5$$

$$\frac{dz}{dw_3} = y = w_2 + g = w_2 - 2 * w_1 = 4 - 2 * 3 = -2$$

Gradient Descent STARTS...



$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$



Target = 10

$$dz / dw_1 = -10$$

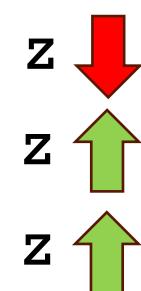
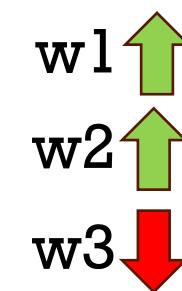
affect of w_1 on z

$$dz / dw_2 = 5$$

affect of w_2 on z

$$dz / dw_3 = -2$$

affect of w_3 on z



GOAL!!

Loss = 20



Predicted = -10



$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$



Target = 10

$w_1 = 3$

$w_2 = 4$

$w_3 = 5$

Predicted = -10



$$dz / dw_1 = -10$$

$$dz / dw_2 = 5$$

$$dz / dw_3 = -2$$

$w_1 \downarrow$
 $w_2 \uparrow$
 $w_3 \downarrow$

Learning rate

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

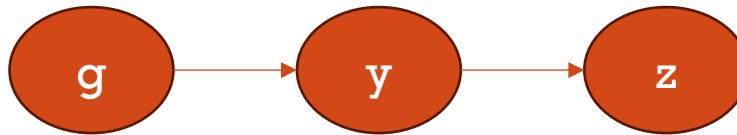
$$w_3 = 5 + 0.01 * (-2) = 4.9$$

$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$



Target = 10

w₁ = 3

w₂ = 4

w₃ = 5

Predicted = $z = (4.1 - 2 * 2.9) * 4.9 = -8.33$

This is one step!!

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$



$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$

1 → Calculate the prediction



Target = 10

Predicted = -8.33

$w_1 = 3$

$w_2 = 4$

$w_3 = 5$

$$dz / dw_1 = -10$$

$$dz / dw_2 = 5$$

$$dz / dw_3 = -2$$

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$

$$g = -2 * w_1$$

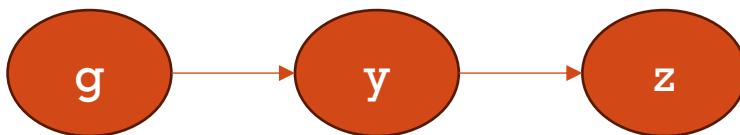
$$y = w_2 + g$$

$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$

1  Calculate the **prediction**

2  Calculate the **loss**  **Loss** = target - **predicted**



Target = 10

Predicted = -8.33

w₁ = 3

w₂ = 4

w₃ = 5

$$dz / dw_1 = -10$$

$$dz / dw_2 = 5$$

$$dz / dw_3 = -2$$

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$



$$g = -2 * w_1$$

$$y = w_2 + g$$

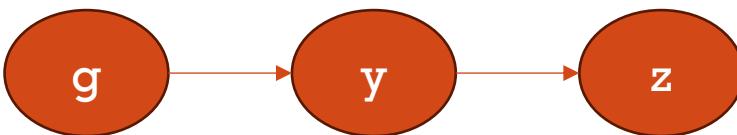
$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$

1 → Calculate the prediction

2 → Calculate the loss

3 → backpropagation



Target = 10

Predicted = -8.33

$w_1 = 3$

$w_2 = 4$

$w_3 = 5$

dz / dw_1

dz / dw_2

dz / dw_3



$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$



$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$

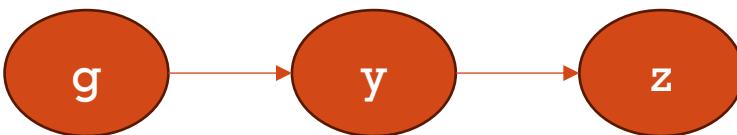
$$z = (w_2 + -2 * w_1) * w_3$$

1 → Calculate the prediction

2 → Calculate the loss

3 → backpropagation

4 → Gradient Descent



Target = 10

Predicted = -8.33

w1 = 3

w2 = 4

w3 = 5

$$dz / dw_1$$

$$dz / dw_2$$

$$dz / dw_3$$

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$

$$g = -2 * w_1$$

$$y = w_2 + g$$

$$z = y * w_3$$

$$z = (w_2 + -2 * w_1) * w_3$$



Target = 10

Predicted = -8.33

$w_1 = 3$

$w_2 = 4$

$w_3 = 5$

- 1 → Calculate the prediction
 - 2 → Calculate the loss
 - 3 → backpropagation
 - 4 → Gradient Descent
-
- A red bracket on the right side of the list groups the last two items (backpropagation and Gradient Descent) under the heading 'dz / dw1', 'dz / dw2', and 'dz / dw3'.

dz / dw_1

dz / dw_2

dz / dw_3

$$w_1 = 3 + 0.01 * (-10) = 2.9$$

$$w_2 = 4 + 0.01 * 5 = 4.1$$

$$w_3 = 5 + 0.01 * (-2) = 4.9$$

NEXT... Neural Networks...

Practically

