AMAL-TP1

Question 1

$$\frac{\partial L \circ h}{\partial x_i}(x) = \sum_{j=1}^{p} \frac{\partial L}{\partial y_j}(h(x)) \frac{\partial h_j}{\partial x_i}(x) = \sum_{j=1}^{p} (\nabla L)_j \frac{\partial h_j}{\partial x_i}(x)$$

Question 2

$$\frac{\partial L_{mse}}{\partial y}(\hat{y}, y) = \nabla L \frac{\partial mse}{\partial y}(\hat{y}, y) = -2 \nabla L (\hat{y} - y)$$

∇L est un scalaire

$$\frac{\partial L_{mse}}{\partial \hat{y}} = 2 \nabla L (\hat{y} - y)$$

$$\frac{\partial L \circ f}{\partial x_j}(x, w, b) = \nabla L \frac{\partial f}{\partial x_j}(x, w, b) = \nabla L \, w_j$$

Question 3

$$\frac{\partial L_{mse}}{\partial Y_{ij}}(\hat{Y}, Y) = \nabla L \frac{\partial mse}{\partial Y_{ij}}(\hat{Y}, Y) = -\frac{2}{q} \nabla L \sqrt{\|\hat{Y} - Y\|^2}'$$

$$\boxed{mse(\hat{Y}, Y) = \frac{1}{q}\left((\hat{Y}_{11} - Y_{11})^2 + \dots + (\hat{Y}_{1p} - Y_{1p})^2 + \dots + (\hat{Y}_{q1} - Y_{q1})^2 + \dots + (\hat{Y}_{qp} - Y_{qp})^2\right)}$$

$$\frac{\partial L_{mse}}{\partial \hat{Y}_{ij}}(\hat{Y}, Y) = \frac{2}{q} \nabla L \sqrt{\|\hat{Y} - Y\|^2}'$$

$$\frac{\partial L \circ f}{\partial X_{ij}}(X, W, b) = \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \frac{\partial f_{ij}}{\partial X_{ij}}(X, W, b) =$$

$$= \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \, w_{ji} = \sum_{i=1}^{q} \sum_{j=1}^{p} (\nabla L)_j \, w_{ji}$$

$$\frac{\partial L \circ f}{\partial W_{ij}}(X, W, b) = \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \frac{\partial f_{ij}}{\partial W_{ij}}(X, W, b) =$$

$$= \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \, x_{ji} = \sum_{j=1}^{p} \sum_{i=1}^{q} x_{ji} (\nabla L)_i$$

$$\frac{\partial L \circ f}{\partial b_j}(X, W, b) = \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \frac{\partial f_{ij}}{\partial b_j}(X, W, b) =$$

$$= \sum_{i=1}^{q} \sum_{j=1}^{p} \frac{\partial L}{\partial y_{ij}}(f(X, W, b)) \cdot 1$$

Question 4

$$\frac{\partial Lomse}{\partial Y}(\hat{Y}, Y) = -\frac{2}{q}\nabla L\,(\hat{Y} - Y) \quad \rightarrow \text{où } \nabla L \text{ est un scalaire}$$

$$\frac{\partial Lomse}{\partial \hat{Y}}(\hat{Y}, Y) = \frac{2}{q}\nabla L\,(\hat{Y} - Y)$$

$$\frac{\partial L \circ f}{\partial X}(X, W, b) = \nabla L \cdot W^T$$

$$\frac{\partial L \circ f}{\partial W}(X, W, b) = X^T \cdot \nabla L$$

$$\frac{\partial L \circ f}{\partial b}(X, W, b) = \underbrace{\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}}_{\in \mathcal{M}_{q,p}(\mathbb{R})}$$

Question 5

$$\frac{\partial C}{\partial W} = \frac{\partial C}{\partial \hat{Y}} \frac{\partial \hat{Y}}{\partial W} \quad \text{où} \quad \frac{\partial C}{\partial \hat{Y}} = -\frac{2}{q}(\hat{Y} - Y) \cdot \nabla L,$$

$$\frac{\partial \hat{Y}}{\partial W} = X^T \cdot \nabla L$$

$$\frac{\partial C}{\partial b} = \frac{\partial C}{\partial \hat{Y}} \frac{\partial \hat{Y}}{\partial b} \quad \text{où} \quad \frac{\partial \hat{Y}}{\partial b} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix}_{q \times p}$$