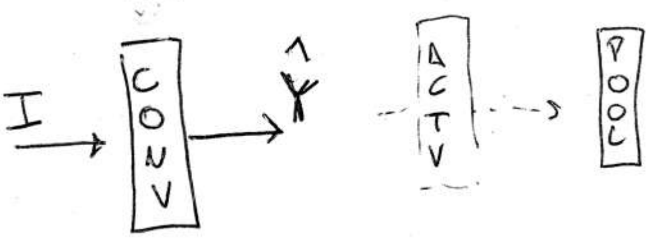
# Back propagation in CNN layer






- I 1 solo canal ( escala de grises) $\qquad I(H, W)$
- W 1 solo kernel (1 solo filtro) $\qquad w(k_1, k_2)$
- padding = valid $\qquad \hat{Y}(H-k_1+1, W-k_2+1)$
- stride = 1

convolución $\qquad (w * I)(i,j) = \sum_{m=0}^{k_1-1}\sum_{n=0}^{k_2-1} I(i-m, j-n)\, w(m,n)$

correlación cruzada $\qquad (w \otimes I)(i,j) = \sum_{m=0}^{k_1-1}\sum_{n=0}^{k_2-1} I(i+m, j+m)\, w(m,n)$

$$(w * I)(i,j) = (w_{flipped} \otimes I)(i,j)$$

La capa de conv hace la operacion

$$\hat{Y} = (w \otimes I)_{(i,j)} + b = \sum_{m=0}^{k_1-1}\sum_{n=0}^{k_2-1} I(i+m, j+n)\, w(m,n) + b$$

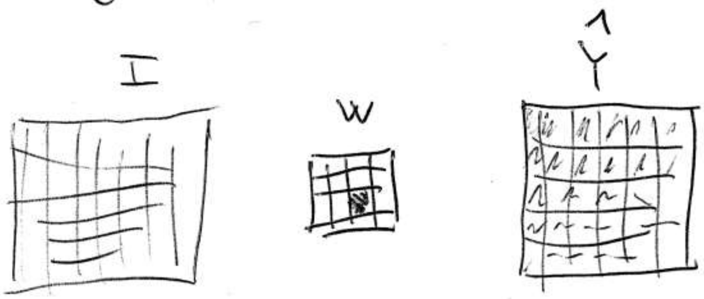
La Loss function de ese punto, dependera del valor real $Y$ y su aproximación

$L(Y, \hat{Y})$ queremos hallar

$\frac{\partial L}{\partial I} \rightarrow$ xq pasar gradiete a capa previa

$\frac{\partial L}{\partial w} \rightarrow$ xa actualizar los pesos de $w$

$1^{o}$ $\frac{\partial L}{\partial w}$ ¿Como afecta un peso $w(m',n')$ a la salida? 




Afecta a todas las salidas ya que cada $Y(i,j)$ es el resultado de aplicar $w$ en distintas posiciones de $I$

entonces tendremos que

$$\frac{\partial L}{\partial w(m',n')} = \sum_{i=0}^{H-k_1} \sum_{j=0}^{W-k_2} \underbrace{\frac{\partial L}{\partial Y(i,j)}}_{\text{lo daría la capa previa... sería nuest } \frac{\partial L}{\partial I}} \underbrace{\frac{\partial \hat{Y}(i,j)}{\partial w(m',n')}}_{\text{hay q averiguarlo}} \quad (1)$$

$$\frac{\partial \hat{Y}(i,j)}{\partial w(m',n')} = \frac{\partial}{\partial w(m',n')}\left[\underbrace{\sum_{m=0}^{k_1-1} \sum_{n=0}^{k_2-1} I(i+m, j+n)\, w(m,n)}_{\text{hay 1 solo termino de } m=m' \text{ y } n=n'} + b\right] =$$

$$= \frac{\partial}{\partial w(m',n')}\left[ I(i+0,j+0)\, w(0,0) + I(i+0, j+1)\, w(0,1) + I(i+0,j+2)\, w(j+2) \\ \dots + I(i+m', j+n') + \dots + b \right]$$

se derivamos se llega a

$$\frac{\partial \hat{Y}(i,j)}{\partial w(m',n')} = I(i+m', j+n')$$ si sustituimos en (1)

$$\frac{\partial L}{\partial w(m',n')} = \sum_{i=0}^{H-k_1} \sum_{j=0}^{W-k_2} \frac{\partial L}{\partial \hat{Y}(i,j)} \cdot I(i+m', j+n')$$

si reordenamos

$$\frac{\partial L}{\partial w(m',n')} = \sum_{i=0}^{H-k_1} \sum_{j=0}^{W-k_2} I(m'+i,\, n'+j)\, \frac{\partial L}{\partial \hat{Y}(i,j)}$$

$$= \frac{\partial L}{\partial Y(m',n')} \otimes I(m',n') = \left\{ \frac{\partial L}{\partial Y(m',n')} \right\}_{\text{flipped}} * I(m',n')$$

ó

$$\boxed{\frac{\partial L}{\partial w(m',n')} = \left\{ \frac{\partial L}{\partial \hat{Y}(m',n')} \right\}_{180^\circ} * I(m',n')}$$

¿Que está pasando?

