MTH 415 — Homework 1 Solutions

Fall 2026

Problem 1

Let
$$\rho(t) = \mathrm{ReLU}(t) = \max\{t, 0\}.$$

From the network diagram, the first hidden layer is

$$h_1 = \rho(x_0 + 2x_1 + 1), \qquad h_2 = \rho(-x_0 + x_1).$$

The second hidden layer is

$$z_1 = \rho(h_1 + h_2), \qquad z_2 = \rho(-h_1 + h_2 - 1).$$

The output layer is linear, so
$$f(x_0, x_1) = 2z_1 - z_2.$$

Therefore,
$$\boxed{\begin{aligned} f(x_0, x_1) = 2\rho\left(\rho(x_0 + 2x_1 + 1) + \rho(-x_0 + x_1)\right) \\ - \rho\left(-\rho(x_0 + 2x_1 + 1) + \rho(-x_0 + x_1) - 1\right). \end{aligned}}$$

Now evaluate at $(x_0, x_1) = (1, -1)$. We have

$$h_1 = \rho(1 + 2(-1) + 1) = \rho(0) = 0,$$

and
$$h_2 = \rho(-1 - 1) = \rho(-2) = 0.$$

Thus
$$z_1 = \rho(0) = 0, \qquad z_2 = \rho(-1) = 0,$$

and hence
$$\boxed{f(1, -1) = 0.}$$

Problem 2

We are given
$$f(x) = \sum_{j=1}^{3} c_j \,\mathrm{ReLU}(w_j x + b_j),$$

where
$$c_j \neq 0, \qquad w_j \neq 0, \qquad j = 1, 2, 3.$$

For each j, define

$$t_j = -\frac{b_j}{w_j}.$$

The jth ReLU term can change slope only when

$$w_j x + b_j = 0,$$

so the possible kink locations are

$$\boxed{t_1 = -\frac{b_1}{w_1}, \qquad t_2 = -\frac{b_2}{w_2}, \qquad t_3 = -\frac{b_3}{w_3}.}$$

However, if two or more of these locations coincide, their changes in slope may cancel.

For the jth term

$$c_j \,\mathrm{ReLU}(w_j x + b_j),$$

the change in slope when crossing t_j from left to right is

$$c_j |w_j|.$$

Indeed, if $w_j > 0$, the slope changes from 0 to $c_j w_j$, while if $w_j < 0$, the slope changes from $c_j w_j$ to 0. In both cases the change in slope is

$$c_j |w_j|.$$

Define

$$a_j = c_j |w_j|.$$

Since $c_j \neq 0$ and $w_j \neq 0$, we have

$$a_j \neq 0.$$

Therefore, at any point t, the total change in slope is

$$\sum_{\{j : t_j = t\}} a_j.$$

Hence t is a genuine kink if and only if

$$\boxed{\sum_{\{j : t_j = t\}} c_j |w_j| \neq 0.}$$

We now classify all possibilities.

No kink

Since a single ReLU kink cannot disappear by itself, all three kink locations must coincide:

$$t_1 = t_2 = t_3.$$

They must also cancel:

$$a_1 + a_2 + a_3 = 0.$$

Thus f has no kink exactly when

$$\boxed{t_1 = t_2 = t_3, \qquad a_1 + a_2 + a_3 = 0.}$$

Equivalently,

$$\boxed{-\frac{b_1}{w_1} = -\frac{b_2}{w_2} = -\frac{b_3}{w_3}, \qquad c_1|w_1| + c_2|w_2| + c_3|w_3| = 0.}$$

Exactly one kink

There are two possibilities.

Case 1: All three locations coincide.

If

$$t_1 = t_2 = t_3$$

but the changes in slope do not cancel, then there is exactly one kink:

$$\boxed{t_1 = t_2 = t_3, \qquad a_1 + a_2 + a_3 \neq 0.}$$

Case 2: Exactly two locations coincide and cancel.

