VIR 2019 **Name:** ______________________

Test 1

Time Limit: **Points:** ______________________

1. **Computational graph:** Draw computational graph for the learning of the classifier $f(\mathbf{x}, \mathbf{w}) = \mathbf{w}^\top \overline{\mathbf{x}}$ with the logistic loss.
   Hint: Logistic loss of classifier output z for label y is $\mathcal{L}(y, z) = -\log(\sigma(yz))$, where σ is the sigmoid function $\sigma(x) = \frac{1}{1-e^{-x}}$.

2. **Feed-forward pass:** Compute feed-forward pass in the computational graph above (Question 1) with the following values: $\mathbf{w} = [-1, +1, 0]^\top, \mathbf{x} = [2, 1]^\top, y = -1$. 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. Make use of the following table:

v	-4	-3	-2	-1	0	1	2	3	4
$\sigma(v)$	0.02	0.05	0.12	0.27	0.5	0.73	0.88	0.95	0.98
$\log(\sigma(v))$	-4.02	-3.05	-2.13	-1.31	-0.69	-0.31	-0.13	-0.05	- 0.02

   - What is the output value of the feed-forward pass (i.e. the logistic-loss) for the given inputs

     $\mathcal{L} =$

3. **Backpropagation:** Compute one iteration of the backpropagation algorithm in the computational graph above (Question 1), with the learning rate $\alpha = \frac{0.5}{0.27}$. 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.

   Hint: $\frac{d\sigma(z)}{dz} = \sigma(z)(1 - \sigma(z))$

   - 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)

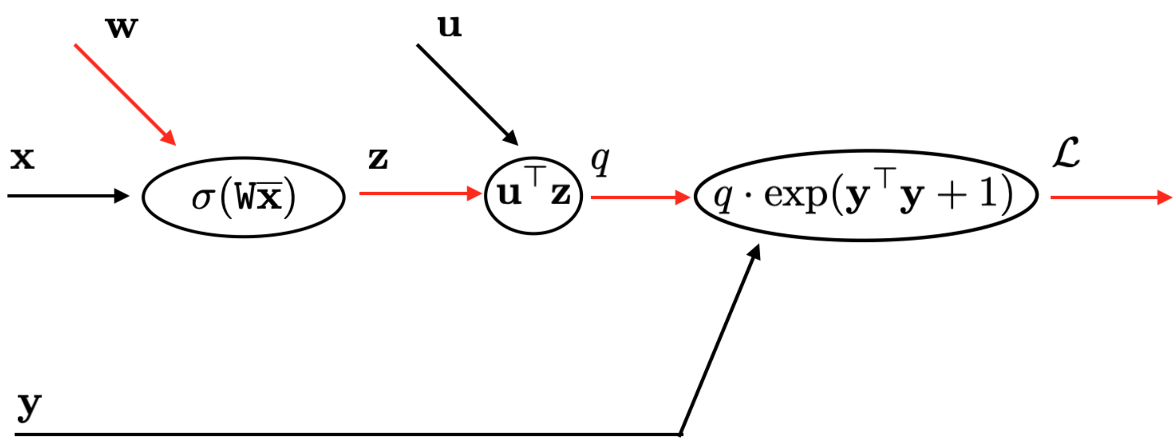
     $\mathbf{w}^{\text{updated}} =$

   - What is the value of the updated logistic loss?

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

4. **Edge jacobians:** You are given the following computational graph, where $\mathtt{W} \in \mathbb{R}^{4\times 6}$, $\mathbf{u} \in \mathbb{R}^{4\times 1}$, $\mathbf{y} \in \mathbb{R}^{17\times 1}$ and $\overline{\mathbf{x}}$ denotes homogeneous coordinates and $\mathbf{w} = \text{vec}(\mathtt{W})$.

   - Fill in dimensionality of $\mathbf{x}, \mathbf{w}, \mathbf{z}, q, \mathcal{L}$ variables in the computational graph.
   - Fill in dimensionality of the following edge gradients $\frac{\partial \mathcal{L}}{\partial q}$, $\frac{\partial q}{\partial \mathbf{z}}$, $\frac{\partial \mathbf{z}}{\partial \mathbf{w}}$.



   - What is dimensionality of the following gradient $\frac{\partial \mathcal{L}}{\partial \mathbf{w}}$ (corresponding edge gradients are emphasized by the red color)?

   - What is dimensionality of $\frac{\partial \mathcal{L}}{\partial \mathbf{y}}$?

5. **ML regression**: You are given probability distribution model $p(y|x, w) = xw\exp(-xwy)$, which models probability of variable $y \in \mathbb{R}^+$, given measurement $x \in \mathbb{R}$ and unknown model parameters $w \in \mathbb{R}$. You are given a training set $\mathcal{D} = \{(x_1, y_1) \ldots (x_N, y_N)\}$. Write down the optimization problem, which corresponds to the maximum likelihood estimate of the model parameters w? Simplify resulting optimization problem if possible.