Matrix-Matrix Multiplication – Pushforward/Jvp rule

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

$$\underline{\underline{A}} \in \mathbb{R}^{m \times n} \qquad \underline{\underline{B}} \in \mathbb{R}^{n \times o} \qquad \longrightarrow \qquad \underline{\underline{C}} \in \mathbb{R}^{m \times o}$$

task: forward-propagate tangent information $\dot{\underline{\underline{A}}} \in \mathbb{R}^{m \times n}$ and $\dot{\underline{\underline{B}}} \in \mathbb{R}^{n \times o}$ to $\dot{\underline{\underline{C}}} \in \mathbb{R}^{m \times o}$

loosely: " $\dot{\underline{\underline{C}}} = \frac{\partial f}{\partial \underline{\underline{A}}} : \dot{\underline{\underline{A}}} + \frac{\partial f}{\partial \underline{\underline{B}}} : \dot{\underline{\underline{B}}}$ "

index notation:

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

pushforward: $$\dot{C}_{ik} = \underbrace{\frac{\partial C_{ik}}{\partial A_{mn}}}_{①} \dot{A}_{mn} + \underbrace{\frac{\partial C_{ik}}{\partial B_{mn}}}_{②} \dot{B}_{mn}$$

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

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

$$\dot{C}_{ik} = \delta_{im} \delta_{jn} B_{jk} \dot{A}_{mn} + A_{ij} \delta_{jm} \delta_{kn} \dot{B}_{mn}$$

$$= B_{jk} \dot{A}_{ij} + A_{ij} \dot{B}_{jk}$$

$$= \dot{A}_{ij} B_{jk} + A_{ij} \dot{B}_{jk}$$

back to symbolic

$$\dot{\underline{\underline{C}}} = \dot{\underline{\underline{A}}}\,\underline{\underline{B}} + \underline{\underline{A}}\,\dot{\underline{\underline{B}}}$$

full pushforward rule

$$\boxed{\mathcal{F}(f, (\underline{\underline{A}}, \underline{\underline{B}}), (\dot{\underline{\underline{A}}}, \dot{\underline{\underline{B}}})) = ((\underbrace{\underline{\underline{A}}\,\underline{\underline{B}}}_{\underline{\underline{C}}},), (\underbrace{\dot{\underline{\underline{A}}}\,\underline{\underline{B}} + \underline{\underline{A}}\,\dot{\underline{\underline{B}}}}_{\dot{\underline{\underline{C}}}},))}$$