

Chapter 12

Gradient

*The input story
dot em with weights*

*This interim called zee
goes for future sets*

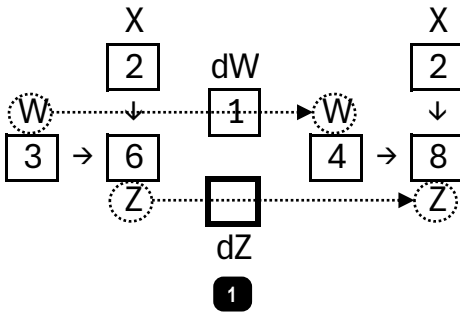
*When the change in weight
is the order from the end*

*The processed input
show a slight different trend*

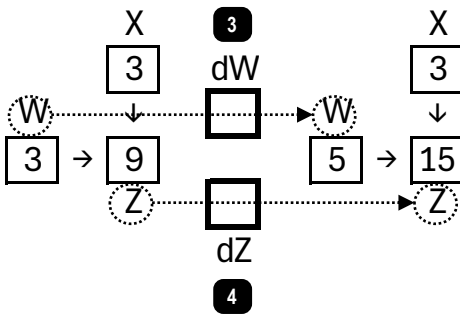
*How much change
of my trend may i know*

*Do the gradient math
and it will show!*

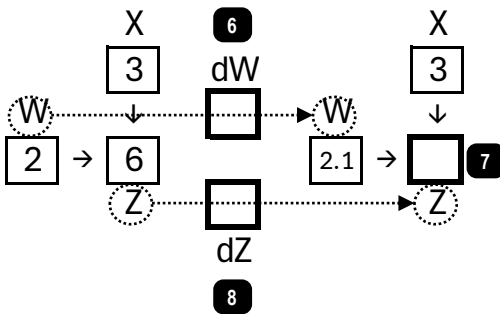
Chapter 12 Gradient - Deep Learning Math Workbook



$$\frac{dZ}{dW} = \frac{\boxed{1}}{\boxed{1}} = \boxed{2}$$



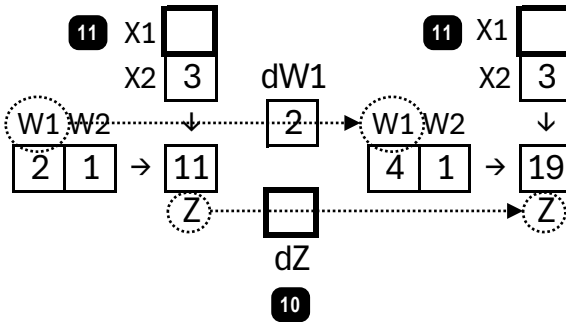
$$\frac{dZ}{dW} = \frac{\boxed{4}}{\boxed{3}} = \boxed{5}$$



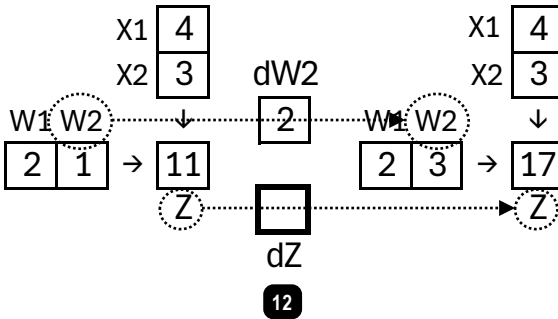
$$\frac{dZ}{dW} = \frac{\boxed{8}}{\boxed{6}} = \boxed{9}$$

1	2	3	4	5	6	7	8	9
2	2	2	6	3	0.1	6.3	0.3	3

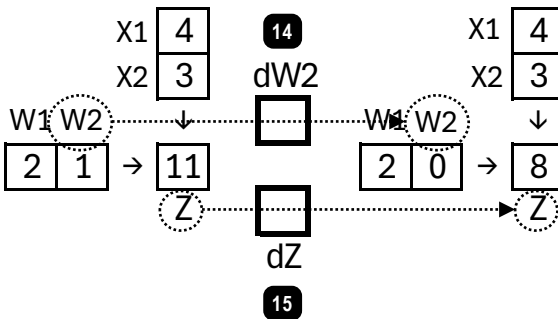
Chapter 12 Gradient - Deep Learning Math Workbook



