Dense Layer

Tuesday, July 18, 2017 6:31 PM

Deriving backpropagation through a 'dense' operation

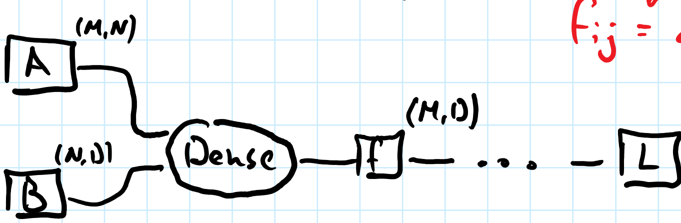
Suppose A is a tensor of shape (M, N)

Suppose B is a tensor of shape (N, D)

Consider the computational graph

"dense" ⟷ matrix multiplication

$$f_{ij} = \sum_k A_{ik} B_{kj}$$



The following is valid if and only if L is a scalar:

Assume that $\frac{dL}{df_{ij}}$ is known for all $1 \le i \le M$, $1 \le j \le D$

Then by the chain rule for multiple variables

$A_{\ell m}$ is the (ℓ, m) entry in A

$$\frac{\partial L}{\partial A_{\ell m}} = \sum_{i,j} \frac{\partial L}{\partial f_{ij}} \frac{\partial f_{ij}}{\partial A_{\ell m}}$$

$$f_{ij} = \sum_k A_{ik} B_{kj}$$

$$\frac{\partial f_{ij}}{\partial A_{\ell m}} = \sum_k \frac{\partial A_{ik}}{\partial A_{\ell m}} B_{kj}$$

$$= \sum_k \delta_{i\ell} \delta_{km} B_{kj}$$

$$= \delta_{i\ell} B_{mj}$$

$$= \sum_{ij} \frac{\partial L}{\partial f_{ij}} \delta_{i\ell} B_{mj}$$

$$B_{jm} = B^T_{mj}$$ transpose

$$= \sum_j \frac{\partial L}{\partial f_{\ell j}} B_{mj} = \sum_j \frac{\partial L}{\partial f_{\ell j}} B^T_{jm}$$

$$\boxed{\frac{\partial L}{\partial A} = \frac{\partial L}{\partial f} B^T}$$

$$\frac{\partial L}{\partial f} = \begin{bmatrix} \frac{\partial L}{\partial f_{00}} & \frac{\partial L}{\partial f_{01}} & \cdots & \frac{\partial L}{\partial f_{0D}} \\ \frac{\partial L}{\partial \ldots} & & & \end{bmatrix}$$

$$\boxed{\frac{\partial L}{\partial A} = \frac{\partial L}{\partial f} B}$$

$$\frac{\partial L}{\partial f} = \begin{bmatrix} \frac{\partial L}{\partial f_{00}} & \frac{\partial L}{\partial f_{01}} & \cdots & \frac{\partial L}{\partial f_{0D}} \\ \frac{\partial L}{\partial f_{10}} & \ddots & & \\ \vdots & & & \\ \frac{\partial L}{\partial f_{M0}} & \cdots & & \frac{\partial L}{\partial f_{MD}} \end{bmatrix}$$

$B_{\ell m}$ is the (ℓ, m) entry of B

$$\frac{\partial L}{\partial B_{\ell m}} = \sum_{ij} \frac{\partial L}{\partial f_{ij}} \frac{\partial f_{ij}}{\partial B_{\ell m}}$$

$$\frac{\partial f_{ij}}{\partial B_{\ell m}} = \sum_k A_{ik} \frac{\partial B_{kj}}{\partial B_{\ell m}}$$

$$= \sum_k A_{ik} \delta_{k\ell} \delta_{jm}$$

$$= A_{i\ell} \delta_{jm}$$

$$= \sum_{ij} \frac{\partial L}{\partial f_{ij}} A_{i\ell} \delta_{jm}$$

$$= \sum_i \frac{\partial L}{\partial f_{im}} A_{i\ell} = \sum_i A^T_{\ell i} \frac{\partial L}{\partial f_{im}}$$

$$\boxed{\frac{\partial L}{\partial B} = A^T \frac{\partial L}{\partial f}}$$

$$\frac{\partial L}{\partial f} = \begin{bmatrix} \frac{\partial L}{\partial f_{00}} & \frac{\partial L}{\partial f_{01}} & \cdots & \frac{\partial L}{\partial f_{0D}} \\ \frac{\partial L}{\partial f_{10}} & \ddots & & \\ \vdots & & & \\ \frac{\partial L}{\partial f_{M0}} & \cdots & & \frac{\partial L}{\partial f_{MD}} \end{bmatrix}$$