Matrix-Matrix Multiplication - Pullback / vJp rule

$$f(\underline{\underline{A}}, \underline{\underline{B}}) = \underline{\underline{A}}\,\underline{\underline{B}} =: \underline{\underline{C}}$$

$$\underline{\underline{A}} \in \mathbb{R}^{m \times n}, \quad \underline{\underline{B}} \in \mathbb{R}^{n \times o} \quad \Rightarrow \quad \underline{\underline{C}} \in \mathbb{R}^{m \times o}$$

task: backpropagate cotangent information on the output $\bar{\underline{\underline{C}}} \in \mathbb{R}^{m \times o}$ to the inputs $\bar{\underline{\underline{A}}} \in \mathbb{R}^{m \times n}$ & $\bar{\underline{\underline{B}}} \in \mathbb{R}^{n \times o}$

$\rightarrow$ backpropagation of Neural Networks for batched

"informally" $\quad \bar{\underline{\underline{A}}} = \bar{\underline{\underline{C}}} : \dfrac{\partial f}{\partial \underline{\underline{A}}} \quad$ & $\quad \bar{\underline{\underline{B}}} = \bar{\underline{\underline{C}}} : \dfrac{\partial f}{\partial \underline{\underline{B}}}$

index notation

forward: $C_{ik} = A_{ij} B_{jk}$ Einstein summation convention

pullback:
$$\bar{A}_{mn} = \bar{C}_{ik} \frac{\partial C_{ik}}{\partial A_{mn}} \quad (1)$$
$$\bar{B}_{mn} = \bar{C}_{ik} \frac{\partial C_{ik}}{\partial B_{mn}} \quad (2)$$

(1)
$$\frac{\partial C_{ik}}{\partial A_{mn}} = \frac{\partial A_{ij}}{\partial A_{mn}} B_{jk} = \delta_{im} \delta_{jn} B_{jk}$$

(2)
$$\frac{\partial C_{ik}}{\partial B_{mn}} = A_{ij} \frac{\partial B_{jk}}{\partial B_{mn}} = A_{ij} \delta_{jm} \delta_{kn}$$

$$\rightarrow \quad \bar{A}_{mn} = \bar{C}_{ik} \delta_{im} \delta_{jn} B_{jk} = \bar{C}_{mk} B_{nk}$$

$$\rightarrow \quad \bar{B}_{mn} = \bar{C}_{ik} A_{ij} \delta_{jm} \delta_{kn} = \bar{C}_{in} A_{im} = A_{im} \bar{C}_{in}$$

back to symbolic notation

$$\bar{\underline{\underline{A}}} = \bar{\underline{\underline{C}}}\, \underline{\underline{B}}^T \qquad \bar{\underline{\underline{B}}} = \underline{\underline{A}}^T \bar{\underline{\underline{C}}}$$

full pullback rule

$$\boxed{\mathcal{B}(f, (\underline{A}, \underline{B}), (\bar{\underline{C}},)) = \left((\underbrace{\underline{A}\,\underline{B}}_{\underline{C}},), (\underbrace{\bar{\underline{C}}\, \underline{B}^T}_{\bar{\underline{A}}}, \underbrace{\underline{A}^T \bar{\underline{C}}}_{\bar{\underline{B}}}) \right)}$$