

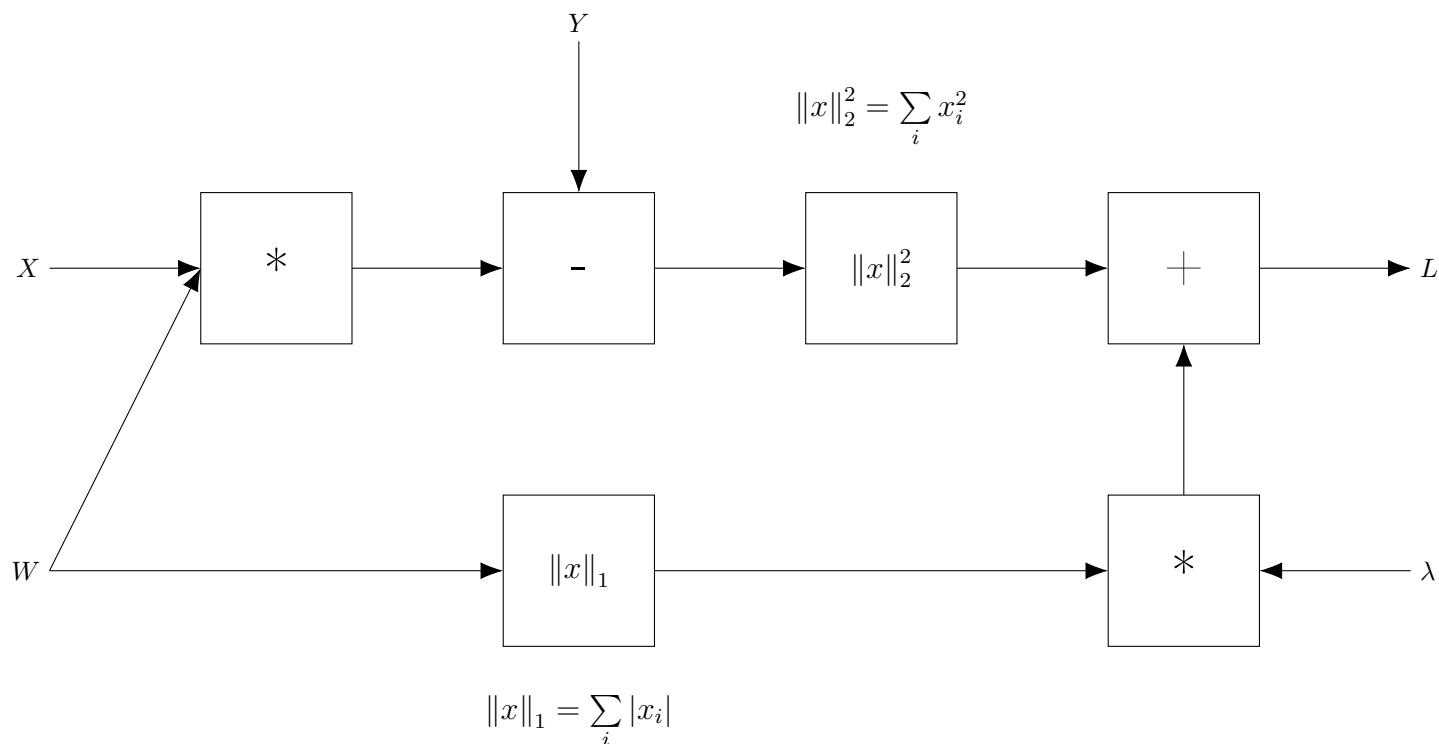
Найти:

$$L = \|X * W - Y\|_2^2 + \lambda * \|W\|_1$$

a) $L = ?$

b) $\nabla_W L = ?$

Схема:



Решение:

$$X = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}; \quad Y = \begin{bmatrix} 0 & -2 \\ 0 & -2 \\ 0 & -2 \end{bmatrix}; \quad W = \begin{bmatrix} 1 & 1 \end{bmatrix}; \quad \lambda = -1$$

a) Найдем L , выполнив все операции для выражения $L = \|X * W - Y\|_2^2 + \lambda * \|W\|_1$ по очереди:

$$1) X * W = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} * \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -2 & -2 \end{bmatrix};$$

$$2) X * W - Y = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -2 & -2 \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & -2 \\ 0 & -2 \end{bmatrix} = \begin{bmatrix} -1 & -3 \\ 1 & -1 \\ -2 & -4 \end{bmatrix};$$

