$$Loss(p, q) = D_{KL}(p, q)$$

$$p = (\hat{p}_1, \cdots, \hat{p}_m)^T, \quad q = (p_1, \cdots, p_m)^T$$

$$\hat{p}(x_j) = \frac{1}{h} \sum_{i=1}^{n} w_i \cdot \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_j - x_i)^2}{2h^2}\right)$$

$$\frac{\partial L}{\partial W} = \frac{\partial \hat{p}}{\partial W} \frac{\partial L}{\partial \hat{p}}$$

Lossをweightで微分

ksdensityをweightで微分

Lossを$\hat{p}$で微分

$$\frac{\partial \hat{p}}{\partial W} = \overset{n \times m}{\begin{pmatrix} \frac{\partial p_1}{\partial w_1} & \cdots & \frac{\partial p_1}{\partial w_m} \\ \vdots & \ddots & \\ \frac{\partial p_m}{\partial w_1} & & \frac{\partial p_m}{\partial w_m} \end{pmatrix}} \qquad \frac{\partial L}{\partial \hat{p}} = \overset{m \times 1}{\left(\frac{\partial L}{\partial \hat{p}_1}, \cdots, \frac{\partial L}{\partial \hat{p}_m}\right)^T}$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial \hat{p}} \frac{\partial \hat{p}}{\partial W}$$

$$\frac{\partial \hat{p}}{\partial W} = \overset{n \times m}{\begin{pmatrix} \frac{\partial \hat{p}_1}{\partial w_1} & \cdots & \frac{\partial \hat{p}_m}{\partial w_1} \\ \vdots & \ddots & \\ \frac{\partial \hat{p}_1}{\partial w_m} & & \frac{\partial \hat{p}_m}{\partial w_m} \end{pmatrix}} \qquad \frac{\partial L}{\partial \hat{p}} = \overset{m \times 1}{\left(\frac{\partial L}{\partial \hat{p}_1}, \cdots, \frac{\partial L}{\partial \hat{p}_m}\right)^T}$$

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial \hat{p}} \frac{\partial \hat{p}}{\partial W}$$

$$= \left(\sum_{j=1}^{m} \frac{\partial \hat{p}_j}{\partial w_1} \cdot \frac{\partial L}{\partial \hat{p}_j}, \cdots, \sum_{j=1}^{m} \frac{\partial \hat{p}_j}{\partial w_m} \cdot \frac{\partial L}{\partial \hat{p}_j} \right)$$