Linear algebra for AI & ML

Convolution:

$a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$

$c = a * b$: convolution.

$c \in \mathbb{R}^{(n+m-1)}$

$$c_k = \sum_{i+j=k+1} a_i b_j \qquad k = 1, 2, \ldots, n+m-1$$

$n = 4$, $m = 3$ $\qquad c \in \mathbb{R}^{4+3-1} \Rightarrow c \in \mathbb{R}^6$

$c_1 = a_1 b_1$; $c_2 = a_2 b_1 + a_1 b_2$

$c_3 = a_3 b_1 + a_2 b_2 + a_1 b_3$

$c_4 = a_2 b_3 + a_3 b_2 + a_4 b_1$

Ex:

$$p(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}$$

$$q(x) = b_1 + b_2 x + \cdots + b_m x^{m-1}$$

$$p(x)\,q(x) = a_1 b_1 + (a_1 b_2 + a_2 b_1)\, x$$

$$\underbrace{\begin{bmatrix} a_1 & 0 & 0 \\ a_2 & a_1 & 0 \\ a_3 & a_2 & a_1 \\ a_4 & a_3 & a_2 \\ 0 & a_4 & a_3 \\ 0 & 0 & a_4 \end{bmatrix}_{6\times 3}}_{\text{Toeplitz}} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \left[\begin{array}{l} \text{convolution} \\ c \\ \\ \\ \end{array}\right.$$

$$\begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = 0$$

$$\begin{bmatrix} 1 & \alpha_1 & \alpha_1^2 & \alpha_1^3 & \alpha_1^4 \end{bmatrix} \begin{bmatrix} a_1 & 0 \\ a_2 & a_1 \\ a_3 & a_2 \\ a_4 & a_3 \\ 0 & a_4 \end{bmatrix} = 0$$

$$\begin{bmatrix} a_1 & 0 & 0 & b_1 & 0 & 0 & 0 \\ a_2 & a_1 & 0 & b_2 & b_1 & 0 & 0 \\ a_3 & a_2 & a_1 & b_3 & b_2 & b_1 & 0 \\ a_4 & a_3 & a_2 & 0 & b_3 & b_2 & b_1 \\ 0 & a_4 & a_3 & 0 & 0 & b_3 & b_2 \\ 0 & 0 & a_4 & 0 & 0 & 0 & b_3 \end{bmatrix}$$

$$\mathbb{R}^n \xrightarrow{T_1} \mathbb{R}^m \xrightarrow{T_2} \mathbb{R}^p$$

$$T_2 \in \mathbb{R}^{p \times m}$$

$$T_1 \in \mathbb{R}^{m \times n}$$

$$T_2 T_1 \in \mathbb{R}^{p \times n}$$

$$T_2 T_1 = C = \begin{bmatrix} & & \\ & \cdot & \\ & & \end{bmatrix}$$

$(i,j)^{th}$

f^n composition not commutative

$\Rightarrow$ matrix multiplication not commutative.

(In general $AB \neq BA$)

f^n composition associative

$\Rightarrow$ matrix multiplication is associative

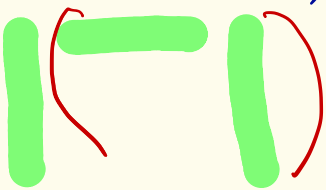
(In general

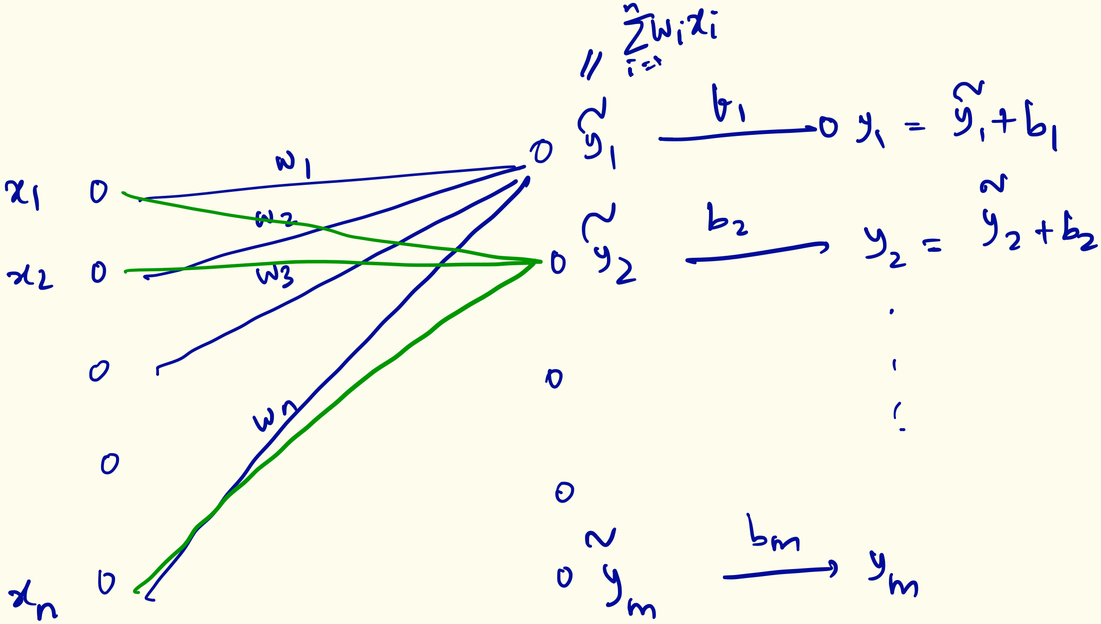
$(AB)C = A(BC) = ABC$

pqr to obtain (AB) which is $p \times r$

prs to obtain $(AB)C$

$pqr + prs$





Layer l

input $x = \begin{pmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$; $y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \in \mathbb{R}^m$: output

$y = Wx + b$ where $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$: bias vector

$W \in \mathbb{R}^{m \times n}$: weight matrix.

Let σ be an activation fn.

outputing the layer: $\sigma(Wx+b) \in \mathbb{R}^m$

<u>Next layer:</u>

$$\sigma\left(W_2\left(\sigma(Wx+b)\right)+b_2\right)$$