# Problem Set 2 – Shallow and Deep Networks

DS542 – DL4DS

Spring, 2025

**Note:** Refer to the equations in the *Understanding Deep Learning* textbook to solve the following problems.

## Problem 3.2

For each of the four linear regions in Figure 3.3j, indicate which hidden units are inactive and which are active (i.e., which do and do not clip their inputs).

## Problem 3.5

Prove that the following property holds for $\alpha \in \mathbb{R}^+$:

$$\text{ReLU}[\alpha \cdot z] = \alpha \cdot \text{ReLU}[z].$$

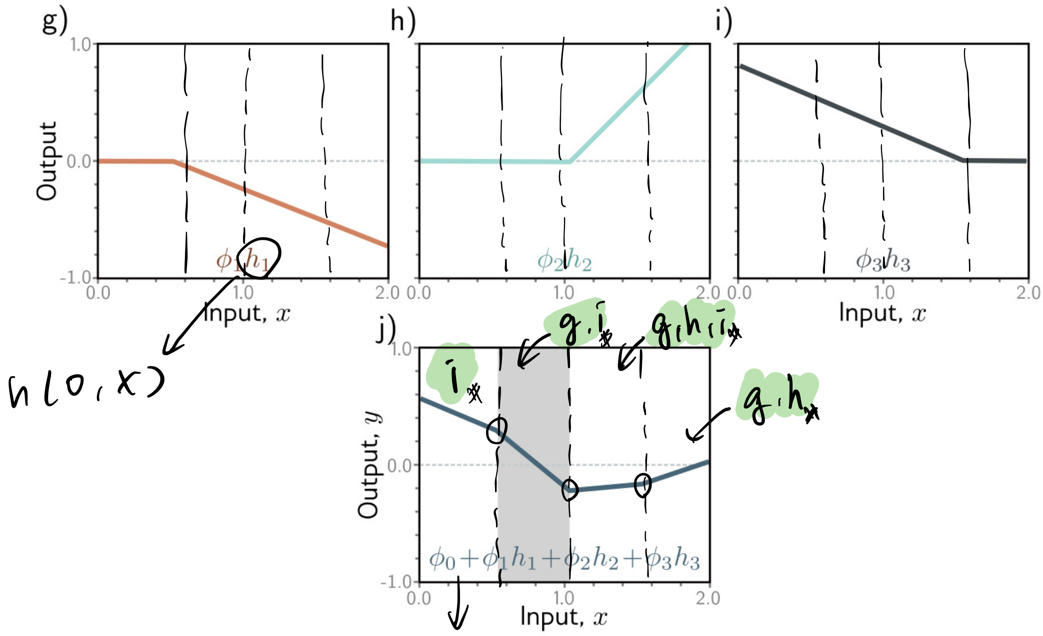
This is known as the non-negative homogeneity property of the ReLU function.

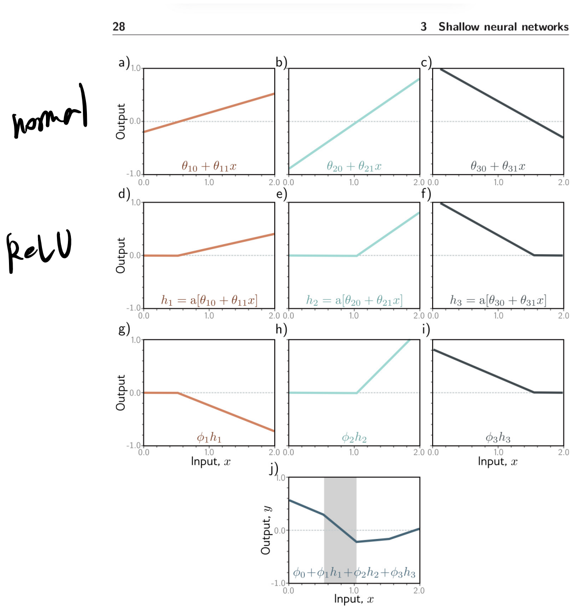
## Problem 4.6

Consider a network with $D_i = 1$ input, $D_o = 1$ output, $K = 10$ layers, and $D = 10$ hidden units in each. Would the number of weights increase more – if we increased the depth by one or the width by one? Provide your reasoning.

# Problem 3.2

For each of the four linear regions in Figure 3.3j, indicate which hidden units are inactive and which are active (i.e., which do and do not clip their inputs).



## Problem 3.5

Prove that the following property holds for $\alpha \in \mathbb{R}^+$:

$$\mathrm{ReLU}[\alpha \cdot z] = \alpha \cdot \mathrm{ReLU}[z].$$

This is known as the non-negative homogeneity property of the ReLU function.

