

# 1 Exercise 1: Derivatives

(a)

$$\begin{aligned}
 \frac{df(x)}{x} &= \cos(x^3) \frac{d \ln(\frac{1}{11}x^4)}{dx} + (-1) \sin(x^3) 3x^2 \ln(\frac{1}{11}x^4) \\
 &= \cos(x^3) \frac{4}{11} x^3 \frac{1}{\frac{1}{11}x^4} - 3x^2 \sin(x^3) \ln(\frac{1}{11}x^4) \\
 &= 11x^{-4} \frac{4}{11} x^3 \cos(x^3) - 3x^2 \sin(x^3) \ln(\frac{1}{11}x^4) \\
 &= \frac{44}{11} x^{-1} \cos(x^3) - 3x^2 \sin(x^3) \ln(\frac{1}{11}x^4) \\
 &= 4x^{-1} \cos(x^3) - 3x^2 \sin(x^3) \ln(\frac{1}{11}x^4)
 \end{aligned}$$

(b)

$$\begin{aligned}
 \frac{d}{dx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} &= e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{d}{dx} \left( -\frac{(x-\mu)^2}{2\sigma^2} \right) \\
 &= e^{-\frac{(x-\mu)^2}{2\sigma^2}} \left( -\frac{1}{2\sigma^2} \frac{d}{dx} (x-\mu)^2 \right) \\
 &= \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}} 2(x-\mu) \frac{d}{dx} (x-\mu)}{2\sigma^2} \\
 &= \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}} (x-\mu) \left( \frac{d}{dx} (x) + \frac{d}{dx} (-\mu) \right)}{\sigma^2} \\
 &= -\frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}} (x-\mu) (1+0)}{\sigma^2} = -\frac{(x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma^2}
 \end{aligned}$$

(c)

$$\begin{aligned}
 \frac{d}{dx} \frac{1}{1+e^{-x}} &= -\frac{\frac{d}{dx} (e^{-x} + 1)}{(e^{-x} + 1)^2} \\
 &= -\frac{\frac{d}{dx} (e^{-x}) + \frac{d}{dx} (1)}{(e^{-x} + 1)^2} \\
 &= -\frac{e^{-x} \frac{d}{dx} (-x) + 0}{(e^{-x} + 1)^2} \\
 &= \frac{e^{-x}}{(e^{-x} + 1)^2}
 \end{aligned}$$

(d)

Let

$$f(\mathbf{x}) = \begin{pmatrix} \sqrt{x_1^2 - x_3} \\ e^{x_2} \cos(4x_3) \end{pmatrix}, \quad f: \mathbb{R}^3 \rightarrow \mathbb{R}^2.$$

The Jacobian matrix  $\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}$  is given by:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial}{\partial x_1} \sqrt{x_1^2 - x_3} & \frac{\partial}{\partial x_2} \sqrt{x_1^2 - x_3} & \frac{\partial}{\partial x_3} \sqrt{x_1^2 - x_3} \\ \frac{\partial}{\partial x_1} e^{x_2} \cos(4x_3) & \frac{\partial}{\partial x_2} e^{x_2} \cos(4x_3) & \frac{\partial}{\partial x_3} e^{x_2} \cos(4x_3) \end{pmatrix}.$$

Computing the partial derivatives:

$$\begin{aligned} \frac{\partial}{\partial x_1} \sqrt{x_1^2 - x_3} &= \frac{1}{2\sqrt{x_1^2 - x_3}} \cdot 2x_1 = \frac{x_1}{\sqrt{x_1^2 - x_3}}, \\ \frac{\partial}{\partial x_2} \sqrt{x_1^2 - x_3} &= 0, \\ \frac{\partial}{\partial x_3} \sqrt{x_1^2 - x_3} &= \frac{-1}{2\sqrt{x_1^2 - x_3}}, \\ \frac{\partial}{\partial x_1} e^{x_2} \cos(4x_3) &= 0, \\ \frac{\partial}{\partial x_2} e^{x_2} \cos(4x_3) &= e^{x_2} \cos(4x_3), \\ \frac{\partial}{\partial x_3} e^{x_2} \cos(4x_3) &= e^{x_2} \cdot (-\sin(4x_3)) \cdot 4 = -4e^{x_2} \sin(4x_3). \end{aligned}$$

