Solutions to Tutorial3 Question 2

Calculate backprop equations:

- Network architecture:

$$\begin{aligned}\mathbf{z} &= \mathbf{W}\mathbf{x} \\ \mathbf{h} &= \text{ReLU}(\mathbf{z}) \\ y &= \mathbf{v}^\top \mathbf{h} \\ \mathcal{E} &= \frac{1}{2}(y-t)^2\end{aligned}$$

- Solution:

  Backprop equations in scalar form with indices:

$$\begin{aligned}\overline{\mathcal{E}} &= 1 \\ \overline{y} &= \overline{\mathcal{E}}(y-t) \\ \overline{h_i} &= \overline{y}\frac{\partial y}{\partial h_i} = \overline{y} v_i \\ \overline{z_i} &= \overline{h_i}\frac{\partial h_i}{\partial z_i} = \overline{h_i}\mathbf{I}(z_i > 0) \\ \overline{x_i} &= \sum_j \overline{z_j}\frac{\partial z_j}{\partial x_i} = \sum_j \overline{z_j} w_{ji} \\ \overline{w_{ij}} &= \overline{z_i}\frac{\partial z_i}{\partial w_{ij}} = \overline{z_i} x_j\end{aligned}$$

  Then we write them in vectorized form:

$$\begin{aligned}\overline{\mathcal{E}} &= 1 \\ \overline{y} &= \overline{\mathcal{E}}(y-t) \\ \overline{\mathbf{h}} &= \overline{y}\frac{\partial y}{\partial \mathbf{h}} = \overline{y}\mathbf{v} \\ \overline{\mathbf{z}} &= \overline{\mathbf{h}} \circ \frac{\partial \mathbf{h}}{\partial \mathbf{z}} = \overline{\mathbf{h}} \circ \mathbf{I}(\mathbf{z} > 0) \\ \overline{\mathbf{x}} &= \mathbf{W}^\top \overline{\mathbf{z}} \\ \overline{\mathbf{W}} &= \overline{\mathbf{z}}\mathbf{x}^\top\end{aligned}$$

  where $\mathbf{I}(x > 0) = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$, and $\circ$ denotes hadamard product (elemenetwise multiplication).