$\frac{\partial L}{\partial \hat{Y}}$

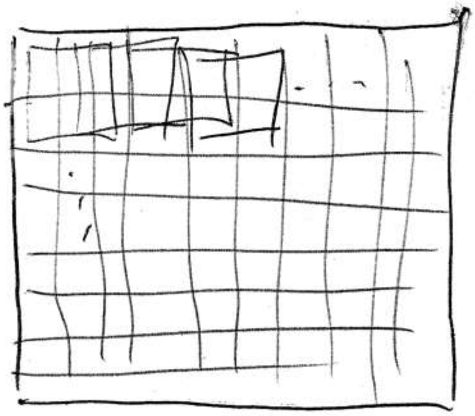
| | | |
|---|---|---|
| 1 | 2 | 3 |
| 4 | 5 | 6 |
| 7 | 8 | 9 |

$\xrightarrow{\text{flip}}$

| | | |
|---|---|---|
| 9 | 8 | 7 |
| 6 | 5 | 4 |
| 3 | 2 | 1 |

$\xrightarrow{\text{Conv}}$

I

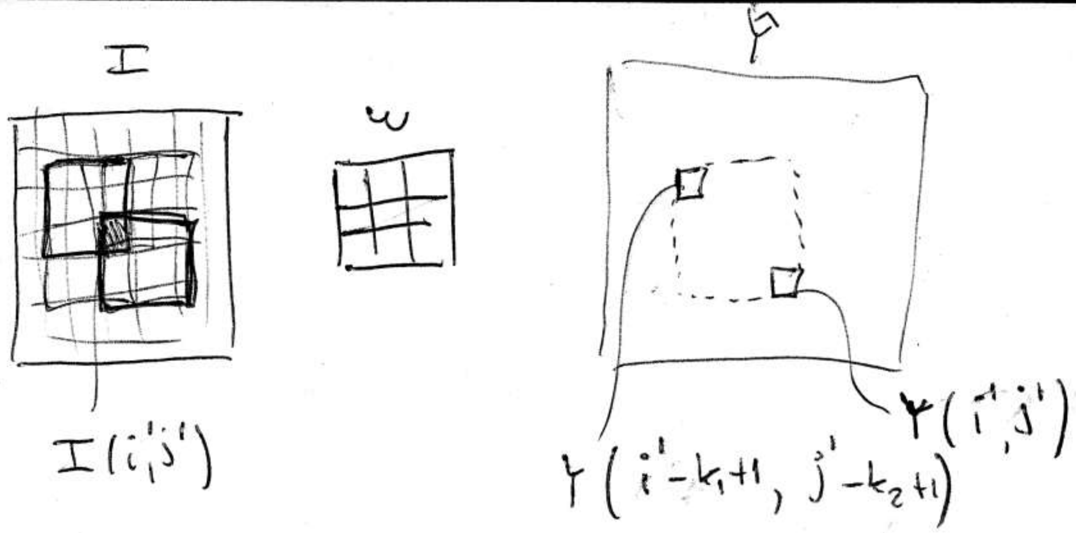


---

Ahora calculamos

$$\frac{\partial L}{\partial I} = \underbrace{\frac{\partial L}{\partial \hat{Y}}}_{\text{conocido}} \; \underbrace{\frac{\partial \hat{Y}}{\partial I}}_{\text{a conocer}}$$

¿como varia la salida cuando se varia 1 pixel de la entrada? → Va a depender del tamaño del kernel




un pixel $I(i',j')$ va a modificar un area de $Y$ con vertise sup izq $(i'-k_1+1, j'-k_2+1)$ y con vertice inf der $(i', j')$

Entonces:

$$\frac{\partial L}{\partial I(i',j')} \overset{②}{=} \underbrace{\sum_{m=0}^{k_1-1} \sum_{n=0}^{k_2-1}}_{\text{Regla de la cadena para cada pixel de } Y \text{ que se ve modificado por el pixel } I(i',j')} \overbrace{\frac{\partial L}{\partial Y(i'-m, j'-n)}}^{\text{de la capa previa}} \cdot \overbrace{\frac{\partial Y(i'-m, j'-n)}{\partial I(i',j')}}^{\text{a trabajar}}$$

$$\frac{\partial Y(i'-m, j'-n)}{\partial I(i',j')} = \frac{\partial}{\partial I(i',j')} \left[ \underbrace{\sum_{p=0}^{k_1-1} \sum_{q=0}^{k_2-1} I(i'-m+p, j'-n+q) \cdot w(p,q) + b}_{\substack{\text{hay 1 solo termino donde} \\ p=m \text{ y } q=n}} \right]$$

¿qué hace q?

$$I(i'-m+m, j-n+n) = I(i',j)$$

Es lo que estoy derivando!!

Entonees

$$\frac{\partial Y(i'-m, j'-n)}{\partial I(i', j')} = \frac{\partial}{\partial I(i', j')}\left[\dots + I(i', j')\, w(m,n) + \dots + b\right]$$ IV

$$= w(m,n)$$ → lo aplico en 2

$$\frac{\partial L}{\partial I(i', j')} = \sum_{m=0}^{k_1-1} \sum_{n=0}^{k_2-1} \frac{\partial L}{\partial Y(i'-m, j'-n)} \cdot w(m,n)$$

$$\boxed{\frac{\partial L}{\partial I(i', j')} = w_{(i', j')} * \frac{\partial L}{\partial Y(i', j')}}$$