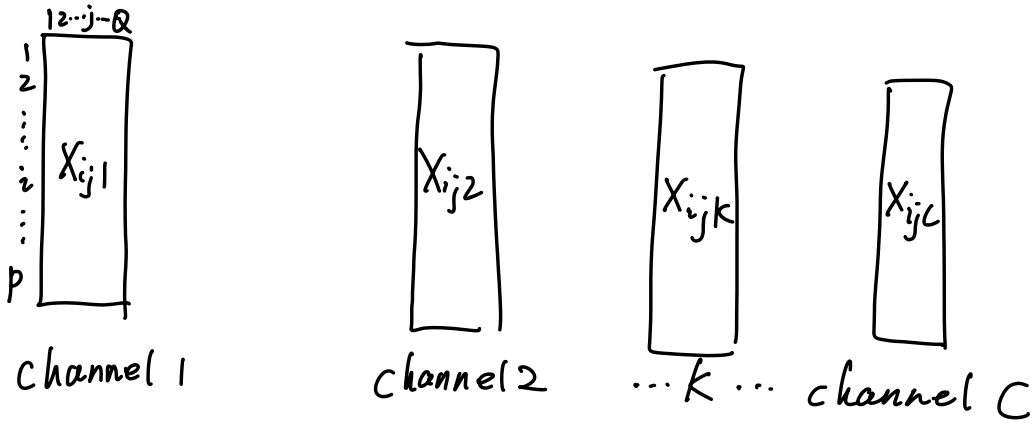
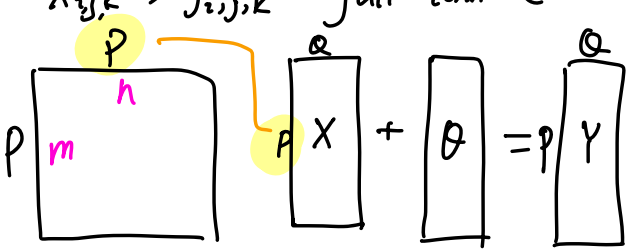
23-S2-Q4

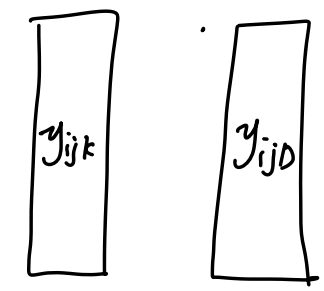
(a) Q: $x_{ijk} \to y_{ijk}$? parameters?

Solution ① understand



$x_{ijk} \to y_{i,j,k}$ full connect





$$\sum \begin{bmatrix} w_{11} & w_{12} \cdots & w_{1n} \cdots & w_{1p} \\ w_{21} & w_{22} \cdots & w_{2n} \cdots & w_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ w_{m1} & w_{m2} \cdots & w_{mn} \cdots & w_{mp} \\ \vdots & \vdots & \vdots & \vdots \\ w_{p1} & w_{p2} & w_{pn} \cdots & w_{pp} \end{bmatrix} \begin{bmatrix} x_{11} \cdots x_{1Q} \\ x_{21} \cdots x_{2Q} \\ \vdots \\ x_{i1} \cdots x_{iQ} \\ \vdots \\ x_{p1} \quad x_{pQ} \end{bmatrix} + \begin{bmatrix} b_{11} \cdots b_{1Q} \\ b_{21} \cdots b_{2Q} \\ \vdots \\ b_{i1} \cdots b_{iQ} \\ \vdots \\ b_{p1} \cdots b_{pQ} \end{bmatrix}$$

$$y_{ijk} = \sum_{l=1}^{C} \sum_{m=1}^{P} \sum_{n=1}^{P} W_{mnlk} X_{ijl} + b_{ijk} (1 \le k \le D)$$

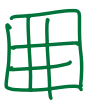
parameters $P \times P \times C \times D + P \times Q \times D$

一张W　　1张bias

C张W生成1张Y　　D张bias

生成D张Y

(b) CNN ?



$$y_{i,j,k} = \sum_{u=-1}^{1} \sum_{v=-1}^{1} \sum_{l=1}^{C} w_{u,v,l,k}\, x_{i-u,j-v,l} + b_k$$

$$3 \times 3 \times C \times D + D = 9 \times C \times D + D$$

(c) $$y_{i,j,k} = \sum_{l=1}^{C} w_{l,k}\, x_{i,j,l} + b_k$$

$$C \times D + D$$

RGB　D个特征图　bias

(d) $$y_{i,k} = \sum_{l=1}^{C} w_{l,k}\, x_{i,l}$$

$$C \times D$$

(e) $Y = XW$

① Both architectures apply matrix multiplications to process input

② $Y = XW$ shows the CNN perform a linear transformation on the input features at each spatial position, identical across all position.

It is similar to the linear layers used in Transformer, where input are transformed via weight matrices