2. Equivalent network

Network 1 is a 6 input and 6 output MLP with multiple hidden layers. Network 2 is a 6 input and 6 output MLP without multiple hidden layers

Assuming a linear activation function, the output values in Network 1 are given by:

$$\vec{a}^{(1)} = W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)} \qquad (1)$$

$$\vec{a}(2) = W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)} \qquad (2)$$

$$\vec{a}^{(3)} = W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)} \qquad (3)$$

The output values in Network 2 are given by:

$$\tilde{\mathrm{a}} = \widetilde{W}\vec{a}^{(0)} + \tilde{b} \qquad (4)$$

Two networks are said to be equivalent if, they have the same number of input nodes and output nodes, and for all inputs, the output of both networks is identical. Given Network 1's weights and bias values, the output and input relationship are expressed in equation (5).

$$\vec{a}^{(3)} = W^{(3)}(W^{(2)}(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}) + \vec{b}^{(2)}) + \vec{b}^{(3)} \qquad (5)$$

Where $\vec{a}^{(0)}$ is the input values and $\vec{a}^{(3)}$ is the output values of Network 1.

For Network 1 and Network 2 to be equivalent, the outputs have to be the same with the same inputs. Therefore the relationship of Network 1 and 2 would be expressed as :

$$\widetilde{W}\vec{a}^{(0)} + \tilde{b} = W^{(3)}(W^{(2)}(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}) + \vec{b}^{(2)}) + \vec{b}^{(3)}$$

$$\widetilde{W}\vec{a}^{(0)} + \tilde{b} = W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)} \qquad (6)$$

As expressed in Equation 6, $\widetilde{W}$ and $\tilde{b}$ in Network 2 would be expressed as:

$$\widetilde{W} = W^{(3)}W^{(2)}W^{(1)}$$

$$\tilde{b} = W^{(3)}W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$