$$\frac{dZ}{dW_1} = \frac{\boxed{10}}{\boxed{2}} = \boxed{5}$$



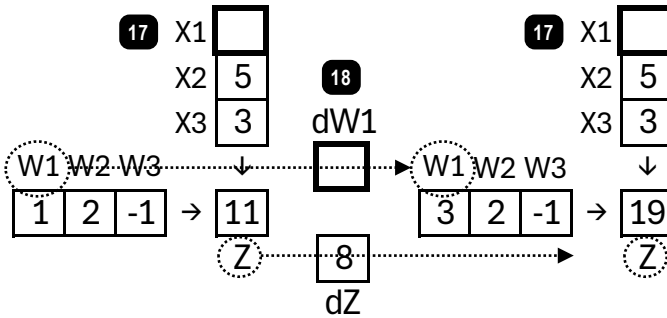
$$\frac{dZ}{dW_2} = \frac{\boxed{12}}{\boxed{2}} = \boxed{6}$$



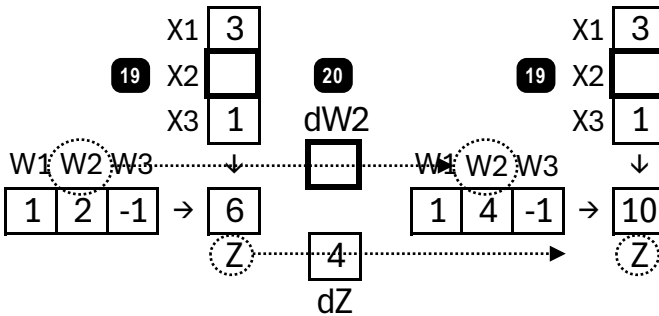
$$\frac{dZ}{dW_2} = \frac{\boxed{15}}{\boxed{14}} = \boxed{1.07}$$

10	11	12	13	14	15	16
∞	4	6	3	1	5	3

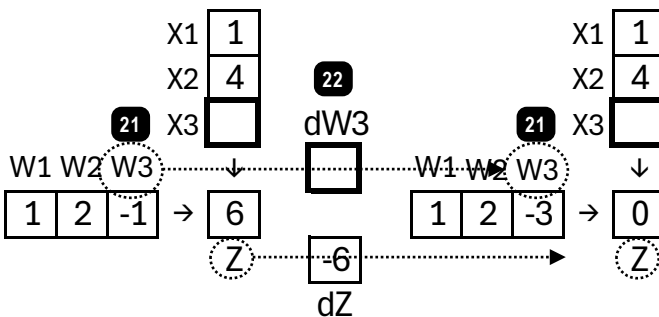
Chapter 12 Gradient - Deep Learning Math Workbook



$$\frac{dZ}{dW_1} = \frac{8}{17} = X_1$$



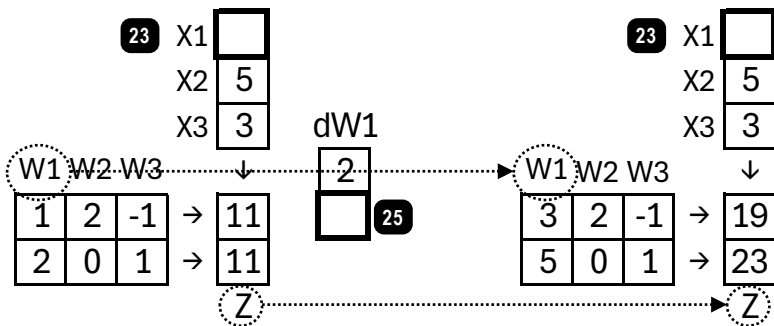
$$\frac{dZ}{dW_2} = \frac{4}{19} = X_2$$



$$\frac{dZ}{dW_3} = \frac{-6}{-2} = X_3$$

17	18	19	20	21	22
4	2	2	2	3	-2

Chapter 12 Gradient - Deep Learning Math Workbook

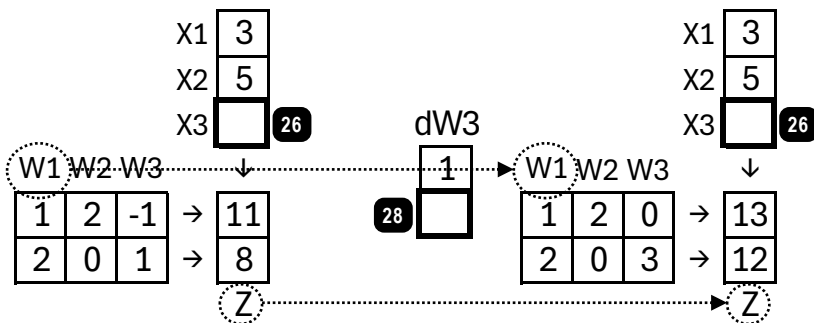


Calculation of $dZ/dW1 = X1$:

$$\frac{dZ}{dW1} = \begin{bmatrix} 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 23 \\ 23 \end{bmatrix}$$

