1D Convolution - Backprop

Wednesday, February 7, 2024 12:57 AM

→ input x_1 x_2 x_3 x_4

kernel w_1 w_2 $\quad y = X * W$

o/p y_1 y_2 y_3

$$y_1 = w_1 \cdot x_1 + w_2 \cdot x_2$$
$$y_2 = w_1 \cdot x_2 + w_2 \cdot x_3$$
$$y_3 = w_1 \cdot x_3 + w_2 \cdot x_4$$

Start with $\frac{\partial L}{\partial y_i}$. Let's compute $\frac{\partial L}{\partial w_i}$ & $\frac{\partial L}{\partial x_i}$

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial w_1} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial w_1} + \frac{\partial L}{\partial y_3} \cdot \frac{\partial y_3}{\partial w_1}$$

$$= \cdot x_1 \quad \cdot x_2 \quad \cdot x_3$$

$$\frac{\partial L}{\partial w_2} = \cdot x_2 \quad \cdot x_3 \quad \cdot x_4$$

$$\Rightarrow \quad \frac{\partial L}{\partial w} = X * \frac{\partial L}{\partial y}$$

$$y_1 = w_1 \cdot x_1 + w_2 \cdot x_2$$
$$y_2 = w_1 \cdot x_2 + w_2 \cdot x_3$$
$$y_3 = w_1 \cdot x_3 + w_2 \cdot x_4$$

$$\frac{\partial L}{\partial x_4} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_4} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_4} + \frac{\partial L}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_4}$$

$$= w_1 \quad 0 \quad 0$$

$$\frac{\partial L}{\partial x_2} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_2} + \frac{\partial L}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_2}$$

$$= \quad w_2 \qquad w_1 \qquad 0$$

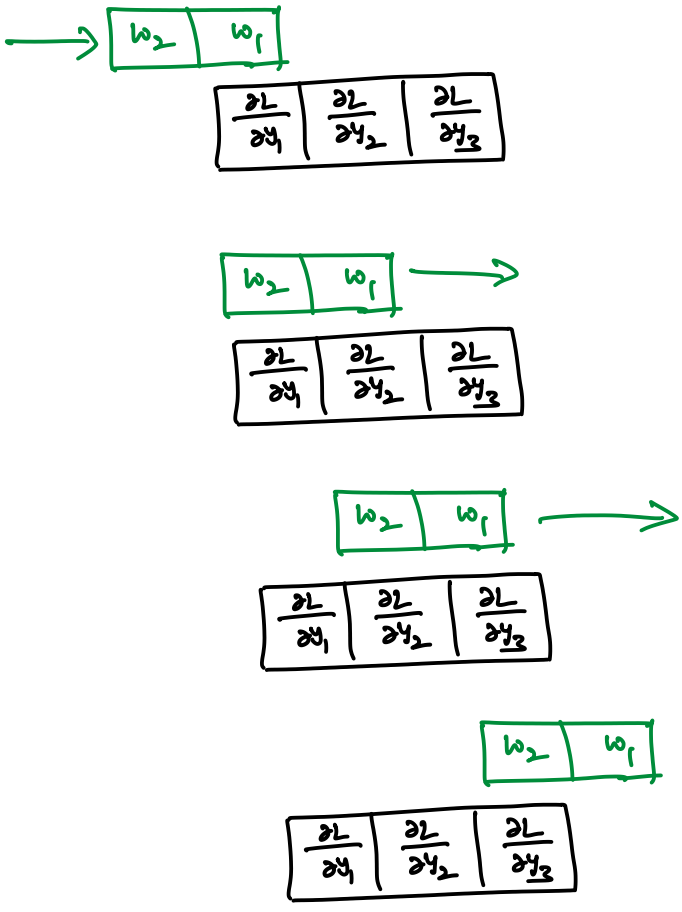
$$\frac{\partial L}{\partial x_3} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_3} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_3} + \frac{\partial L}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_3}$$

$$= \quad 0 \qquad w_2 \qquad w_1$$

$$\frac{\partial L}{\partial x_4} = \frac{\partial L}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_4} + \frac{\partial L}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_4} + \frac{\partial L}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_4}$$

$$= \quad 0 \qquad 0 \qquad w_2$$

These operations can be visualized as



→ Notice that the flipped kernel 'full' convolves the up-stream gradients to result in the down-stream gradients w.r.t. the input.