2.4.

If the first Layer is $\tilde{z}^{(1)} = W^{(1)}x + b^{(1)}$ for $\tilde{z}^{(0)} = x$

than the second layer is:

$$\tilde{z}^{(2)} = W^{(2)}(W^{(1)}x + b^{(1)}) + b^2$$

$$= \underbrace{W^{(2)}W^{(1)}}_{W^*}x + \underbrace{W^{(2)}b^{(1)} + b^2}_{b^* \text{ (some bias)}}$$

which has the same form as the first layer

$$W^*x + b^*$$

If all functions in the network are linear, the whole network collapses into a single linear model and a linear model cannot separate the XOR-problem