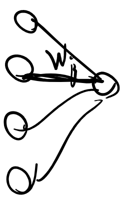
Backprop notes

Monday, November 27, 2023 14:54

Continuous Perceptron

$$y(x^\mu) = f\left(\sum_{j=1}^{N} w_j x_j^\mu\right)$$

$$L = \sum_{\mu=1}^{p} \left(y^\mu - y(x^\mu)\right)$$



$$w_i \to w_i + dw_i$$

$$dy = \frac{\partial y}{\partial w_i} dw_i = f' \cdot x_i^\mu dw_i$$

$$\delta^\mu$$

$$dL = \frac{\partial L}{\partial y} dy = \sum_{1}^{p} \underbrace{(y^\mu - y)}_{}\cdot(-2)\, f' x_i^\mu dw_i$$

$$\Delta w_i = -\alpha \frac{dL}{dw_i} = 2\alpha \sum_{\mu=1}^{p} \underbrace{\delta^\mu f'}_{\text{Redefined } \delta^\mu} x_i^\mu$$

Multi-layer case

INDEX μ IS IMPLICIT

$$z_i^L = f\left(\sum_{j=1}^{N} w_{ij}^L z_j^{L-1}\right)$$

Perturb synapse w_{ij}^L

$$w_{ij}^L \to w_{ij}^L + dw_{ij}^L$$

$$d\ \ \partial z_i^L \quad dw^L \quad f' z^{L-1} dw_{ij}^L$$

$$dz_k^L = \frac{\partial z_i^L}{\partial w_{ij}^L} dw_{ij}^L = f' z_j^{L-1} dw_{ij}^L$$

$$dz_k^{L+1} = \sum_{m=1}^{N} \frac{\partial z_k^{L+1}}{\partial z_m^L} dz_m^L$$

$$\frac{\partial z_k^{L+1}}{\partial z_m^L} = w_{km}^{L+1} f_k' \stackrel{\text{DEF}}{=} W^{L+1} \quad (N \times N \text{ MATRIX})$$

$$d\underline{z}^{L+1} = W^{L+1} \, d\underline{z}^{L}$$

$$d\underline{y} = W^{F} \, d\underline{z}^{F-1}$$

$$dL = \sum_{k=1}^{N} \frac{\partial L}{\partial y_k} dy_k = -2 \delta_k^{M} dy_k$$

$$= -2 \delta^T d\underline{y}$$

δ IS A VECTOR $N \times 1$

$$dL = -2 \underset{1\times N}{\delta^T} \underset{N\times N}{W^F} W^{F-1} \ldots W^{L+1} f' z_j^{L-1} dw_{ij}^L$$

$$\Delta w_{ij}^L = \alpha \underbrace{\underline{\delta}^T W^F}_{= \underset{1\times N}{\underline{\delta}_F^T}} W^{F-1} \ldots W^{L+1} f' z_j^{L-1}$$

$$\delta_F \stackrel{\text{DEF}}{=} W$$

$$\underline{\delta}_F^T W^{F-1}$$

$$= \alpha \, \delta_{L,j}^T f' z_j^{L-1}$$