#Author: Jerry Wang, written in R markdown file

Given the piecewise equation:

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 2x - 1 & \text{if } x \geq 1 \end{cases} \tag{1}$$

#Step 1: try to express each segment using ReLU independently:

for x<0, this is just 0, and can be represented as

$$0 \times ReLU(x)$$

for 0<=x<1, x can be represented as

$$ReLU(x)$$

for x>=1, f(x) = 2x-1 can be written as x + (x-1), and then be represented as

$$ReLU(x) + ReLU(x-1)$$

I have found it very lucky that the last ReLU function already captures all complexity of this piecewise equation. ReLU(x-1) segment only activates when x>=1. So if we have x = 0.5 for instance, this function will be automatically converted to ReLU(x). If x<0, neither segment of this function activates, turning straight to 0. As such, f(x) can be written in the form:

$$f(x) = ReLU(x) + ReLU(x-1)$$

test different scenarios:

$$x = -1 \rightarrow f(x) = \text{ReLU}(-1) + \text{ReLU}(-1-1) = 0 + 0 = 0$$

$$x = 0.4 \rightarrow f(x) = \text{ReLU}(0.4) + \text{ReLU}(0.4-1) = 0.4 + 0 = 0.4$$

$$x = 1.1 \rightarrow f(x) = \text{ReLU}(1.1) + \text{ReLU}(1.1-1) = 1.1 + 0.1 = 1.2$$

$$x = 3 \rightarrow f(X) = \text{ReLU}(3) + \text{ReLU}(3-1) = 3 + 2 = 5$$

my ReLU representation works for f(x)

#Step 2: convert f(x) = ReLU (x) + ReLU(x-1) to form:

$$f(x) = M^{(2)}\text{ReLU}(M^{(1)}x + B^{(1)}) + B^{(2)}$$

where $M^{(1)}$ and $B^{(1)}$ are 2x1 matrices; $M^{(2)}$ is a 1x2 matrix and $B^{(2)}$ is a scalar. Since we have a combination of two ReLU segments: ReLU(x) and ReLU(x-1), we may represent x as a 2x1 matrix and output f(x) as a 1x1 matrix.

define z as the 2x1 matrix formed by stacking x and x-1:

$$z = \begin{bmatrix} x + 0 \\ x - 1 \end{bmatrix}$$

The matrix $M^{(1)}$ is therefore a 2x1 matrix which, when multiplied by x (a 1x1 matrix), should yield the x values in the 2x1 matrix z above. In addition, the matrix $B^{(1)}$ is a 2x1 matrix that yields the bias values in the 2x1 matrix z. We can easily find that:

$$M^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, B^{(1)} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

So,

$$z = M^{(1)}x + B^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

now apply ReLU to z element wise:

$$\text{ReLU}(z) = \begin{bmatrix} \text{ReLU}(x) \\ \text{ReLU}(x-1) \end{bmatrix}$$

Recall from step 1 that ReLU(x) + ReLU(x-1) yields f(x). At this point it is clear that we just need a 1x2 matrix $M^{(2)}$ of [1 1] and a zero scalar $B^{(2)}$, since there is no further bias added:

$$M^{(2)} = \begin{bmatrix} 1 \ 1 \end{bmatrix}, B^{(2)} = \begin{bmatrix} 0 \end{bmatrix}$$

Final form:

$$f(x) = \begin{bmatrix} 1 \ 1 \end{bmatrix} \cdot \begin{bmatrix} \text{ReLU}(x) \\ \text{ReLU}(x-1) \end{bmatrix} + 0$$

$$f(x) = \begin{bmatrix} 1 \ 1 \end{bmatrix} \cdot \text{ReLU}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ -1 \end{bmatrix} \right) + 0$$