Calculation of $dZ/dW1 = X1$ (repeated):

$$\frac{dZ}{dW1} = \begin{bmatrix} 8 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 23 \\ 23 \end{bmatrix}$$



Calculation of $dZ/dW3 = X3$:

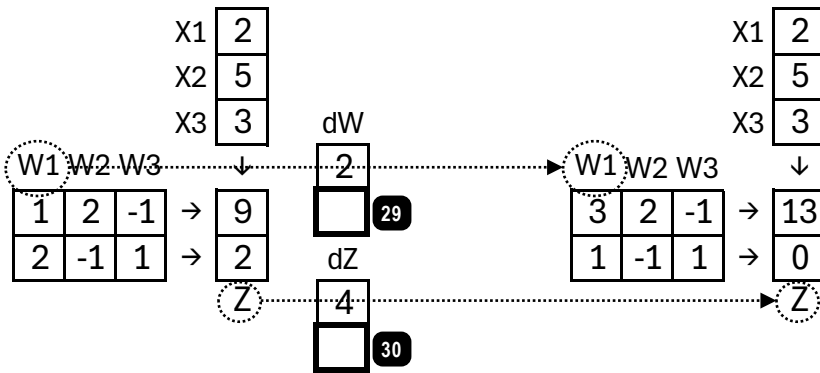
$$\frac{dZ}{dW3} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 26 \\ 26 \end{bmatrix}$$

Calculation of $dZ/dW3 = X3$ (repeated):

$$\frac{dZ}{dW3} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 26 \\ 26 \end{bmatrix}$$

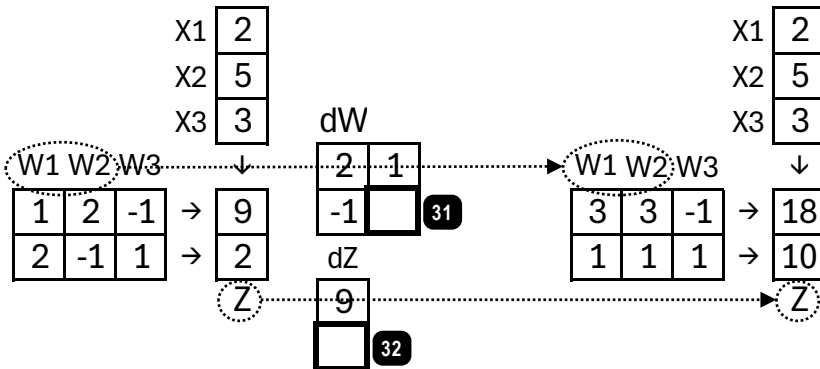
23	24	25	26	27	28
4	12	3	2	4	2

Chapter 12 Gradient - Deep Learning Math Workbook



$$dZ = \begin{bmatrix} 4 \\ \square \\ \square \end{bmatrix} = dW = \begin{bmatrix} 2 \\ \square \\ \square \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ \square \end{bmatrix} \rightarrow \Sigma = \begin{bmatrix} 4 \\ \square \\ \square \end{bmatrix}$$

30 29 30

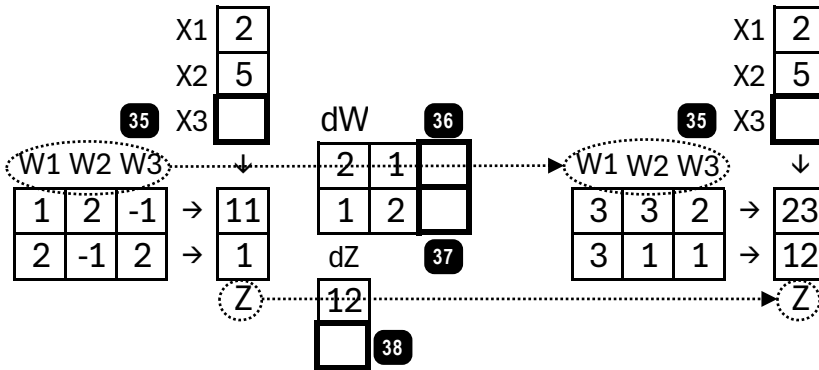


$$dZ = \begin{bmatrix} 9 \\ \square \\ \square \end{bmatrix} = dW = \begin{bmatrix} 2 & 1 \\ -1 & \square \\ \square & \square \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 4 & \square \\ -2 & \square \\ \square & \square \end{bmatrix} \rightarrow \Sigma = \begin{bmatrix} 9 \\ \square \\ \square \end{bmatrix}$$