Thus, the Jacobian is:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{x_1}{\sqrt{x_1^2 - x_3}} & 0 & \frac{-1}{2\sqrt{x_1^2 - x_3}} \\ 0 & e^{x_2} \cos(4x_3) & -4e^{x_2} \sin(4x_3) \end{pmatrix}.$$

(e)

Let

$$f(\mathbf{x}) = (a(x_1), \dots, a(x_n)), \quad f: \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

The Jacobian matrix  $\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}$  is:

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial a(x_1)}{\partial x_1} & 0 & \dots & 0 \\ 0 & \frac{\partial a(x_2)}{\partial x_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{\partial a(x_n)}{\partial x_n} \end{pmatrix}.$$

The resulting Jacobian matrix is diagonal, where each diagonal entry is  $a'(x_i)$ . The characteristic feature of this matrix is that it is a \*\*diagonal matrix\*\*.

(f)

Consider the following functions:

$$f_1(\mathbf{x}) = \sin(x_1) \cos(x_2), \quad \mathbf{x} \in \mathbb{R}^2,$$

$$f_2(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^2,$$

$$f_3(\mathbf{x}) = \sin(\log(\mathbf{x}^T \mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^n.$$

Compute the gradient  $\nabla_{\mathbf{x}} f_i(\cdot)$  with respect to  $\mathbf{x}$  for all three functions.

**Gradient of  $f_1(\mathbf{x})$**

$$f_1(\mathbf{x}) = \sin(x_1) \cos(x_2),$$

$$\nabla_{\mathbf{x}} f_1(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} \sin(x_1) \cos(x_2) \\ \frac{\partial}{\partial x_2} \sin(x_1) \cos(x_2) \end{pmatrix}.$$

Compute the partial derivatives:

$$\begin{aligned} \frac{\partial}{\partial x_1} \sin(x_1) \cos(x_2) &= \cos(x_1) \cos(x_2), \\ \frac{\partial}{\partial x_2} \sin(x_1) \cos(x_2) &= -\sin(x_1) \sin(x_2). \end{aligned}$$

Thus, the gradient is:

$$\nabla_{\mathbf{x}} f_1(\mathbf{x}) = \begin{pmatrix} \cos(x_1) \cos(x_2) \\ -\sin(x_1) \sin(x_2) \end{pmatrix}.$$

The dimension of the gradient is  $\mathbb{R}^2$ .

**Gradient of  $f_2(\mathbf{x}, \mathbf{y})$**

$$f_2(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} = \sum_{i=1}^2 x_i y_i.$$

The gradient with respect to  $\mathbf{x}$  is:

$$\nabla_{\mathbf{x}} f_2(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} \frac{\partial}{\partial x_1} (\mathbf{x}^T \mathbf{y}) \\ \frac{\partial}{\partial x_2} (\mathbf{x}^T \mathbf{y}) \end{pmatrix}.$$

Compute the partial derivatives:

$$\frac{\partial}{\partial x_i} (\mathbf{x}^T \mathbf{y}) = y_i, \quad i = 1, 2.$$

Thus, the gradient is:

$$\nabla_{\mathbf{x}} f_2(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{y}.$$

The dimension of the gradient is  $\mathbb{R}^2$ .

### Gradient of $f_3(\mathbf{x})$

$$f_3(\mathbf{x}) = \sin(\log(\mathbf{x}^T \mathbf{x})).$$

Let  $g(\mathbf{x}) = \log(\mathbf{x}^T \mathbf{x})$ , so:

$$f_3(\mathbf{x}) = \sin(g(\mathbf{x})).$$

Using the chain rule:

$$\nabla_{\mathbf{x}} f_3(\mathbf{x}) = \cos(g(\mathbf{x})) \cdot \nabla_{\mathbf{x}} g(\mathbf{x}).$$

Now compute  $\nabla_{\mathbf{x}} g(\mathbf{x})$ :

$$g(\mathbf{x}) = \log(\mathbf{x}^T \mathbf{x}),$$

$$\nabla_{\mathbf{x}} g(\mathbf{x}) = \frac{1}{\mathbf{x}^T \mathbf{x}} \cdot \nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{x}).$$

$$\nabla_{\mathbf{x}} (\mathbf{x}^T \mathbf{x}) = 2\mathbf{x}.$$

Thus:

$$\nabla_{\mathbf{x}} g(\mathbf{x}) = \frac{2\mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

Finally:

$$\nabla_{\mathbf{x}} f_3(\mathbf{x}) = \cos(\log(\mathbf{x}^T \mathbf{x})) \cdot \frac{2\mathbf{x}}{\mathbf{x}^T \mathbf{x}}.$$

The dimension of the gradient is  $\mathbb{R}^n$ .