For some distinct indices i, j, k,

$$t_i = t_j \neq t_k.$$

The two terms at the common location must cancel:

$$a_i + a_j = 0.$$

Thus

$$\boxed{t_i = t_j \neq t_k, \qquad a_i + a_j = 0.}$$

Explicitly, the possibilities are

$$\begin{aligned} t_1 = t_2 \neq t_3, &\quad a_1 + a_2 = 0, \\ t_1 = t_3 \neq t_2, &\quad a_1 + a_3 = 0, \\ t_2 = t_3 \neq t_1, &\quad a_2 + a_3 = 0. \end{aligned}$$

Exactly two kinks

There must be exactly two distinct candidate kink locations. Thus two of the t_j's coincide, while the third is different.

The two changes in slope at the repeated location must not cancel. Therefore, for some distinct i, j, k,

$$\boxed{t_i = t_j \neq t_k, \qquad a_i + a_j \neq 0.}$$

Explicitly,

$$\begin{aligned} t_1 = t_2 \neq t_3, &\quad a_1 + a_2 \neq 0, \\ t_1 = t_3 \neq t_2, &\quad a_1 + a_3 \neq 0, \\ t_2 = t_3 \neq t_1, &\quad a_2 + a_3 \neq 0. \end{aligned}$$

Exactly three kinks

All three candidate kink locations must be distinct:

$$\boxed{t_1,\ t_2,\ t_3 \text{ are pairwise distinct.}}$$

Since $a_j \neq 0$, none of the three individual kinks can disappear.

Summary

The complete classification is

Number of kinks	Condition
0	$t_1 = t_2 = t_3, \quad a_1 + a_2 + a_3 = 0$
1	$t_1 = t_2 = t_3, \quad a_1 + a_2 + a_3 \neq 0,$ or $t_i = t_j \neq t_k, \quad a_i + a_j = 0$
2	$t_i = t_j \neq t_k, \quad a_i + a_j \neq 0$
3	t_1, t_2, t_3 pairwise distinct.

Problem 3

(a) Network representing f

The function is

$$f(x) = \begin{cases} 0, & x \leq -1, \\ x + 1, & -1 < x \leq 0, \\ 1 - x, & 0 < x \leq 1, \\ 0, & x > 1. \end{cases}$$

It can be written as

$$\boxed{f(x) = \mathrm{ReLU}(x+1) - 2\,\mathrm{ReLU}(x) + \mathrm{ReLU}(x-1).}$$

Thus we may use one hidden layer with three ReLU neurons:

$$h_1 = \mathrm{ReLU}(x+1), \qquad h_2 = \mathrm{ReLU}(x), \qquad h_3 = \mathrm{ReLU}(x-1),$$

followed by the linear output

$$f = h_1 - 2h_2 + h_3.$$

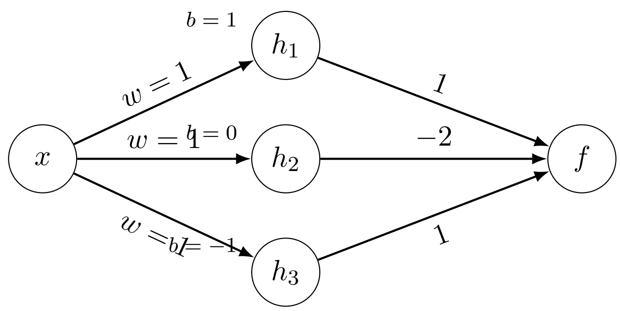
The weights and biases are

$$W_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad b_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

and

$$W_2 = \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}, \qquad b_2 = 0.$$

A corresponding network is



Here the three hidden neurons use ReLU activation and the output neuron uses the identity activation.

(b) Network representing g

Let

$$\sigma(t) = \frac{1}{1 + e^{-t}}$$

be the sigmoid activation.

The given function is

$$g(x_0, x_1) = \frac{1}{1 + \exp\left(-\left[\frac{1}{1+e^{-x_0}} + \frac{1}{1+e^{-x_1}} - 1\right]\right)}.$$