$* \mathrm{ReLu}[\alpha \cdot z] = \max(\alpha \cdot z, 0]$

know that $* \mathrm{ReLU}[z] = \max[z, 0]$

$\downarrow$ 乘 $\alpha$

$\alpha \mathrm{ReLU}(z) = \alpha \max[z, 0]$

$* \quad a \cdot \max(b, 0) = \max(a \cdot b, a \cdot 0)$

$\alpha \max[z, 0] = \max(\alpha z, \alpha \cdot 0)$

therefore $\mathrm{ReLU}[\alpha \cdot z] = \alpha \cdot \mathrm{ReLU}[z]$

## Problem 4.6

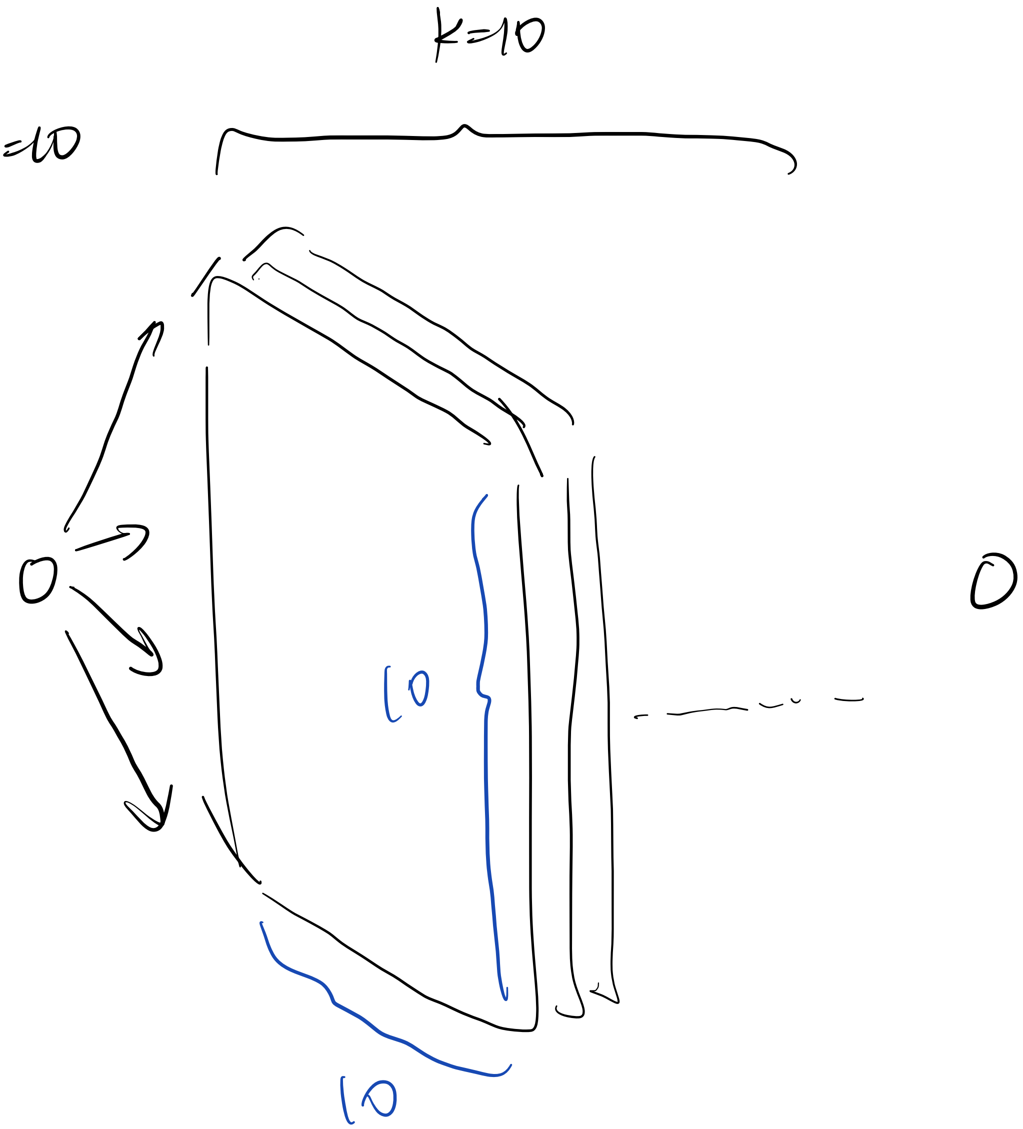
Consider a network with $D_i = 1$ input, $D_o = 1$ output, $K = 10$ layers, and $D = 10$ hidden units in each. Would the number of weights increase more – if we increased the depth by one or the width by one? Provide your reasoning.

Input $D_1 = 1$

output $D_o = 1$

Hidden layer $K = 10$

$D = 10$




initial:

$\underbrace{1 \times 10}_{\text{input}} + 10 \times 10 \times 10 + \underbrace{1 \times 10}_{\text{output}} = 1020$

depth+1 =

$1 \times 10 + 10 \times 10 \times 11 + 1 \times 10 = 1120$

Width + 1 =

$1 \times 11 + 11 \times 11 \times 10 + 11 = 1232$

more than depth

number of weights will increase more while widpth will increase more than depth *