### Summary of Gradients

$$\begin{aligned}\nabla_{\mathbf{x}} f_1(\mathbf{x}) &\in \mathbb{R}^2, \\ \nabla_{\mathbf{x}} f_2(\mathbf{x}, \mathbf{y}) &\in \mathbb{R}^2, \\ \nabla_{\mathbf{x}} f_3(\mathbf{x}) &\in \mathbb{R}^n.\end{aligned}$$

### (g)

We define:

$$f(\mathbf{z}, \mathbf{y}) := \log p(\mathbf{x}, \mathbf{z}) - \log q(\mathbf{z}, \mathbf{y}),$$

$$\mathbf{z} := t(\mathbf{w}, \mathbf{y}),$$

where the functions  $p, q, t$  are differentiable and:

$$\mathbf{x} \in \mathbb{R}^D, \quad \mathbf{z} \in \mathbb{R}^E, \quad \mathbf{y} \in \mathbb{R}^F, \quad \mathbf{w} \in \mathbb{R}^F.$$

We compute the gradient  $\nabla_{\mathbf{y}} f(\mathbf{z}, \mathbf{y})$  using the chain rule.

Step 1: Expand  $\nabla_{\mathbf{y}} f(\mathbf{z}, \mathbf{y})$  By the chain rule:

$$\nabla_{\mathbf{y}} f(\mathbf{z}, \mathbf{y}) = \nabla_{\mathbf{z}} f(\mathbf{z}, \mathbf{y}) \cdot \frac{\partial \mathbf{z}}{\partial \mathbf{y}} + \frac{\partial f(\mathbf{z}, \mathbf{y})}{\partial \mathbf{y}}.$$

Step 2: Compute  $\nabla_{\mathbf{z}} f(\mathbf{z}, \mathbf{y})$  From the definition of  $f(\mathbf{z}, \mathbf{y})$ :

$$f(\mathbf{z}, \mathbf{y}) = \log p(\mathbf{x}, \mathbf{z}) - \log q(\mathbf{z}, \mathbf{y}),$$

we compute:

$$\nabla_{\mathbf{z}} f(\mathbf{z}, \mathbf{y}) = \nabla_{\mathbf{z}} \log p(\mathbf{x}, \mathbf{z}) - \nabla_{\mathbf{z}} \log q(\mathbf{z}, \mathbf{y}).$$

For each term:

$$\nabla_{\mathbf{z}} \log p(\mathbf{x}, \mathbf{z}) = \frac{1}{p(\mathbf{x}, \mathbf{z})} \nabla_{\mathbf{z}} p(\mathbf{x}, \mathbf{z}),$$

$$\nabla_{\mathbf{z}} \log q(\mathbf{z}, \mathbf{y}) = \frac{1}{q(\mathbf{z}, \mathbf{y})} \nabla_{\mathbf{z}} q(\mathbf{z}, \mathbf{y}).$$

Thus:

$$\nabla_{\mathbf{z}} f(\mathbf{z}, \mathbf{y}) = \frac{\nabla_{\mathbf{z}} p(\mathbf{x}, \mathbf{z})}{p(\mathbf{x}, \mathbf{z})} - \frac{\nabla_{\mathbf{z}} q(\mathbf{z}, \mathbf{y})}{q(\mathbf{z}, \mathbf{y})}.$$

Step 3: Compute  $\frac{\partial \mathbf{z}}{\partial \mathbf{y}}$  From  $\mathbf{z} = t(\mathbf{w}, \mathbf{y})$ , we compute:

$$\frac{\partial \mathbf{z}}{\partial \mathbf{y}} = \nabla_{\mathbf{y}} t(\mathbf{w}, \mathbf{y}),$$

which is the Jacobian matrix of  $t(\mathbf{w}, \mathbf{y})$  with respect to  $\mathbf{y}$ .

Step 4: Compute  $\frac{\partial f(\mathbf{z}, \mathbf{y})}{\partial \mathbf{y}}$  For the second term in  $f(\mathbf{z}, \mathbf{y})$ , only  $\log q(\mathbf{z}, \mathbf{y})$  depends explicitly on  $\mathbf{y}$ . Thus:

$$\frac{\partial f(\mathbf{z}, \mathbf{y})}{\partial \mathbf{y}} = -\nabla_{\mathbf{y}} \log q(\mathbf{z}, \mathbf{y}),$$

and:

$$\nabla_{\mathbf{y}} \log q(\mathbf{z}, \mathbf{y}) = \frac{1}{q(\mathbf{z}, \mathbf{y})} \nabla_{\mathbf{y}} q(\mathbf{z}, \mathbf{y}).$$

Step 5: Combine Results Substituting all terms into the chain rule expansion:

$$\nabla_{\mathbf{y}} f(\mathbf{z}, \mathbf{y}) = \left( \frac{\nabla_{\mathbf{z}} p(\mathbf{x}, \mathbf{z})}{p(\mathbf{x}, \mathbf{z})} - \frac{\nabla_{\mathbf{z}} q(\mathbf{z}, \mathbf{y})}{q(\mathbf{z}, \mathbf{y})} \right) \nabla_{\mathbf{y}} t(\mathbf{w}, \mathbf{y}) - \frac{\nabla_{\mathbf{y}} q(\mathbf{z}, \mathbf{y})}{q(\mathbf{z}, \mathbf{y})}.$$

This is the gradient  $\nabla_{\mathbf{y}} f(\mathbf{z}, \mathbf{y})$ .

(h)

Let:

$$f(\mathbf{y}) = \sin(\mathbf{y}), \quad \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b},$$

where  $\mathbf{A} \in \mathbb{R}^{M \times N}$ ,  $\mathbf{x} \in \mathbb{R}^N$ , and  $\mathbf{b} \in \mathbb{R}^M$ . The sine function,  $\sin(\cdot)$ , is applied element-wise. Compute the derivative  $\frac{df}{d\mathbf{x}}$  using the chain rule.

Step 1: Apply the Chain Rule Using the chain rule:

$$\frac{df}{d\mathbf{x}} = \frac{\partial f}{\partial \mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial \mathbf{x}}.$$

Step 2: Compute  $\frac{\partial f}{\partial \mathbf{y}}$  The function  $f(\mathbf{y}) = \sin(\mathbf{y})$  is element-wise. Its derivative with respect to  $\mathbf{y}$  is also element-wise:

$$\frac{\partial f}{\partial \mathbf{y}} = \cos(\mathbf{y}).$$

This is a diagonal matrix, as the derivatives are taken element-wise:

$$\frac{\partial f}{\partial \mathbf{y}} = \text{diag}(\cos(y_1), \cos(y_2), \dots, \cos(y_M)),$$

where  $\mathbf{y} = (y_1, y_2, \dots, y_M)^T$ .

Step 3: Compute  $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$  From  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ , the derivative is:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{A}.$$

Step 4: Combine Results Substitute into the chain rule:

$$\frac{df}{d\mathbf{x}} = \frac{\partial f}{\partial \mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial \mathbf{x}}.$$

Substituting the values:

$$\frac{df}{d\mathbf{x}} = \text{diag}(\cos(y_1), \cos(y_2), \dots, \cos(y_M)) \cdot \mathbf{A}.$$

Since  $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ , we can write this as:

$$\frac{df}{d\mathbf{x}} = \text{diag}(\cos(\mathbf{A}\mathbf{x} + \mathbf{b})) \cdot \mathbf{A}.$$

Step 5: Dimension of  $\frac{df}{d\mathbf{x}}$  The resulting matrix  $\frac{df}{d\mathbf{x}}$  has the same dimensions as  $\mathbf{A}$ , i.e.,  $M \times N$ .

Final Answer

$$\frac{df}{d\mathbf{x}} = \text{diag}(\cos(\mathbf{A}\mathbf{x} + \mathbf{b})) \cdot \mathbf{A}, \quad \text{dimension: } M \times N.$$

## 2. Computational Graph and Backpropagation

### (a) Introducing Intermediate Variables

For the equations:

$$A = \sqrt{a + b + c^2} + \log(a + b + c^2) + \frac{a + b + c^2}{bc^2},$$

$$B = \sum_{i=1}^3 (w_0 + w_1 x_i - y_i)^2,$$

we introduce the following intermediate variables:

For  $A$ :

$$t_1 = a + b + c^2,$$

$$t_2 = \sqrt{t_1},$$

$$t_3 = \log(t_1),$$

$$t_4 = bc^2,$$

$$t_5 = \frac{t_1}{t_4},$$

$$A = t_2 + t_3 + t_5.$$

For  $B$ :

$$t_6^i = w_0 + w_1 x_i, \quad i = 1, 2, 3,$$

$$t_7^i = t_6^i - y_i, \quad i = 1, 2, 3,$$

$$t_8^i = (t_7^i)^2, \quad i = 1, 2, 3,$$

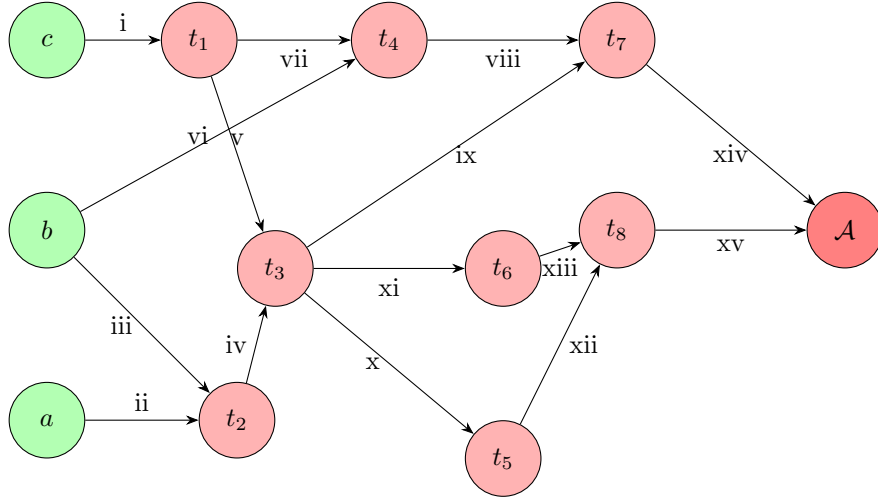
$$B = \sum_{i=1}^3 t_8^i.$$

### (b) Drawing the Computational Graph

The computational graph consists of the following steps:

1. Nodes for the inputs:  $a, b, c, w_0, w_1, x_i, y_i$ . 2. Nodes for intermediate variables  $t_1, t_2, \dots, t_8^i$  as defined above. 3. Final nodes for  $A$  and  $B$ .

## Computational Graph for $A$ and $B$



(a)

$$\begin{aligned}
 t_1 &= c^2, \\
 t_2 &= a + b, \\
 t_3 &= t_1 + t_2, \\
 t_4 &= b \cdot t_1, \\
 t_5 &= \sqrt{t_3}, \\
 t_6 &= \log(t_3), \\
 t_7 &= \frac{t_3}{t_4}, \\
 t_8 &= t_5 + t_6, \\
 \mathcal{A} &= t_7 + t_8.
 \end{aligned}$$



$$\begin{aligned}
\text{xv} &= \frac{\partial \mathcal{A}}{\partial \mathcal{A}} \frac{\partial \mathcal{A}}{\partial t_8} \\
\text{xiv} &= \frac{\partial \mathcal{A}}{\partial \mathcal{A}} \frac{\partial \mathcal{A}}{\partial t_7} \\
\text{xiii} &= \frac{\partial \mathcal{A}}{\partial t_8} \frac{\partial t_8}{\partial t_6} \\
\text{xii} &= \frac{\partial \mathcal{A}}{\partial t_8} \frac{\partial t_8}{\partial t_5} \\
\text{xi} &= \frac{\partial \mathcal{A}}{\partial t_6} \frac{\partial t_6}{\partial t_3} \\
\text{x} &= \frac{\partial \mathcal{A}}{\partial t_5} \frac{\partial t_5}{\partial t_3} \\
\text{ix} &= \frac{\partial \mathcal{A}}{\partial t_7} \frac{\partial t_7}{\partial t_3} \\
\text{viii} &= \frac{\partial \mathcal{A}}{\partial t_7} \frac{\partial t_7}{\partial t_4} \\
\text{vii} &= \frac{\partial \mathcal{A}}{\partial t_3} \frac{\partial t_3}{\partial t_2} \\
\text{vi} &= \frac{\partial \mathcal{A}}{\partial t_3} \frac{\partial t_3}{\partial t_1} \\
\text{v} &= \frac{\partial \mathcal{A}}{\partial t_4} \frac{\partial t_4}{\partial t_1} \\
\text{iv} &= \frac{\partial \mathcal{A}}{\partial t_4} \frac{\partial t_4}{\partial b} \\
\text{iii} &= \frac{\partial \mathcal{A}}{\partial t_2} \frac{\partial t_2}{\partial b} \\
\text{ii} &= \frac{\partial \mathcal{A}}{\partial t_2} \frac{\partial t_2}{\partial a} \\
\text{i} &= \frac{\partial \mathcal{A}}{\partial t_1} \frac{\partial t_1}{\partial c}
\end{aligned}$$

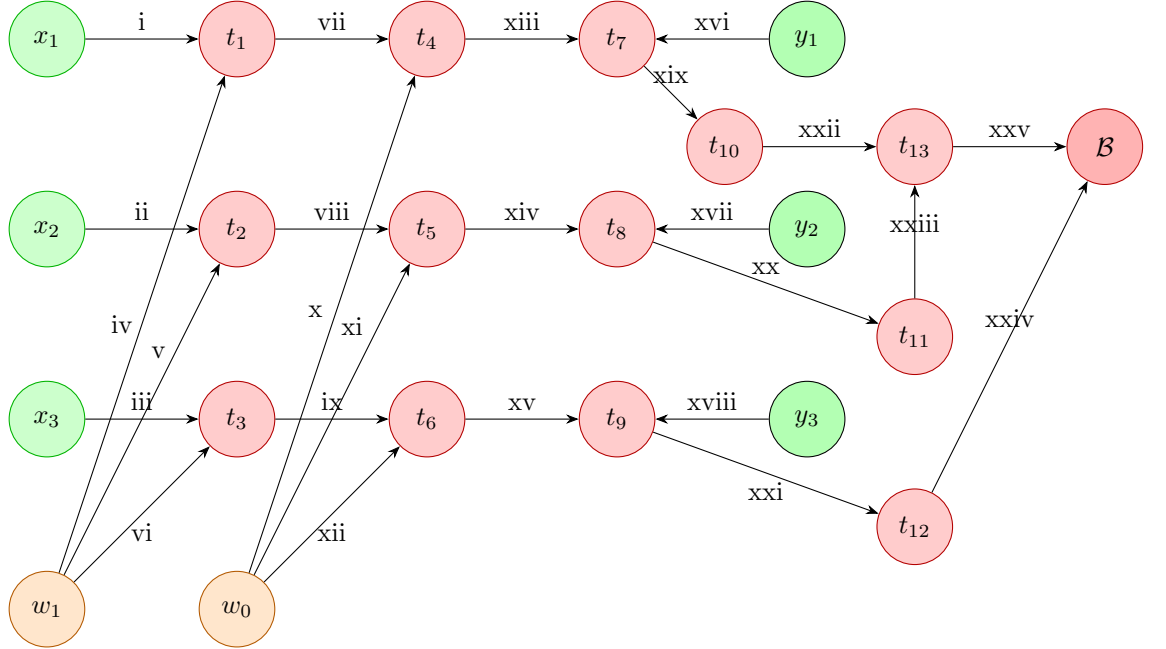
$$\begin{aligned}
\mathcal{A} &: \frac{\partial \mathcal{A}}{\partial \mathcal{A}} \\
t_8 &: \frac{\partial \mathcal{A}}{\partial t_8} \\
t_7 &: \frac{\partial \mathcal{A}}{\partial t_7} \\
t_6 &: \frac{\partial \mathcal{A}}{\partial t_6} \\
t_5 &: \frac{\partial \mathcal{A}}{\partial t_5} \\
t_4 &: \frac{\partial \mathcal{A}}{\partial t_4} \\
t_3 &: \frac{\partial \mathcal{A}}{\partial t_7} \frac{\partial t_7}{\partial t_3} + \frac{\partial \mathcal{A}}{\partial t_6} \frac{\partial t_6}{\partial t_3} + \frac{\partial \mathcal{A}}{\partial t_5} \frac{\partial t_5}{\partial t_3} \\
t_2 &: \frac{\partial \mathcal{A}}{\partial t_2} \\
t_1 &: \frac{\partial \mathcal{A}}{\partial t_4} \frac{\partial t_4}{\partial t_1} + \frac{\partial \mathcal{A}}{\partial t_3} \frac{\partial t_3}{\partial t_1} \\
c &: \frac{\partial \mathcal{A}}{\partial c} \\
b &: \frac{\partial \mathcal{A}}{\partial t_4} \frac{\partial t_4}{\partial b} + \frac{\partial \mathcal{A}}{\partial t_2} \frac{\partial t_2}{\partial b} \\
a &: \frac{\partial \mathcal{A}}{\partial a}
\end{aligned}$$

### (a) Computations

$$\mathcal{B} = \sum_{i=1}^3 (w_0 + w_1 x_i - y_i)^2$$

$$\begin{aligned}
t_1 &= w_1 \cdot x_1, \\
t_2 &= w_1 \cdot x_2, \\
t_3 &= w_1 \cdot x_3, \\
t_4 &= w_0 + t_1, \\
t_5 &= w_0 + t_2, \\
t_6 &= w_0 + t_3, \\
t_7 &= t_4 - y_1, \\
t_8 &= t_5 - y_2, \\
t_9 &= t_6 - y_3, \\
t_{10} &= t_7^2, \\
t_{11} &= t_8^2, \\
t_{12} &= t_9^2, \\
t_{13} &= t_{10} + t_{11}, \\
\mathcal{B} &= t_{13} + t_{12}.
\end{aligned}$$

**(b) Computational Graph**



$$\begin{aligned}
\text{xxv} &:= \frac{\partial \mathcal{B}}{\mathcal{B}} \frac{\partial B}{\partial t_{13}}, & \text{xxiv} &:= \frac{\partial \mathcal{B}}{\partial \mathcal{B}} \frac{\partial \mathcal{B}}{\partial t_{12}}, & \text{xxiii} &:= \frac{\partial \mathcal{B}}{\partial \sqcup_{\infty \exists}} \frac{\partial t_{13}}{\partial t_{11}}, \\
\text{xxii} &:= \frac{\partial \mathcal{B}}{\partial t_{13}} \frac{\partial t_{13}}{\partial t_{10}}, & \text{xxi} &:= \frac{\partial \mathcal{B}}{\partial t_{12}} \frac{\partial t_{12}}{\partial t_9}, & \text{xx} &:= \frac{\partial \mathcal{B}}{\partial t_{11}} \frac{\partial t_{11}}{\partial t_8}, \\
\text{xix} &:= \frac{\partial \mathcal{B}}{\partial t_{10}} \frac{\partial t_{10}}{\partial t_7}, & \text{xviii} &:= \frac{\partial \mathcal{B}}{\partial t_9} \frac{\partial t_9}{\partial t_3}, & \text{xvii} &:= \frac{\partial \mathcal{B}}{\partial t_8} \frac{\partial t_8}{\partial y_2}, \\
\text{xvi} &:= \frac{\partial \mathcal{B}}{\partial t_7} \frac{\partial t_7}{\partial y_1}, & \text{xv} &:= \frac{\partial \mathcal{B}}{\partial t_9} \frac{\partial t_9}{\partial t_6}, & \text{xiv} &:= \frac{\partial \mathcal{B}}{\partial t_8} \frac{\partial t_8}{\partial t_5}, \\
\text{xiii} &:= \frac{\partial \mathcal{B}}{\partial t_7} \frac{\partial t_7}{\partial t_4}, & \text{xii} &:= \frac{\partial \mathcal{B}}{\partial t_6} \frac{\partial t_6}{\partial w_0}, & \text{xi} &:= \frac{\partial \mathcal{B}}{\partial t_5} \frac{\partial t_5}{\partial w_0}, \\
\text{x} &:= \frac{\partial \mathcal{B}}{\partial t_4} \frac{\partial t_4}{\partial w_0}, \\
\text{ix} &:= \frac{\partial \mathcal{B}}{\partial t_6} \frac{\partial t_6}{\partial t_3}, \\
\text{viii} &:= \frac{\partial \mathcal{B}}{\partial t_5} \frac{\partial t_5}{\partial t_2}, \\
\text{vii} &:= \frac{\partial \mathcal{B}}{\partial t_4} \frac{\partial t_4}{\partial t_1}, \\
\text{vi} &:= \frac{\partial \mathcal{B}}{\partial t_3} \frac{\partial t_3}{\partial w_1}, \\
\text{v} &:= \frac{\partial \mathcal{B}}{\partial t_2} \frac{\partial t_2}{\partial w_1}, \\
\text{iv} &:= \frac{\partial \mathcal{B}}{\partial t_1} \frac{\partial t_1}{\partial w_1}, \\
\text{iii} &:= \frac{\partial \mathcal{B}}{\partial t_3} \frac{\partial t_3}{\partial x_3}, \\
\text{ii} &:= \frac{\partial \mathcal{B}}{\partial t_2} \frac{\partial t_2}{\partial x_2}, \\
\text{i} &:= \frac{\partial \mathcal{B}}{\partial t_1} \frac{\partial t_1}{\partial x_1}
\end{aligned}$$