Using σ, this becomes

$$\boxed{g(x_0, x_1) = \sigma\left(\sigma(x_0) + \sigma(x_1) - 1\right).}$$

Thus the first hidden layer is

$$h_1 = \sigma(x_0), \qquad h_2 = \sigma(x_1),$$

and the output is

$$g = \sigma(h_1 + h_2 - 1).$$

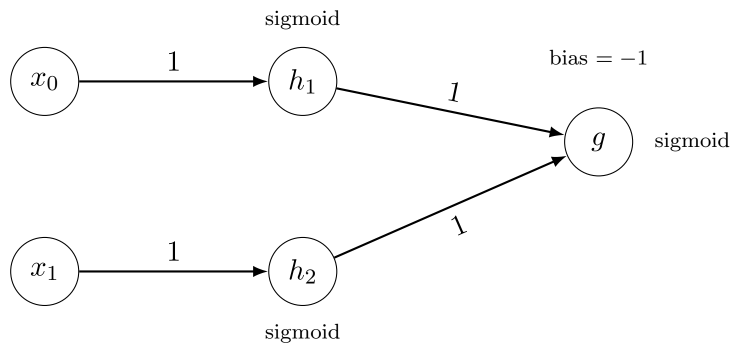
The weights and biases are

$$W_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad b_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

and

$$W_2 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \qquad b_2 = -1.$$

A corresponding network is



(c) Network representing h

Define the step activation

$$S(t) = \begin{cases} 1, & t > 0, \\ 0, & t \le 0. \end{cases}$$

We want

$$h(x_0, x_1) = \begin{cases} 1, & 0 < x_0 < 1 \quad \text{or} \quad -1 < x_1 < 2, \\ 0, & \text{otherwise.} \end{cases}$$

The first layer checks the four inequalities

$$a_1 = S(x_0), \qquad a_2 = S(1 - x_0),$$

and

$$a_3 = S(x_1 + 1), \qquad a_4 = S(2 - x_1).$$

Thus

$$W_1 = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}, \qquad b_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

Notice that

$$0 < x_0 < 1$$

holds exactly when

$$a_1 = a_2 = 1.$$

Therefore define

$$r_1 = S\left(a_1 + a_2 - \frac{3}{2}\right).$$

Similarly,

$$-1 < x_1 < 2$$

holds exactly when

$$a_3 = a_4 = 1,$$

so define

$$r_2 = S\left(a_3 + a_4 - \frac{3}{2}\right).$$

Hence

$$W_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} -\frac{3}{2} \\ -\frac{3}{2} \end{pmatrix}.$$

Finally, we take the logical OR of r_1 and r_2:

$$h = S\left(r_1 + r_2 - \frac{1}{2}\right).$$

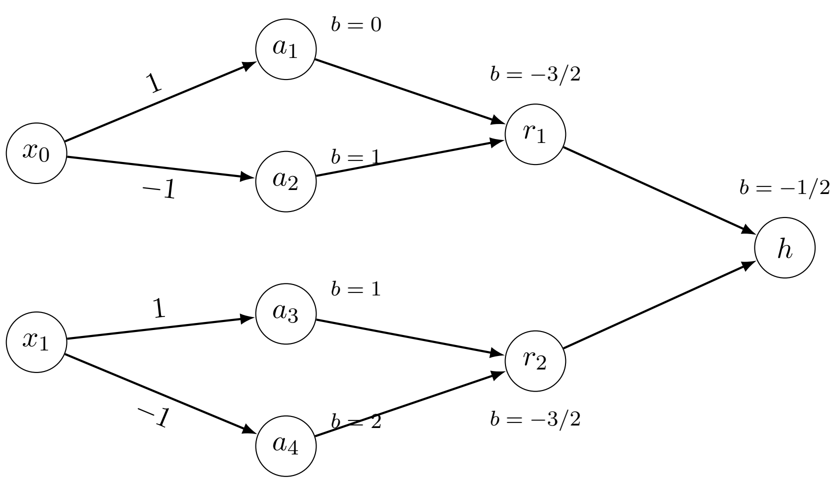
Thus

$$W_3 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \qquad b_3 = -\frac{1}{2}.$$

The complete network can therefore be written as

$$\boxed{\begin{aligned} a_1 &= S(x_0), & a_2 &= S(1 - x_0), \\ a_3 &= S(x_1 + 1), & a_4 &= S(2 - x_1), \\ r_1 &= S\left(a_1 + a_2 - \frac{3}{2}\right), & r_2 &= S\left(a_3 + a_4 - \frac{3}{2}\right), \\ h &= S\left(r_1 + r_2 - \frac{1}{2}\right). \end{aligned}}$$

A corresponding network is



All neurons in this network use the step activation S.