Question 2

For Network 1:

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

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

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

Bring equation (1), (2) into equation (3)

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

For Network 2:

$\vec{a}^{(3)} = W \quad \vec{a}^{(0)} + B$

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

The two networks are equivalent.