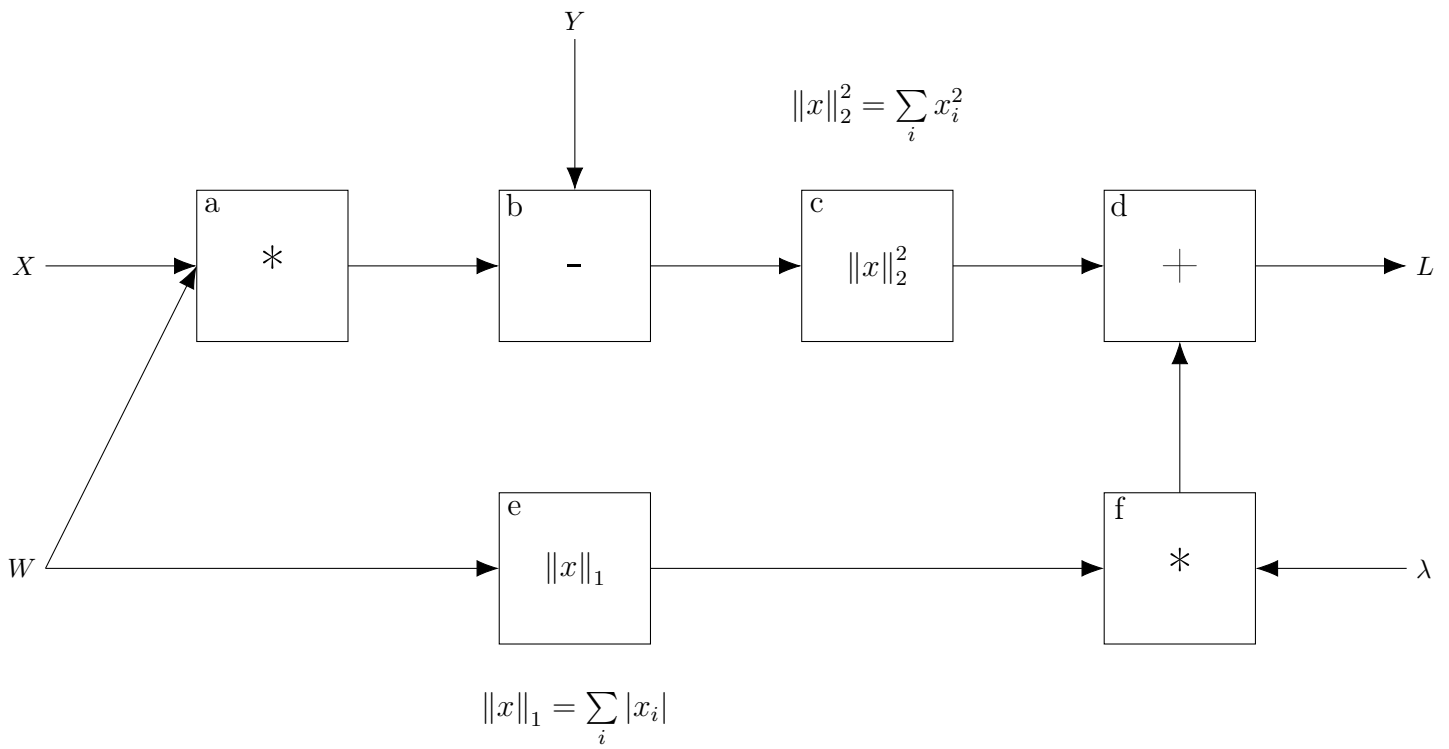
$$3) \|X * W - Y\|_2^2 = \left\| \begin{bmatrix} -1 & -3 \\ 1 & -1 \\ -2 & -4 \end{bmatrix} \right\|_2^2 = (-1)^2 + (-3)^2 + 1^2 + (-1)^2 + (-2)^2 + (-4)^2 = 1 + 9 + 1 + 1 + 4 + 16 = 32;$$

$$4) \|W\|_1 = \left\| \begin{bmatrix} 1 & 1 \end{bmatrix} \right\|_1 = 1 + 1 = 2;$$

$$5) \lambda * \|W\|_1 = -1 * 2 = -2;$$

$$6) \|X * W - Y\|_2^2 + \lambda * \|W\|_1 = 32 + (-2) = 30;$$

b) Чтобы посчитать $\nabla_W L$ обозначим нейроны по буквам (a, b, c, d, e, f):



Найдем градиент $\nabla_W L$ методом обратного распространения ошибки (backpropagation):

Путь через f:

$$\frac{\partial L}{\partial W} = \frac{\partial d}{\partial f} * \frac{\partial f}{\partial e} * \frac{\partial e}{\partial W};$$

- $d = c + f : \quad \frac{\partial d}{\partial f} = 1;$
- $f = \lambda * e : \quad \frac{\partial f}{\partial e} = -1;$
- $e = \|W\|_1 = |W_1| + |W_2| : \quad \frac{\partial e}{\partial W} = \begin{bmatrix} 1 & 1 \end{bmatrix};$
- $\frac{\partial L}{\partial W} = \begin{bmatrix} -1 & -1 \end{bmatrix};$

Путь через c:

$$\frac{\partial L}{\partial W} = \frac{\partial d}{\partial c} * \frac{\partial c}{\partial b} * \frac{\partial b}{\partial a} * \frac{\partial a}{\partial W};$$

- $d = c + f : \quad \frac{\partial d}{\partial c} = 1;$
- $c = \|b\|_2^2 = b_{11}^2 + b_{12}^2 + b_{21}^2 + b_{22}^2 + b_{31}^2 + b_{32}^2 : \quad \frac{\partial c}{\partial b} = \begin{bmatrix} -2 & -6 \\ 2 & -2 \\ -4 & -8 \end{bmatrix};$
- $b = a - Y : \quad \frac{\partial b}{\partial a} = 1;$
- $a = X * W : \quad \frac{\partial a}{\partial W} = \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ -2 & -2 \end{bmatrix};$
- $\frac{\partial L}{\partial W} = \begin{bmatrix} 12 & 20 \end{bmatrix};$

Ответ: а) $L = 30$; б) $\nabla_W L = \begin{bmatrix} -1 & -1 \end{bmatrix} + \begin{bmatrix} 12 & 20 \end{bmatrix} = \begin{bmatrix} 11 & 19 \end{bmatrix}.$