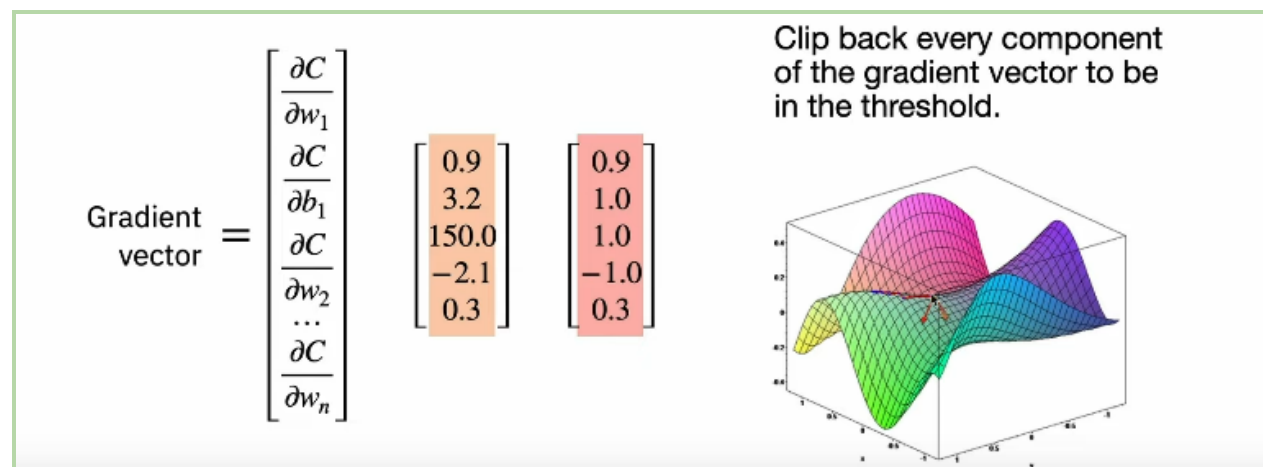


Gradient Clipping :

Gradient clipping is a technique used in deep learning to mitigate the problem of exploding gradients during training. When gradients become very large, they can lead to unstable training, making it difficult for the neural network to converge to a good solution. Gradient clipping helps prevent this issue by limiting the magnitude of gradients during backpropagation.

Gradient clipping is a technique in deep learning that limits the size of gradients during training to prevent them from becoming too large and causing instability. It helps ensure smoother and more stable convergence of neural networks.

One of the solution of unstable gradient problem is gradient clipping and it is not done for all type of networks , it is only done for exploding gradient problem in RNNs



$$\text{Gradient vector} = \begin{bmatrix} \frac{\partial C}{\partial w_1} \\ \frac{\partial C}{\partial b_1} \\ \frac{\partial C}{\partial w_2} \\ \dots \\ \frac{\partial C}{\partial w_n} \end{bmatrix} \quad \begin{bmatrix} 0.9 \\ 3.2 \\ 150.0 \\ -2.1 \\ 0.3 \end{bmatrix} \quad \begin{bmatrix} 0.006 \\ 0.021 \\ 1.0 \\ -0.014 \\ 0.002 \end{bmatrix}$$

Clip back every component of the gradient vector to be in the threshold.

"clip by norm" to keep the same direction of the gradient.

We can experiment with both approaches and with different threshold values