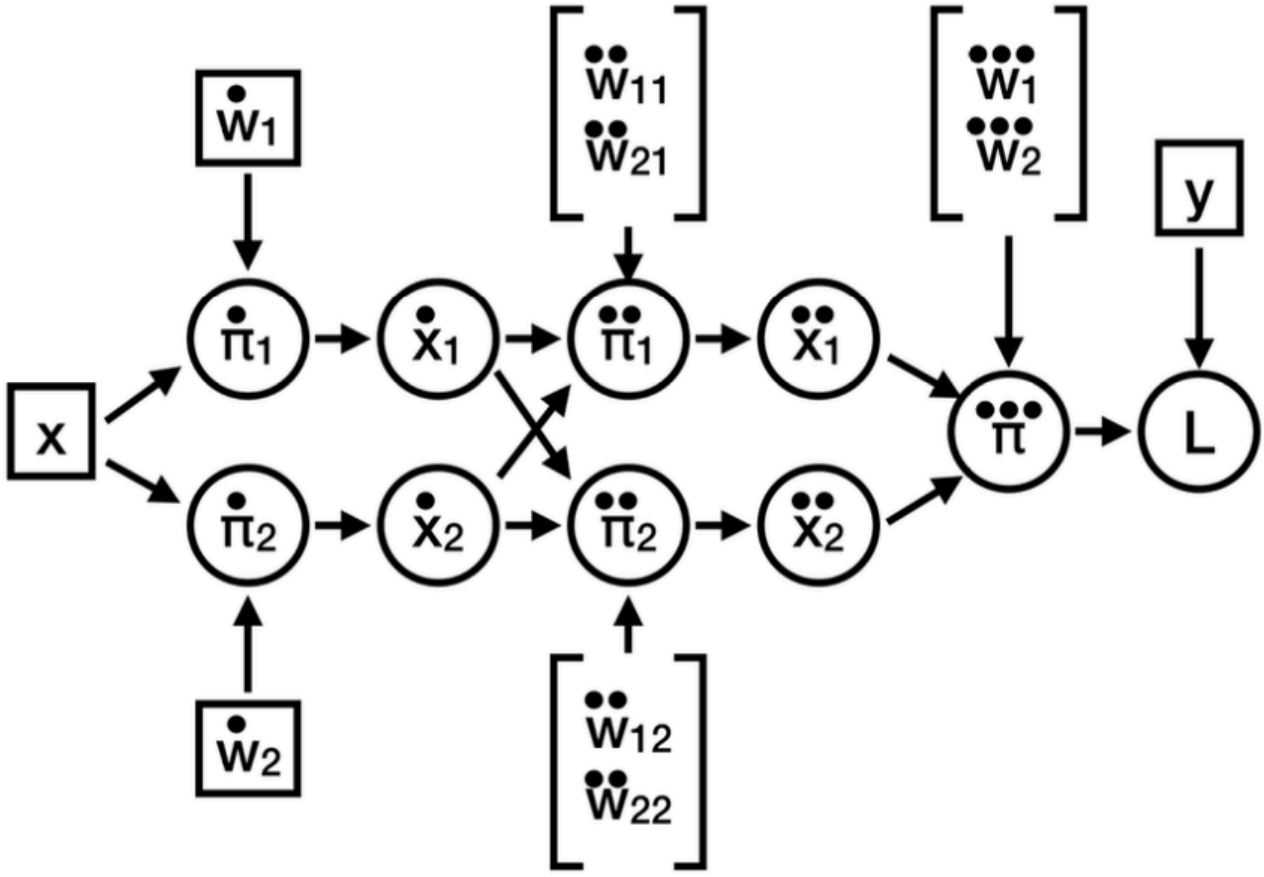
HW15: Backpropagation

Consider the following feedforward neural network.



All explicitly represented quantities are real-valued scalars, e.g. $\ddot{w}_{11}$ is a scalar and $\dot{\pi}_1$ is a scalar. The loss function L is the ordinary linear regression loss, i.e. $L = (y - \dddot{\pi})^2$. The other functions are defined as usual for a feedforward neural network, i.e.:

- $\dot{\pi}_i = w_i * x$
- $\dot{x}_i = a(\dot{\pi}_i)$
- $\ddot{\pi}_i = \ddot{w}_{1i}\dot{x}_1 + \ddot{w}_{2i}\dot{x}_2$
- $\ddot{x}_i = a(\ddot{\pi}_i)$
- $\dddot{\pi} = \dddot{w}_1\ddot{x}_1 + \dddot{w}_2\ddot{x}_2$

where $a(z)$ is the ReLU function.

For a particular datum, $x = -2$ and $y = 7$. Suppose that gradient descent is at the point in the weight space where: $\dot{w}_1 = 2, \dot{w}_2 = -1, \ddot{w}_{11} = 1, \ddot{w}_{21} = 2, \ddot{w}_{12} = 2, \ddot{w}_{22} = 1, \dddot{w}_1 = 1, \dddot{w}_2 = 2$.

What is $\frac{\partial L}{\partial \dot{w}_1}$, evaluated at this point in the weight space? **The answer should be a single number.**

Solution: