Глубокое обучение для обработки изображений

Лекция 2

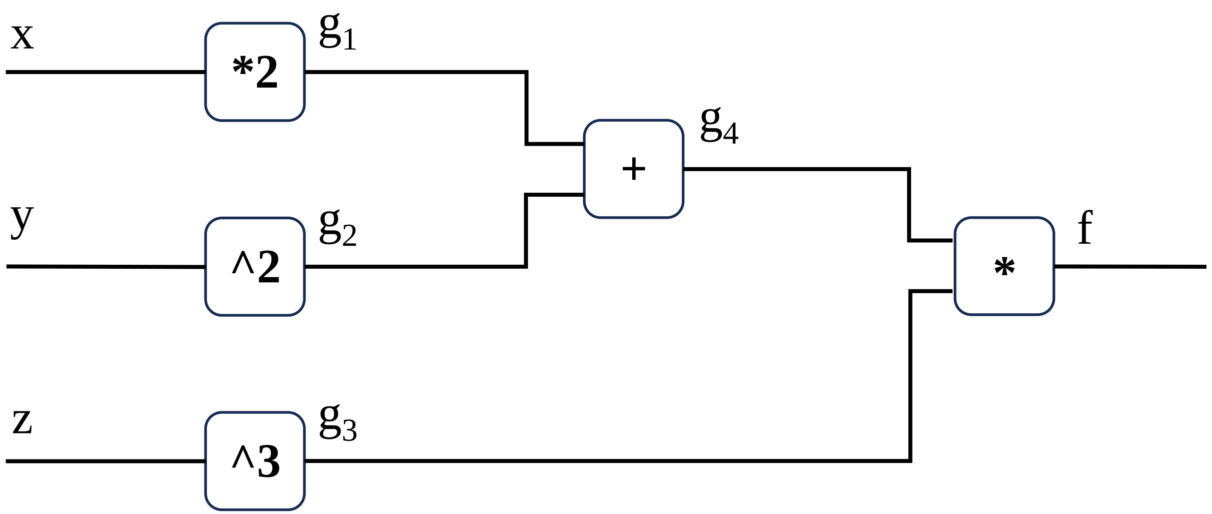
Метод обратного распространения ошибки

При обучении нейронных сетей мы хотим найти оптимальные параметры модели. Для этого используются **градиентные методы**; с помощью них мы итеративно обновляем веса нашей модели. Для обновления весов нам необходимо знать градиенты функции потерь, которые можно найти с помощью алгоритма **обратного распространения ошибки (backpropagation).**

Процесс обучения будет состоять таким образом из следующих шагов:

- На вход подаются данные, которые передаются по сети в прямом направлении, в результате чего получаются выходные данные (**Forward pass**)
- Сигнал ошибки передается в обратном направлении, находятся значения градиентов (**Backward pass**)
- С помощью оптимизатора (SGD, Adam и др.) обновляем значения весов

Пример



Пусть дана функция многих переменных:

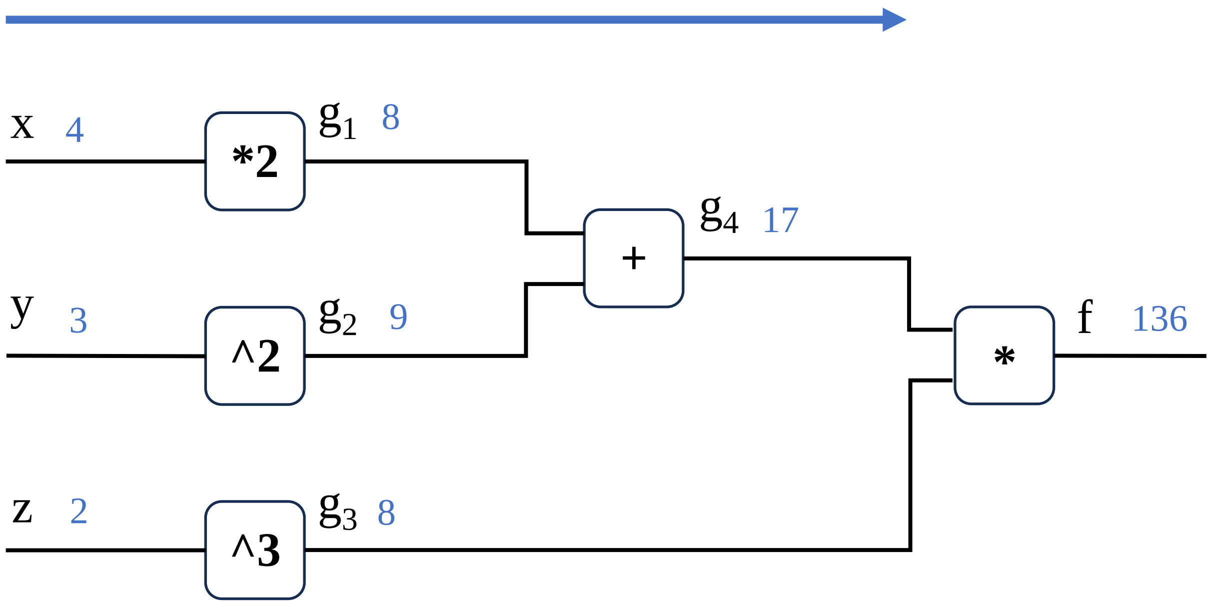
$$f(x, y, z) = (2x + y^2)z^3$$

Необходимо найти производные:

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$

Удобно представить функцию в виде вычислительного графа

Пример



$$f(x,y,z) = (2x + y^2)z^3$$

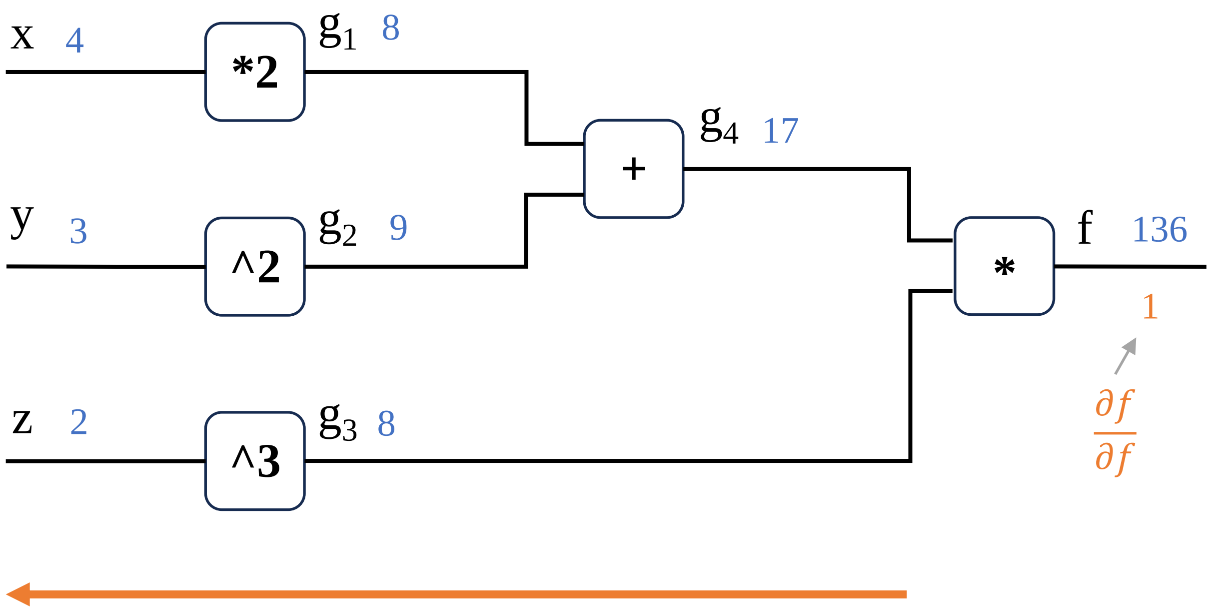
Сначала найдем значение функции при некоторых входных значениях – **forward pass.** Пусть $x = 4, y = 3, z = 2$

Для удобства введем обозначения для промежуточных математических операций

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x,y,z) = g_3 g_4$$

Пример



$$f(x, y, z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x, y, z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

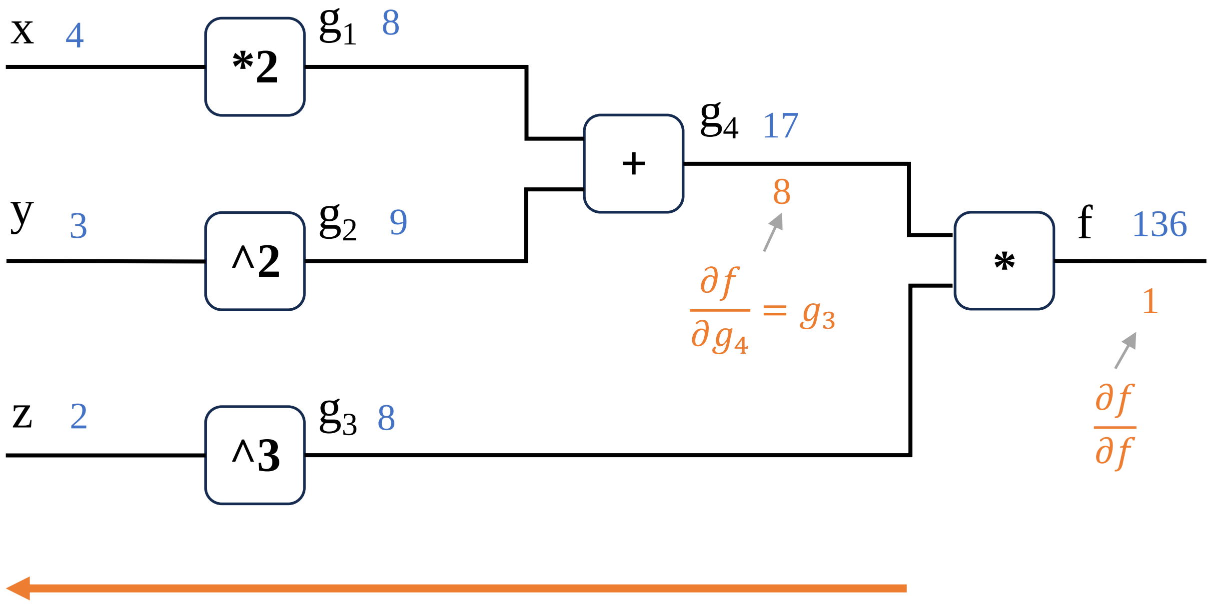
В нашем примере производные имеют вид:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial g_4} \frac{\partial g_4}{\partial g_1} \frac{\partial g_1}{\partial x}$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial g_4} \frac{\partial g_4}{\partial g_2} \frac{\partial g_2}{\partial y}$$

$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial g_3} \frac{\partial g_3}{\partial z}$$

Пример



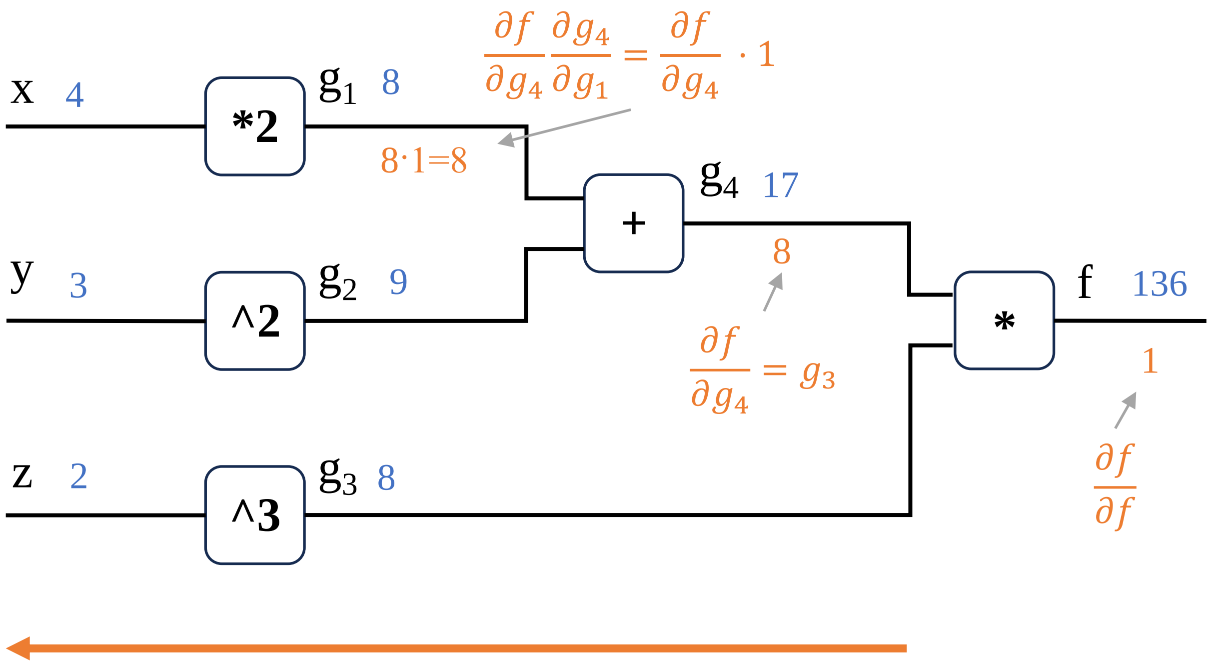
$$f(x,y,z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x,y,z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



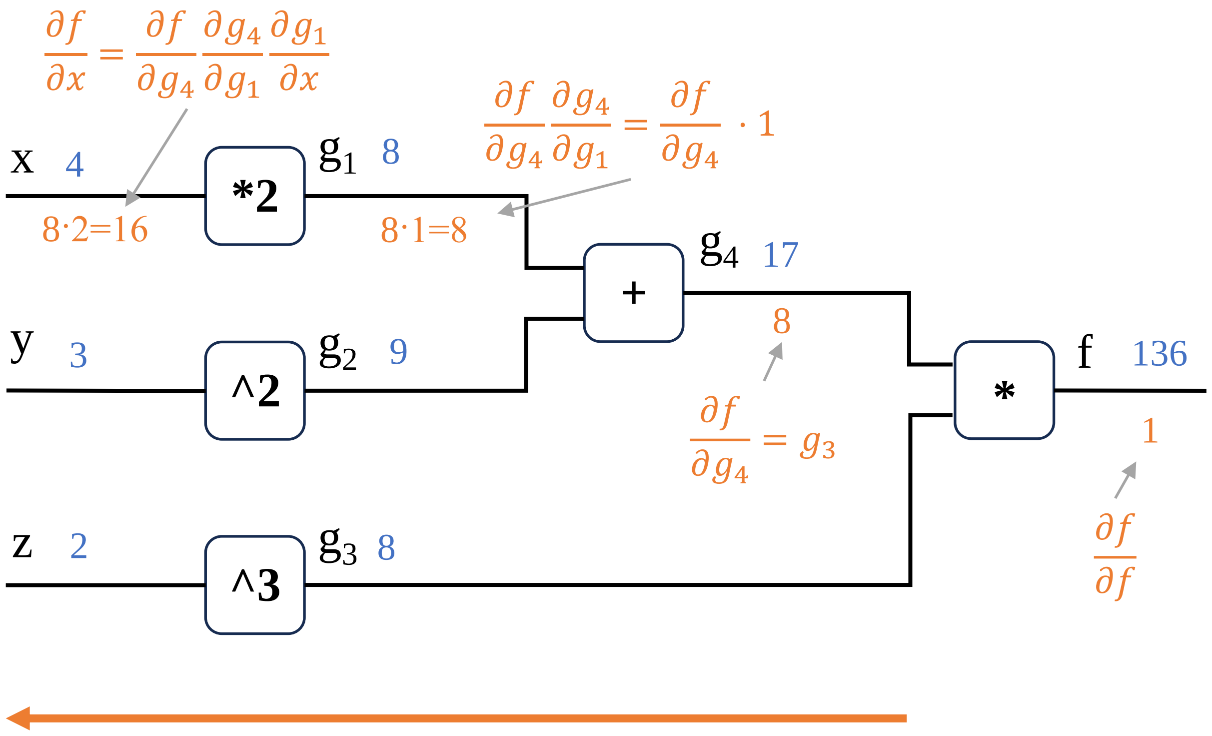
$$f(x,y,z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x,y,z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



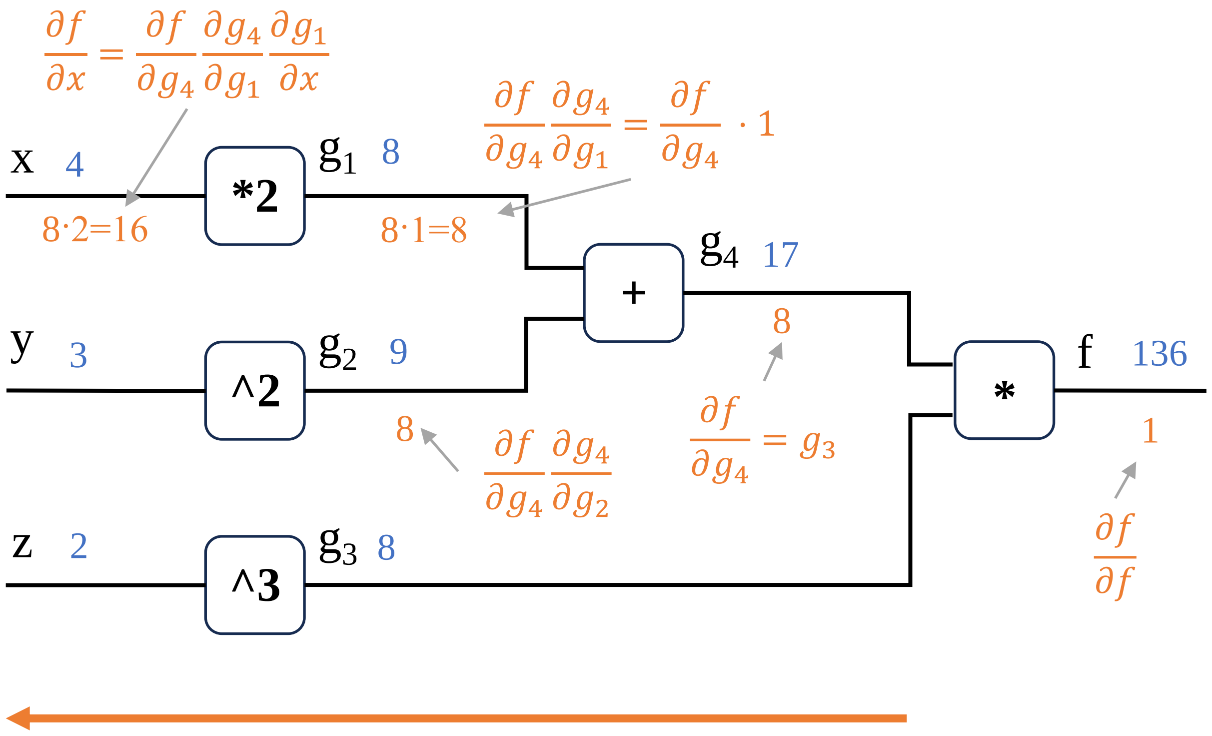
$$f(x, y, z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x, y, z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



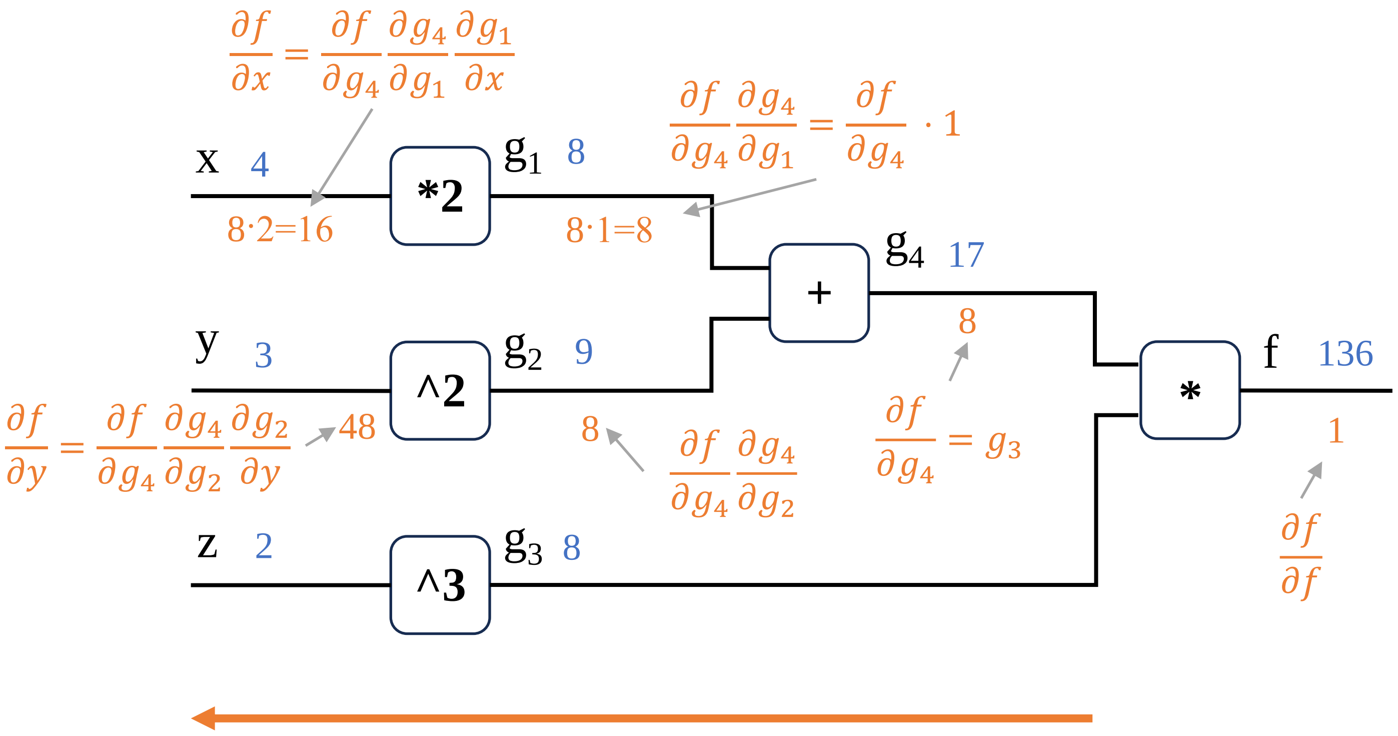
$$f(x, y, z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x, y, z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



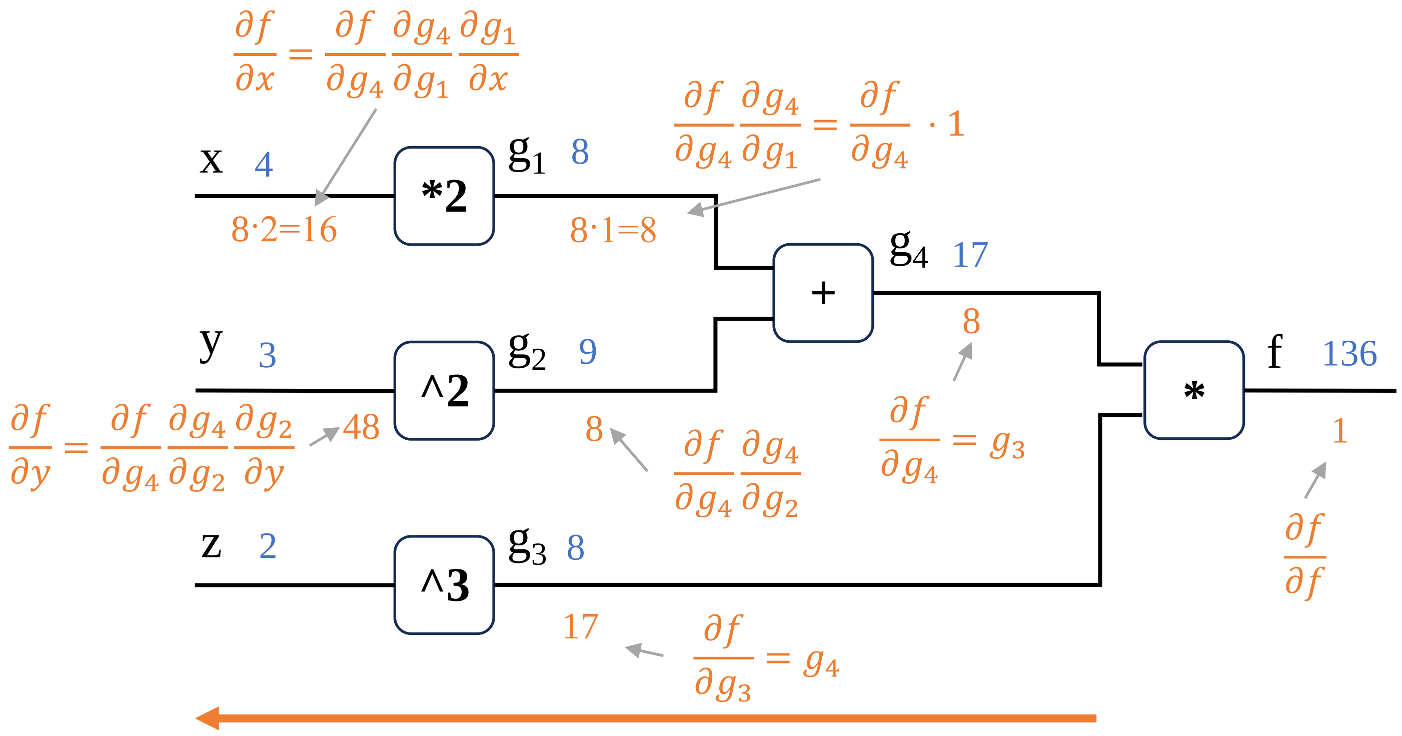
$$f(x,y,z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x,y,z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



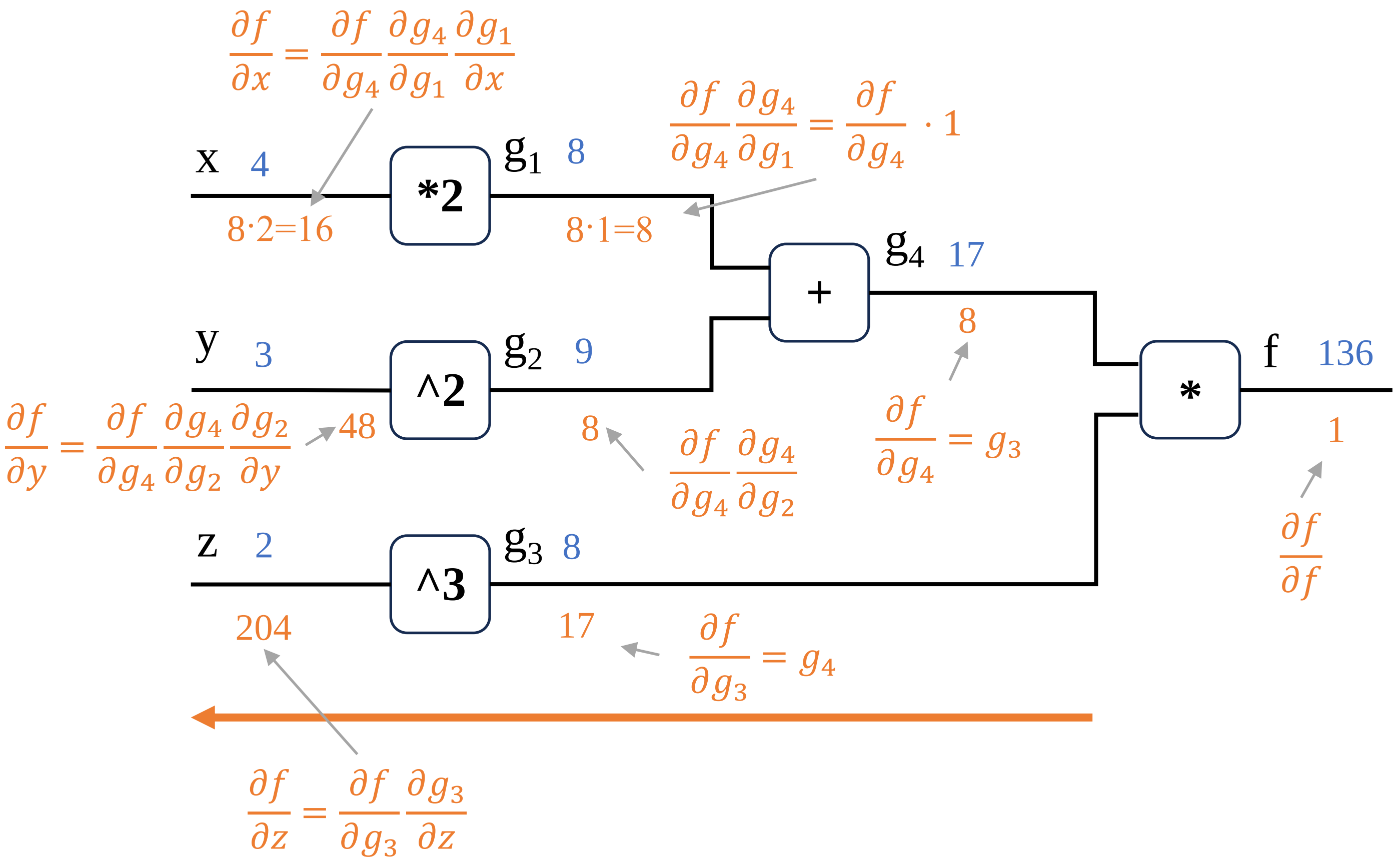
$$f(x, y, z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x, y, z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример



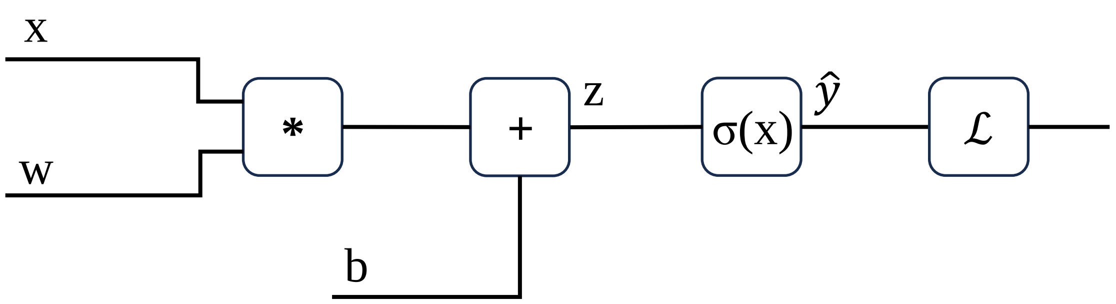
$$f(x,y,z) = (2x + y^2)z^3$$

$$g_1 = 2x$$
$$g_2 = y^2$$
$$g_3 = z^3$$
$$g_4 = g_1 + g_2$$

$$f(x,y,z) = g_3 g_4$$

Теперь, зная результат прямого прохода, будем искать градиенты – **backward pass**

Пример: логистическая регрессия



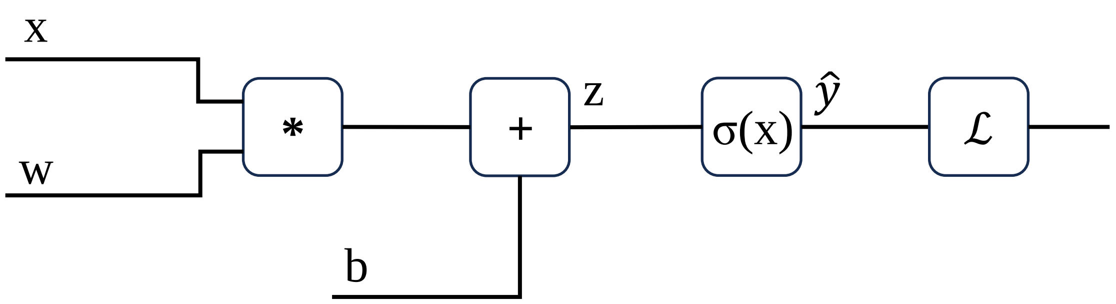
Рассмотрим упрощенный пример логистической регрессии, где все величины являются скалярами. В качестве функции потерь выберем *бинарную кросс-энтропию.* Тогда лосс можем найти следующим образом:

$$z = wx + b$$
$$\hat{y} = \sigma(z)$$
$$\mathcal{L} = -y\ln\hat{y} - (1-y)\ln(1-\hat{y})$$

Градиенты имеют вид:

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

Пример: логистическая регрессия



Рассмотрим упрощенный пример логистической регрессии, где все величины являются скалярами. В качестве функции потерь выберем *бинарную кросс-энтропию*. Тогда лосс можем найти следующим образом:

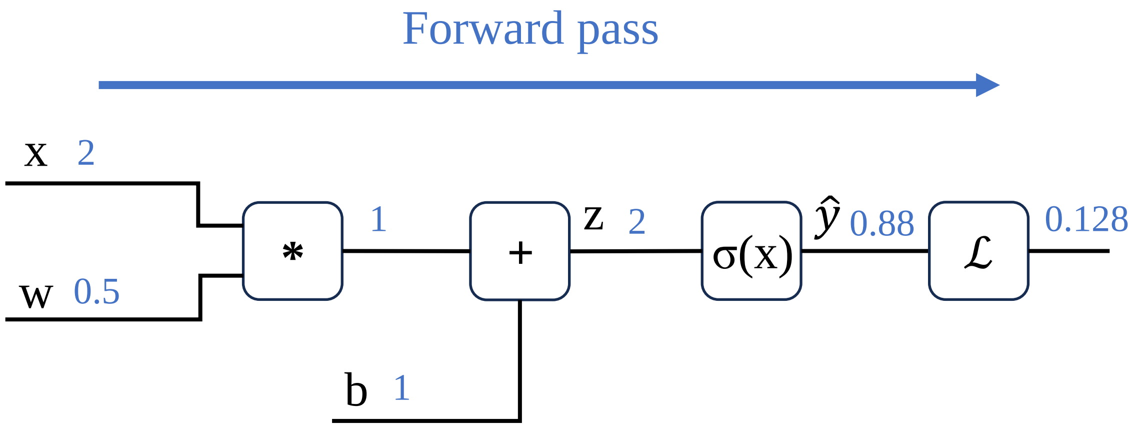
$$z = wx + b$$
$$\hat{y} = \sigma(z)$$
$$\mathcal{L} = -y\ln\hat{y} - (1-y)\ln(1-\hat{y})$$

Градиенты имеют вид:

$$\frac{\partial\mathcal{L}}{\partial\hat{y}} = -\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}$$
$$\frac{\partial\mathcal{L}}{\partial z} = \frac{\partial\mathcal{L}}{\partial\hat{y}}\sigma'(z) = \frac{\partial\mathcal{L}}{\partial\hat{y}}\sigma(z)(1-\sigma(z))$$
$$\frac{\partial\mathcal{L}}{\partial w} = \frac{\partial\mathcal{L}}{\partial z}x$$
$$\frac{\partial\mathcal{L}}{\partial b} = \frac{\partial\mathcal{L}}{\partial z}$$

Задача: найти $\frac{\partial\mathcal{L}}{\partial w}$ и $\frac{\partial\mathcal{L}}{\partial b}$

Пример: логистическая регрессия



$z = wx + b$

$\hat{y} = \sigma(z)$

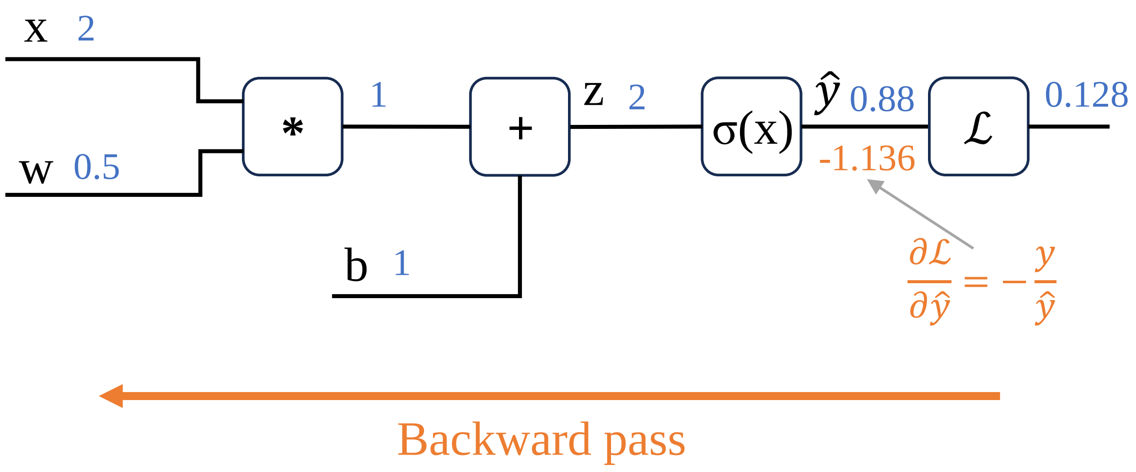
$\mathcal{L} = -y\ln\hat{y} - (1 - y)\ln(1 - \hat{y})$

$\boldsymbol{w = 0.5, b = 1}$

$\boldsymbol{x = 2, y = 1}$

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

Пример: логистическая регрессия



$z = wx + b$

$\hat{y} = \sigma(z)$

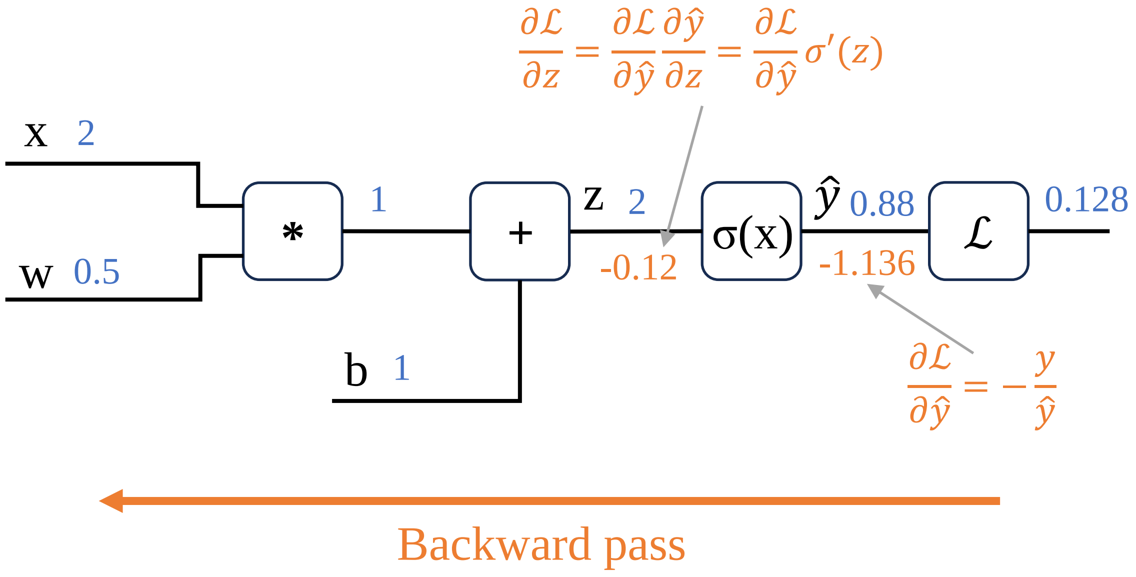
$\mathcal{L} = -y\ln\hat{y} - (1-y)\ln(1-\hat{y})$

$\boldsymbol{w = 0.5, b = 1}$

$\boldsymbol{x = 2, y = 1}$

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

Пример: логистическая регрессия



$z = wx + b$

$\hat{y} = \sigma(z)$

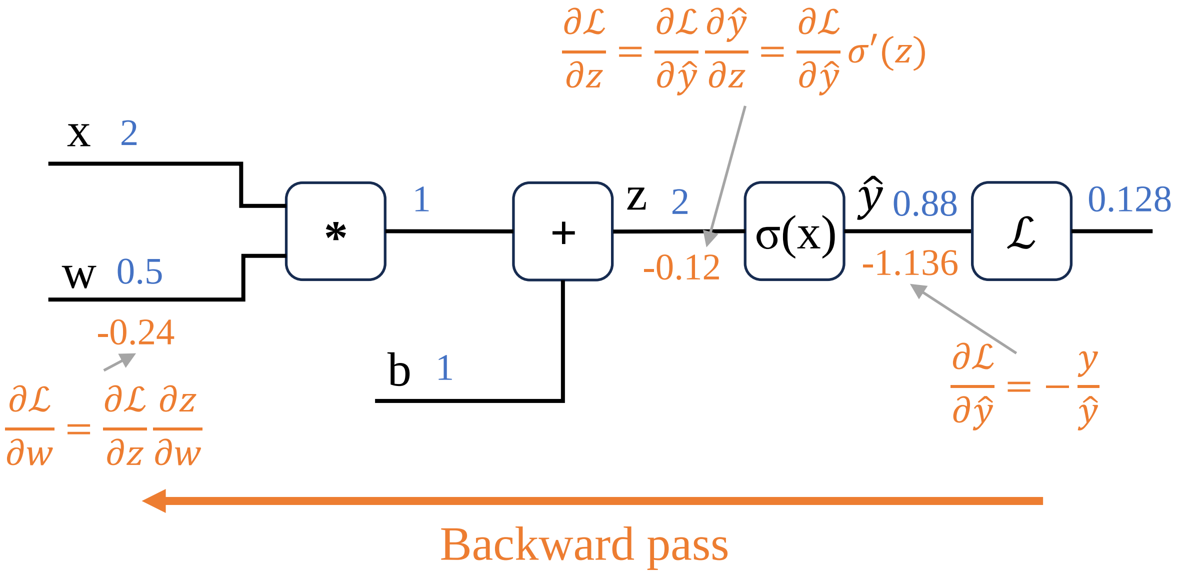
$\mathcal{L} = -y\ln\hat{y} - (1-y)\ln(1-\hat{y})$

$\boldsymbol{w = 0.5, b = 1}$

$\boldsymbol{x = 2, y = 1}$

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

Пример: логистическая регрессия



$z = wx + b$

$\hat{y} = \sigma(z)$

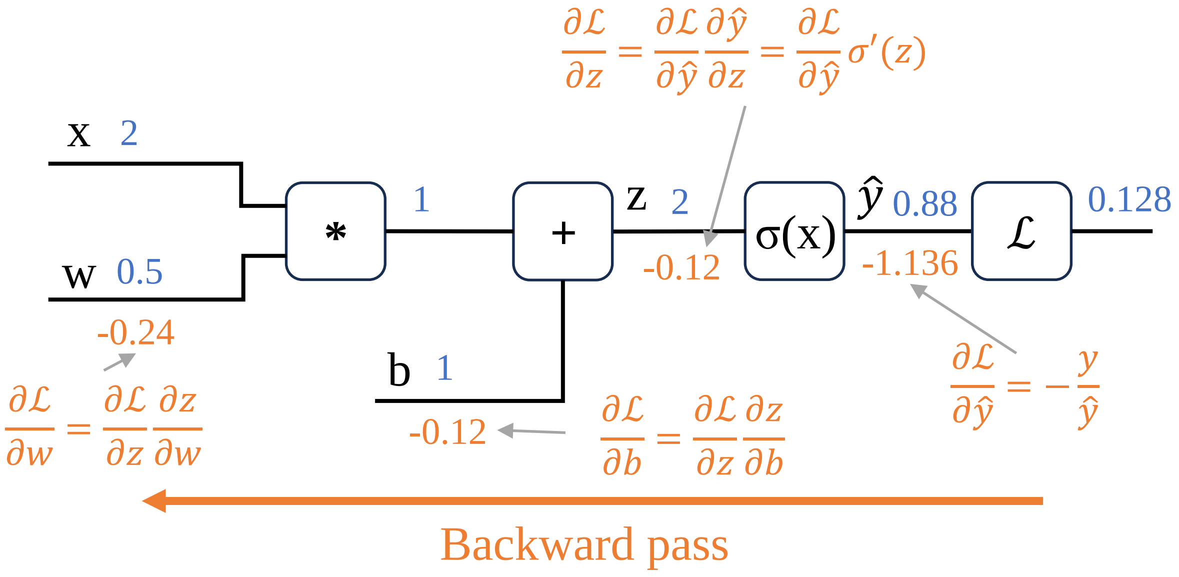
$\mathcal{L} = -y\ln\hat{y} - (1 - y)\ln(1 - \hat{y})$

$\boldsymbol{w = 0.5, b = 1}$

$\boldsymbol{x = 2, y = 1}$

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

Пример: логистическая регрессия



$$z = wx + b$$
$$\hat{y} = \sigma(z)$$
$$\mathcal{L} = -y\ln\hat{y} - (1 - y)\ln(1 - \hat{y})$$
$$\boldsymbol{w = 0.5, b = 1}$$
$$\boldsymbol{x = 2, y = 1}$$

Задача: найти $\frac{\partial \mathcal{L}}{\partial w}$ и $\frac{\partial \mathcal{L}}{\partial b}$

После нахождения градиентов можно переходить к обновлению весов с помощью оптимизатора