$$\begin{aligned}
\mathcal{B} &: \frac{\partial \mathcal{B}}{\partial \mathcal{B}} \\
t_{13} &: \frac{\partial \mathcal{B}}{\partial t_{13}}, \quad t_{12} : \frac{\partial \mathcal{B}}{\partial t_{12}}, \\
t_{11} &: \frac{\partial \mathcal{B}}{\partial t_{11}}, \quad t_{10} : \frac{\partial \mathcal{B}}{\partial t_{10}}, \\
t_9 &: \frac{\partial \mathcal{B}}{\partial t_9}, \quad t_8 : \frac{\partial \mathcal{B}}{\partial t_8}, \\
t_7 &: \frac{\partial \mathcal{B}}{\partial t_7}, \quad t_6 : \frac{\partial \mathcal{B}}{\partial t_6}, \\
t_5 &: \frac{\partial \mathcal{B}}{\partial t_5}, \quad t_4 : \frac{\partial \mathcal{B}}{\partial t_4}, \\
t_3 &: \frac{\partial \mathcal{B}}{\partial t_3}, \quad t_2 : \frac{\partial \mathcal{B}}{\partial t_2}, \\
t_1 &: \frac{\partial \mathcal{B}}{\partial t_1} \\
x_3 &: \frac{\partial \mathcal{B}}{x_3} \\
x_2 &: \frac{\partial \mathcal{B}}{x_2} \\
x_1 &: \frac{\partial \mathcal{B}}{x_1} \\
y_3 &: \frac{\partial \mathcal{B}}{y_3} \\
y_2 &: \frac{\partial \mathcal{B}}{y_2} \\
y_1 &: \frac{\partial \mathcal{B}}{y_1} \\
w_1 &: \frac{\partial \mathcal{B}}{\partial t_3} \frac{\partial t_3}{\partial w_1} + \frac{\partial \mathcal{B}}{\partial t_2} \frac{\partial t_2}{w_1} + \frac{\partial \mathcal{B}}{\partial t_1} \frac{\partial t_1}{\partial w_1} \\
w_0 &: \frac{\partial \mathcal{B}}{\partial t_6} \frac{\partial t_6}{\partial w_0} + \frac{\partial \mathcal{B}}{\partial t_5} \frac{\partial t_5}{w_0} + \frac{\partial \mathcal{B}}{\partial t_4} \frac{\partial t_4}{\partial w_0}
\end{aligned}$$