

ARO

Name: \_\_\_\_\_

### MLE and backprop examples

---

1. You are given Laplace probability distribution model  $p(y|x, w) = \frac{1}{2} \exp(-|y - \mathbf{w}\bar{\mathbf{x}}|)$ , which models probability of variable  $y \in \mathbb{R}$ , given measurement  $x \in \mathbb{R}$  and unknown model parameters  $w \in \mathbb{R}^2$ . You are given a training set  $\mathcal{D} = \{(x_1, y_1) \dots (x_N, y_N)\}$ .

- **ML regression:** Write down the optimization problem, which corresponds to the maximum likelihood estimate of the model parameters  $\mathbf{w}$ . Simplify resulting loss function  $\mathcal{L}(\mathbf{w})$  if possible.

$$\mathcal{L}(\mathbf{w}) =$$

- **Feed-forward pass:** Draw computational graph of  $\mathcal{L}(\mathbf{w})$  and compute feed-forward pass with the following values:  $\mathbf{w} = [1, -1]^\top$ ,  $\mathbf{x}_1 = 2$ ,  $y_1 = -1$ . Keep vector notation to keep the graph simple.  
**Hint:** assign a variable to each edge and evaluate its value and write it directly into the computational graph.

2. **Backpropagation:** Compute one iteration of the backpropagation algorithm in the computational graph above, with the learning rate  $\alpha = \frac{1}{4}$ . One iteration consists of the following steps:

- (i) compute gradient w.r.t  $\mathbf{w}$  by the backward-pass,
- (ii) update weights  $\mathbf{w}$ ,
- (iii) substitute updated weights and compute the value of the new logistic loss.

- What is the gradient (expression + value) of the back-propagated logistic loss?

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}} =$$

- What are updated weights (expression + value)

$$\mathbf{w}^{\text{updated}} =$$

- What is the value of the updated logistic loss?

$$\mathcal{L}^{\text{updated}} =$$

3. You are given input feature map (image)  $\mathbf{x}$  and kernel  $\mathbf{w}$ :

$$\mathbf{x} = \begin{array}{|c|c|c|} \hline 1 & 0 & 2 \\ \hline 2 & 1 & -1 \\ \hline 0 & 0 & 2 \\ \hline \end{array} \quad \mathbf{w} = \begin{array}{|c|c|} \hline 1 & -1 \\ \hline 0 & 2 \\ \hline \end{array}$$

Compute outputs of the following layers:

- $\text{conv}(\mathbf{x}, \mathbf{w}) =$

- $\text{conv}(\mathbf{x}, \mathbf{w}, \text{stride} = 3, \text{pad} = 1) =$

- $\text{max}(\mathbf{x}, 2 \times 2) =$