32 31 33 34 32

29	30	31	32	33	34
1	2	2	8	5	10

Chapter 12 Gradient - Deep Learning Math Workbook



$$dZ = \begin{bmatrix} 12 \\ 11 \end{bmatrix} = \begin{bmatrix} 2 & 1 & \square \\ 1 & 2 & \square \end{bmatrix} \cdot \begin{bmatrix} 2 & 5 & \square \\ 2 & 5 & \square \end{bmatrix} = \begin{bmatrix} 4 & 5 & \square \\ 2 & 10 & \square \end{bmatrix} \rightarrow \begin{bmatrix} 12 \\ 11 \end{bmatrix}$$

(Labels: 36, 35, 39, 40)

$$dZ/dW = \begin{bmatrix} 2 \\ 5 \\ \square \end{bmatrix} \quad \begin{bmatrix} 2 \\ 5 \\ \square \end{bmatrix} \quad \begin{bmatrix} 2 \\ 5 \\ \square \end{bmatrix}$$

(Label: 35)

$$dZ = \begin{bmatrix} 12 \\ \square \end{bmatrix} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 12 \\ \square \end{bmatrix}$$

(Labels: 38, 38)

35 36 37 38 39 40

1 3 1 1 3 1

Chapter 12 Gradient - Deep Learning Math Workbook

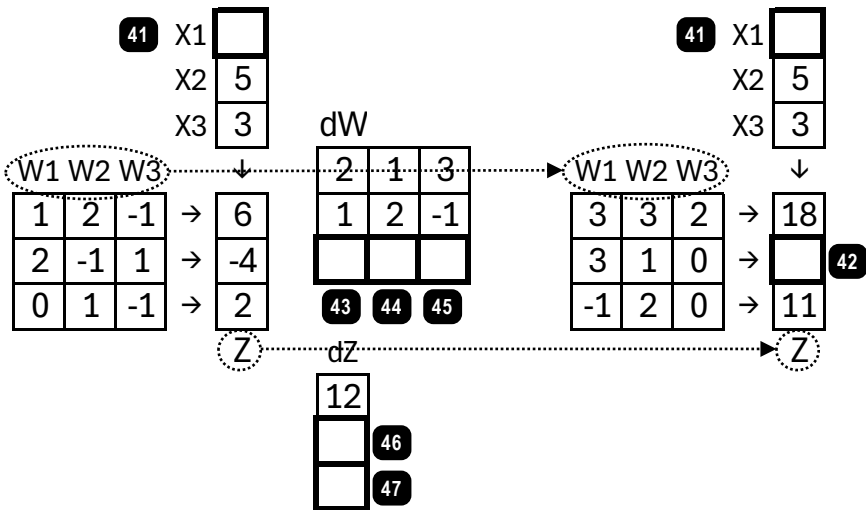


Diagram illustrating the backpropagation step for the second layer. The diagram shows the calculation of dZ for the second layer using the output of the first layer (dZ/dW) and the error from the second layer (dZ). The diagram shows a 3x3 grid of weights (dZ/dW) multiplied by a 3x3 grid of errors (dZ) to produce a 3x3 grid of errors (dZ). The final result is a 3x3 grid of errors (dZ) which is then summed to produce the total error (Σ).

Diagram illustrating the calculation of the derivative dZ/dW for a specific weight matrix W and input vector X .

The input vector X is shown as a column vector:

$$X = \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

The weight matrix W is shown as a 3x3 matrix:

$$W = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & -1 \\ \text{43} & \text{44} & \text{45} \end{bmatrix}$$

The derivative dZ/dW is calculated as the product of the derivative of the output with respect to the input, dZ/dX , and the input vector X :

$$dZ/dW = \begin{bmatrix} 12 \\ \text{46} \\ \text{47} \end{bmatrix} \times \begin{bmatrix} X1 \\ X2 \\ X3 \end{bmatrix} = \begin{bmatrix} 12 \\ \text{46} \\ \text{47} \end{bmatrix} \times \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}$$

The resulting derivative matrix dZ/dW is shown as a 3x3 matrix:

$$dZ/dW = \begin{bmatrix} 12 & 1 & 3 \\ \text{46} & 5 & -1 \\ \text{47} & \text{43} & \text{44} \end{bmatrix}$$

41	42	43	44	45	46	47	48	49	50
-1	2	-1	1	1	6	9	1	5	3